Short-Term Forecasting of Precipitation Using Satellite Data

Abstract

Short-range forecasting of rain or snow, known as precipitation nowcasting, is typically displayed on geographical maps by weather services for up to a 2-hour timeframe. Current methods for precipitation nowcasting predominantly use the extrapolation of ground-based radar observations, employing techniques like optical flow or neural networks. However, the effectiveness of these methods is geographically restricted to areas surrounding radar installations. This paper introduces a novel precipitation nowcasting technique that utilizes geostationary satellite imagery. This method has been integrated into the Yandex.Weather precipitation map, which includes an alert system with push notifications for Yandex ecosystem products. The integration of satellite imagery significantly broadens the coverage area, marking a step towards developing a comprehensive global nowcasting service.

1 Introduction

Weather conditions significantly impact the daily routines and planning of urban populations. Similar to how ancient humans relied on environmental cues for hunting, modern individuals adjust their daily and leisure activities based on the likelihood of rain or cloud cover. Weather forecasting services provide essential data, including temperature, precipitation intensity and type, cloudiness, humidity, pressure, and wind conditions. These services offer current weather updates, short-term predictions up to 2 hours (nowcasting), medium-range forecasts up to 10 days, and long-range predictions spanning several months.

A crucial component of weather services is the precipitation map, which combines radar data with neural network-based, very short-term precipitation forecasting to deliver a detailed map of anticipated precipitation for the next two hours, updated every 10 minutes. This feature enables personalized, user-friendly notifications, such as alerts about impending rain. The popularity of this feature is evident, as it significantly influences user engagement and reliance on weather services.

The information from the precipitation map is used to refine current weather condition reports (e.g., sunny, cloudy, rainy) on the main weather website. Additionally, partners and offline users, including radio and television, depend on this data, effectively doubling the audience for the precipitation nowcasting product.

Traditional weather forecasting, which involves numerical modeling of the atmosphere, cannot accurately predict exact rain locations on short time scales. For instance, it struggles to determine which part of a city will be affected by rain within the next hour. Moreover, traditional methods provide hourly updates, making it difficult to pinpoint brief periods without rain during short, intense precipitation events. People often need straightforward answers to simple questions like when it will rain or stop raining, requiring specific predictions such as "heavy rain will start in 10 minutes and last for 30 minutes."

Conventional numerical weather prediction (NWP) models are limited in their ability to forecast precipitation events at specific locations and times. Radar extrapolation products are effective for the first couple of hours but fail to predict precipitation accurately due to physical processes. Consequently, the current trend in nowcasting is to merge high-resolution radar data with traditional NWP models.

However, radar-based products are limited by the location of radar installations and are not easily scalable. Radars are costly, their installation requires governmental and public approval, and their operation needs trained personnel. Coverage is particularly sparse in large, unevenly populated countries like Russia, where many remote areas lack the necessary infrastructure. Similar challenges exist in many developing countries that need weather services but lack the infrastructure for radar networks.

The objective of this research is to develop and implement a practical system for precipitation nowcasting that relies on satellite imagery and NWP products. The goal is to replicate the precipitation fields obtained from radar using satellite data and then to provide nowcasting over a much larger area using a similar predictive model. The system's effectiveness is validated by comparing predicted precipitation with data from ground-based weather stations. The primary focus areas with limited radar coverage are the Siberian and Ural federal districts of Russia, which have a combined population of approximately 30 million.

2 Related Work

This section provides an overview of related work, divided into two main parts corresponding to the primary components of our pipeline.

2.1 Precipitation detection

The global and continuous coverage offered by geostationary satellite imagery makes it a highly desirable data source for precipitation nowcasting algorithms. Since satellites do not directly observe rainfall, precipitation data must be extracted using heuristic or machine learning methods. This extraction can be framed as either precipitation estimation (regression) or precipitation detection (binary classification). This paper concentrates on the binary classification approach to precipitation detection.

The interaction of light with the atmosphere, specifically absorption and scattering, is governed by established physical principles. These principles can be used to develop heuristics for detecting precipitation. One such implementation is the multi-sensor precipitation estimate (MPE), which is, however, limited to detecting convective rain and may produce inaccurate results in areas with other forms of precipitation. This limitation is particularly significant in middle and high latitudes, where convective precipitation is predominantly a summer phenomenon resulting from surface heating, leading to the formation of cumulonimbus clouds and heavy rainfall. During much of the year, frontal precipitation, driven by cyclonic movements and interactions between warm and cold fronts, is more common. The MPE algorithm often fails to capture these frontal precipitation events.

A more advanced physics-based heuristic is the precipitation properties (PP) algorithm, which integrates NWP model data, cloud physical properties, and satellite measurements. This algorithm uses radar observations to calibrate its parameters. However, because it relies on satellite observations at visible wavelengths to determine cloud properties, it can only retrieve precipitation data during daylight hours.

Machine learning techniques, including decision trees, neural networks, and SVMs, have been evaluated for precipitation detection. However, these studies often used pixel-wise data splits for training and testing, which may lead to overfitting due to neglecting the spatial and temporal smoothness of atmospheric phenomena. While these studies examined day, twilight, and night conditions separately, with the best results during the day, a more sophisticated method using a fully-connected stacked denoising autoencoder has also been applied to precipitation detection. Although the autoencoder's unsupervised training helps mitigate overfitting, there is no comparison with other architectures.

From a machine learning perspective, precipitation detection is similar to semantic segmentation, where a multichannel image is input, and each pixel is assigned an output label. Convolutional neural networks have become the standard for semantic segmentation in recent years, making them a natural choice for precipitation detection as well.

Convolutional neural networks have been effectively used in various satellite image processing tasks, such as road and building detection. Despite numerous public challenges that have advanced the field, the range of architectures used for aerial image processing remains narrower compared to those used for semantic segmentation datasets like Microsoft COCO or Cityscapes. A common issue in these datasets is the presence of objects of the same class at different scales, which has led to the development of multiscale approaches. However, these approaches are less applicable to precipitation detection and other satellite imagery tasks, as the distance between the sensor and the Earth's surface is usually known. Consequently, simpler models like UNet and fully-convolutional ResNet remain relevant.

2.2 Nowcasting

Precipitation nowcasting is typically accomplished in two stages by extrapolating radar observations. Initially, wind patterns are estimated by comparing multiple precipitation fields captured by radar. The techniques used for this in meteorology are similar to optical flow estimation algorithms in computer vision. Subsequently, the precipitation field is moved according to the estimated wind directions.

A novel approach to nowcasting using a convolutional recurrent neural network (Conv-LSTM) was introduced and later refined. While this neural network adds complexity, it can theoretically improve rainfall prediction accuracy by accounting for radar artifacts and the appearance or disappearance of precipitation areas. However, the most significant of these processes, the vanishing of precipitation, can also be managed by adding basic filtering to the optical flow method.

3 Methodology

This section details the methodology used for precipitation detection and nowcasting, focusing on data preprocessing, model training, and evaluation metrics.

3.1 Data Sources

Precipitation nowcasting imposes distinct data requirements compared to Numerical Weather Prediction (NWP), including high spatial and temporal resolution, direct rainfall measurement, and global coverage. Since no single source can fulfill all these requirements, it is necessary to combine multiple data sources.

Weather stations provide direct precipitation observations, typically measuring accumulated precipitation every 12 hours according to the SYNOP protocol. Although many stations report more frequently, usually every 3 hours, this frequency is insufficient for nowcasting due to the lack of detailed spatial and temporal data needed to generate high-resolution precipitation fields.

Radar observations are the primary source of high-resolution precipitation data. The Russian network of DMRL-C radars, operated by Roshydromet, uses C-band Doppler technology to measure raindrop reflectivity and radial velocity. Each radar covers a circular area with a radius of up to 250 km and 10 km above the ground, with accuracy diminishing with distance. The radar echo can be converted to surface precipitation using the Marshall-Palmer relation. The resulting precipitation field has a resolution of 2 x 2 km, with scans repeated every ten minutes. However, radar coverage is limited, especially outside densely populated areas of Europe and North America, with most Russian radars located in the western part of the country.

Low Earth orbit satellites equipped with radars and sensors provide another source of precipitation measurements. These satellites scan a narrow band beneath their orbital path, offering global coverage in the sense that every location within a certain latitude range is eventually scanned. However, the time between consecutive passes of a single satellite can be quite long. The Global Precipitation Measurements (GPM) mission, operated by NASA and JAXA, uses a constellation of about 10 operational satellites to provide global precipitation coverage from 65°S to 65°N with a 3-hour temporal resolution.

Geostationary satellites are widely used for weather observation. Positioned 35,786 km above the equator, these satellites match the Earth's rotation, allowing continuous monitoring of a large area. However, at such altitudes, the only feasible instrument for cloud and precipitation detection is a high-resolution imager that captures visible and infrared spectrum snapshots. Accurately detecting precipitation from these images is challenging. Previous studies on this topic have not achieved the accuracy needed for user-facing products that aim to alert users about precipitation within 10 minutes.

This study uses data from the Meteosat-8 satellite, operated by EUMETSAT, positioned over the Indian Ocean at 41.5° longitude, covering the western part of Russia and Europe. The SEVIRI instrument on Meteosat-8 scans the Earth's surface in 12 channels, with a spatial resolution of 3 km per pixel and a full scan time of 15 minutes.

This paper describes a precipitation nowcasting system that integrates radar, satellite, and NWP model data. A new approach to precipitation detection is introduced and its accuracy is demonstrated.

3.2 Precipitation Detection

The approach to precipitation detection is summarized in Table 1. The key components of the pipeline are described in detail in the following subsections.

Table 1: Summary of our precipitation detection approach.

Input features	Satellite imagery, GFS fields, solar altitude, topography
Ground truth	Binarized radar measurements
Model	UNet
Loss function	Binary crossentropy + Dice loss
Evaluation measure	F1 score

3.2.1 Preprocessing

The data preparation process involves several steps aimed at minimizing the discrepancies between different data domains.

Radar data preprocessing begins by discarding radar observations taken beyond 200 km from the radar, as these are deemed unreliable. Subsequently, observations from various radars are consolidated onto a single map, resolving any conflicts between radars with overlapping coverage areas. Due to frequent false negatives in radar observations, the maximum value between two data points is used for aggregation. Finally, radar observations are binarized using three thresholds: 0.08 mm/h for light rain, 0.5 mm/h for moderate rain, and 2.5 mm/h for heavy rain.

Satellite images and radar observations are remapped onto a uniform grid using an equirectangular projection. Given the oblique observation angles and the fact that precipitation can occur up to 2 km above the ground, there can be a parallax shift of up to 3 pixels between radar and satellite data. However, in practice, accurately estimating precipitation height is complex, and accounting for parallax did not improve the alignment.

Satellite and radar data have different observation frequencies: satellite images are available every 15 minutes, while radar images are available every 10 minutes. To align these data sources temporally, a frame rate conversion is implemented using optical flow interpolation. The goal is to match the radar data's temporal resolution, so satellite data is converted to a 10-minute time

step. However, optical flow cannot be directly computed from satellite imagery due to the presence of both transient atmospheric phenomena and the permanent underlying relief. This issue is circumvented by performing precipitation detection before the optical flow step, allowing the optical flow to be computed directly from the precipitation detection results, which do not include the relief.

To generate the missing image I_t between two adjacent anchor images taken at times t_0 and t_1, the following equation is used:

$$I_t(r) = aI_{t_0}(r + bu_{01}) + bI_{t_1}(r + au_{10})$$

where $a = \frac{t_1 - t}{t_1 - t_0}$ and $b = \frac{t - t_0}{t_1 - t_0}$ are coefficients dependent on the time of the generated image, and u_{01} and u_{10} are the forward and backward optical flows, computed using the TV-L1 optical flow algorithm implemented in OpenCV.

Roshydromet radars record the timestamp at the end of a scan, whereas EUMETSAT marks the start. Since the Earth is scanned in a series of lateral sweeps starting from the south, the actual observation time varies with latitude, with northern latitudes observed last. The combined discrepancy between timestamps can reach 20 minutes. Experimental validation has confirmed that this value corresponds to the minimum discrepancy between radar data and precipitation field reconstruction.

Additional features are incorporated into the satellite imagery to enhance the signal. The Global Forecast System (GFS) model is used to provide a comprehensive description of atmospheric conditions, including physical properties not easily inferred from satellite imagery. The GFS model produces forecasts four times a day with a spatial resolution of 0.25° x 0.25° and temporal intervals of 3 hours. Key fields from GFS include convective precipitation rate, cloud work function, cloud water, precipitable water, and convective potential energy at different levels. Additionally, a topography map and solar altitude data are included as features.

3.2.2 Training

A modified UNet architecture is employed as the primary model for precipitation detection. Through testing, it was determined that using 5 upsample/downsample blocks, compared to the original 4, yields the best results on the validation dataset. The model utilizes standard 3x3 convolutions, 2x2 pooling, and batch normalization layers. The number of channels begins at 16 in the first block and doubles with each downsampling step. This reduced number of channels helps mitigate overfitting and accelerates training and evaluation.

The network is trained for 250,000 iterations using the Adam algorithm, with an initial learning rate of 10^{-4}, which is reduced by a factor of 10 after 200,000 iterations. The addition of the Dice loss to the standard binary cross-entropy improves the F1 scores for the converged model. Training is performed using the Keras framework with a TensorFlow backend and Horovod for multi-GPU learning.

The model is trained to detect three levels of precipitation (light, medium, and heavy) simultaneously, producing three output maps with binary classification loss applied to each map independently.

Typically, precipitation estimation algorithms are developed separately for day, twilight, and night conditions. However, this separation is challenging for machine learning in high-latitude zones due to the underrepresentation of night during summer and day during winter, making it difficult to compile a balanced dataset. Therefore, a single model is trained, with solar altitude provided as an additional input feature.

Overfitting is a significant concern due to the limited geographical area of the dataset. The network can easily memorize the relief, which is visible in some wavelengths even if not explicitly provided as a feature, and use it to overfit on ground truth labels within the radar coverage areas. Moreover, memorizing the correspondence between geographical location and output labels may cause the model to ignore areas outside radar coverage, leading to constant output in these regions. This contradicts the goal of extending nowcasting beyond radar coverage. To address this, the model is trained on relatively small data crops (96x96 pixels).

Due to the large number of channels in the input data, which is atypical for computer vision problems, data loading can be slow. To manage this, a small batch of 5 multi-channel images (including all additional features) is loaded, and each image is then cropped 10 times at random locations.

3.2.3 Metrics

This section presents the evaluation metrics for the precipitation detection algorithm. Due to class imbalance, standard classification accuracy is not informative. Therefore, the primary metric used is the F1 score, averaged across temporal and spatial dimensions.

Several approaches are compared:

- **UNet with GFS**: The UNet architecture with a complete set of features, trained as described earlier. - **UNet w/o GFS**: The same UNet approach without GFS features. - **Pointwise**: A neural network with two convolutional layers using 1x1 convolutions, equivalent to a pointwise perceptron model. GFS features are not used in this model. - **PP and MPE**: Physics-based algorithms (Precipitation Properties and Multi-sensor Precipitation Estimate).

Given that PP and MPE algorithms are designed for daylight conditions, the metrics are also averaged separately for day, night, and twilight periods. The neural network approaches consistently outperform the physics-based methods across all time periods and metrics. The generally poor performance of PP and MPE in these experiments may be due to their tuning for predicting convective rainfall aggregated over extended periods, which does not align with the requirements of this service.

The pointwise model's performance falls between that of UNet and the physics-based approaches. Since it is trained on radar data, it detects similar types of precipitation and performs well during testing.

The UNet architecture's superiority over the pointwise model likely stems from its ability to gather information from a large receptive field. While precipitation reconstruction does not require the same extent of multiscale data processing as many semantic segmentation tasks, the interconnectedness of adjacent atmospheric locations makes a large receptive field beneficial for precipitation detection.

Finally, the addition of GFS features further enhances the F1 score of the UNet model, as demonstrated in the results.

4   Experiments

4.1   Nowcasting

Upon completing the reconstruction of the precipitation field in the area of interest, a separate algorithm is employed to forecast future precipitation fields based on several consecutive reconstructed fields. Two options are considered for this algorithm: extrapolation with optical flow, as used for frame rate conversion, and a convolutional neural network previously developed for radar data prediction. The network consists of a sequence of blocks, each modeling the extrapolation process with optical flow via a spatial transformer layer. Although the neural network's prediction mechanism is intentionally similar, end-to-end learning on real data theoretically allows it to surpass the performance of simpler algorithms. While the neural network approach was found to be superior in the single radar setting, preliminary experiments did not show the same success with composited radar images and satellite data. Despite the optical flow approach being simpler and not requiring retraining with the introduction of new data sources, it is believed that neural nowcasting remains promising and could outperform simpler techniques with proper tuning of the network architecture and training regimen.

5   Results

5.1   Post-Launch Performance

Although the satellite-based rain detection model was trained to match radar fields, its reception by users was uncertain. A/B testing alone was insufficient to evaluate the product's performance, as it was essentially a new feature for several regions of Russia and could be well-received initially even if the map quality was low. Therefore, the performance of the new precipitation map was assessed using ground station data. While the optimal metrics for a user-facing precipitation prediction algorithm are still debated, there was evidence of the nowcasting product's popularity, and the aim was to replicate the properties of the radar-based precipitation map using satellite data. Specifically, the radar data differs from longer-term forecasts based on proprietary Meteum technology in having higher accuracy and lower systematic error rates (precipitation imbalance) at the cost of a lower F1 score when compared to ground station weather observations. The same comparison strategy was used to evaluate the performance of the new satellite-based rain detection algorithm over the federal districts of Russia. Results showed that while the accuracy of the satellite-based product is lower than that of radar, it is still better than traditional forecasts, with precipitation imbalance and F1 scores similar to those for radar. It is important to note that the radar located in Siberia was used only for verification at this stage; its data was not included in the training dataset. This comparison allows for evaluating precipitation detection quality in regions without radar observation.

This result confirmed the success of the new rain map. Additionally, A/B testing on users showed a statistically significant increase in daily active users (DAU) in areas where the rain map was previously unavailable (Siberia and Ural regions), justifying its rollout in late September.

Table 2: Comparison of precipitation detection methods with various metrics averaged over time.

Method	Accuracy	F1 Score	Precision	Recall
MPE	0.92	0.21	0.28	0.17
PP	0.86	0.30	0.24	0.40
Pointwise	0.91	0.48	0.40	0.61
U-Net w/o GFS	0.94	0.56	0.64	0.50
U-Net with GFS	0.94	0.60	0.62	0.59

6   Conclusion

A precipitation nowcasting system has been developed, implemented, and launched, utilizing both ground-based radar observations and geostationary satellite imagery. The system employs advanced machine learning algorithms and incorporates the physical properties of the atmosphere and ground surface based on NWP models. The inclusion of satellite data enables nowcasting for areas not covered by ground-based radars, achieving quality comparable to traditional radar-based nowcasts.

Table 3: Comparison of F1 scores of precipitation detection methods during different time periods.

Method	Day	Twilight	Night	All
MPE	0.19	0.22	0.21	0.21
PP	0.32	0.31	0.27	0.30
Pointwise	0.54	0.48	0.41	0.48
U-Net w/o GFS	0.65	0.55	0.49	0.56
U-Net with GFS	0.67	0.60	0.54	0.60

Currently, the system is limited to the region centered on European Russia within the Meteosat-8 field of view. Compared to previous solutions, the potential audience has been expanded from approximately 70 million to 300 million people, based on coverage area and population density. The approach can be extended to the rest of the Meteosat-8 coverage area. Scaling the technology to other geostationary satellites with similar measurement systems, such as Himawari and GOES, offers the possibility of providing global precipitation nowcasting and alerting services worldwide. However, differences in weather patterns across geographical regions will likely necessitate retraining the detection model and adjusting the set of input features.

One encountered problem is the sharp edge between radar and satellite data. This stationary edge on the weather map can confuse users, indicating the need for more sophisticated data fusion techniques. Experiments with image blending to erase conflicting observations along the border and inpainting the missing parts have been conducted.

7 Acknowledgments