An Online Training Procedure for Rain Detection Models Applied to Satellite Microwave Links

Christian Gianoglio, *Member, IEEE*, Matteo Colli, Sara Zani, and Daniele D. Caviglia, *Life Member, IEEE*

***Abstract*— In recent years, opportunistic rainfall sensing exploiting commercial microwave links (CMLs) has become popular for real-time monitoring of rainfall events. Among such systems, satellite-to-Earth microwave links (SMLs) are gaining great interest. An important information to extract from the collected data is whether or not an SML observation corresponds to a rain condition. This letter proposes an online training procedure to keep the algorithms that classify this condition up to date, overcoming the problem of data seasonality. Specifically, we propose to retrain the algorithms after one day of observations by applying a forgetting mechanism, where the oldest data are removed from the training dataset. Data were collected by the Smart Rainfall System (SRS) SML-type sensors located in the city of Genoa (Italy). We compare three classification algorithms: one based on an anomaly detection method and two based on machine learning (ML) algorithms. The outcomes of this analysis pave the way for implementing artificial intelligence models on embedded systems with limited resources such as the electronic units of SML sensors, hence reducing the data flow from the peripheral sensor network to the central data acquisition and integration server.**

***Index Terms*— Machine learning (ML), rainfall prediction, satellite microwave links, Smart Rainfall System (SRS).**

I. Introduction

REAL-TIME monitoring of precipitations in local areas is essential for effective emergency management [1], [2], [3]. Typically, professional rain gauge networks and long-range weather radars are used to collect and transmit precipitation measurements to National Weather Services (NWS), to provide regional and interregional monitoring of severe weather events [4], [5].

In recent years, there has been growing interest in opportunistic systems for real-time precipitation monitoring in local areas [6]. An example of such an approach is the Smart Rainfall System (SRS) that exploits the microwave links of the satellite broadcasting infrastructure and was conceived and validated at the University of Genoa (Italy) along with Artys (now a division of the Italian SME Darts Engineering srl) [7], [8]. Other researchers have experimented with other systems to reconstruct rainfall fields, such as those used in mobile communications networks (e.g., [9], [10]).

Manuscript received 9 October 2023; revised 26 October 2023; accepted 27 October 2023. Date of publication 30 October 2023; date of current version 9 November 2023. This work was supported by the European Union-FSE-REACT-EU, PON Research and Innovation 2014–2020 DM1062/2021 under Contract 11-G-13751-2. *(Corresponding author: Christian Gianoglio.)*

Christian Gianoglio and Daniele D. Caviglia are with DITEN, University of Genoa, 16145 Genoa, Italy (e-mail: christian.gianoglio@unige.it; daniele.caviglia@unige.it).

Matteo Colli and Sara Zani are with Artys, Darts Engineering srl, 16121 Genoa, Italy (e-mail: m.colli@artys.it; s.zani@artys.it).

Digital Object Identifier 10.1109/LGRS.2023.3328718

With the developments of enabling technologies for IoT applications [11], it is becoming feasible to embed algorithms in the sensor electronic units to process information and provide real-time monitoring for precipitation monitoring systems. As an example, Ostrometzky et al. [12] proposed an IoT-based system where Kalman filters were implemented on the cloud to recognize wet/dry periods to further estimate the rain intensity. In real-time precipitation monitoring adopting satellite-to-Earth microwave link (SML)-based systems, the initial step involves identifying wet and dry periods to estimate rainfall intensity (RI) from raw data. Specifically, RI is determined by measuring the attenuation of the received signal power during rainy conditions in comparison to power measurements taken during dry periods [7]. Importantly, raw data processing is energy-efficient, which is advantageous for hardware equipment and allows for the use of self-powered batteries. Machine learning (ML) and signal processing algorithms have been explored in the literature to cope with the instantaneous recognition of rain/nonrain conditions, as summarized below. In [13], an artificial neural network (ANN) identified rain/nonrain periods from an SML in the Ku-band, showing promising results to estimate rainfall from signal attenuation. Arslan et al. [14] adopted the instantaneous bit error rate (BER) measurements computing the average and the standard deviation on the received signals. A logistic regression analysis based on BER measurements is used to identify wet and dry periods. In [15], two Kalman filters were used to detect precipitation events and estimate rainfall from attenuation in the downlink channel of a commercial DVB satellite signal. He et al. [16] adopted a long short-term memory network (LSTM) for classifying rain/nonrain observations from SML signals. Xian et al. [17] used a support vector machine (SVM) to identify rain/nonrain periods from SML signals and determined the attenuation baseline during rainy periods by adopting an LSTM. Giro et al. [18] trained a randomized trees classifier to assess precipitation presence and estimated the rain attenuation from the received power signal to compute the rain rate with a 1-min time resolution. Gianoglio et al. [19] compared four ML algorithms (MLAs) with reduced computational complexity to classify rainy and nonrainy periods based on SML data. The results demonstrated the potential of using shallow MLAs to improve rain monitoring by SMLs. In general, the MLAs need to be trained with a large amount of data to include the large variety of signal singularities that are commonly due to spurious factors and that may affect the rain/no-rain classification accuracy [7]. This letter extends the outcomes of [19] presented by the same authors. Here, the authors propose a two-step method for

training the MLAs for classifying rain/nonrain events to avoid collecting a large amount of data before training the models. This helps the adoption of a monitoring system without the requirement of a large historical dataset available and independently of the SML setup. First, the algorithms are trained with a small amount of labeled data related to a few days of observations sensed by the SRS. Next, new observations are collected and classified by the algorithms. At the end of each test day, the measurements of a tipping-bucket rain gauge (TBRG) are taken into consideration and used as labels for the training procedure. The SRS data are then added to the training set by applying a forgetting mechanism where the oldest observations are removed from the training set and the new ones are added. This approach overcomes the problem of the seasonality of data that can significantly affect classification accuracy. Finally, the results obtained with the three algorithms, two based on supervised ML and one on an anomaly detection approach, are compared. All of them are designed to be implemented on resource-constrained devices and allow for fast retraining when a new observation day is acquired, as the forgetting mechanism avoids storing a large amount of data on the edge device.

II. Data Acquisition System

The data were collected between 2017 and 2019 using two SRS sensors, described in detail in [7]. The SRS sensors measure the power of the Ku-band signal received from satellites and provide the corresponding data to a server. The sensors were installed on the rooftop of the DITEN department at the Polytechnic School in the city of Genoa, Italy, and were connected to the low-noise block (LNB) converters mounted on parabolic dishes with diameters of 60 and 85 cm, respectively. The dishes were aligned toward the Turksat 42°E constellation with an elevation angle $\theta = 29.2048°$. To label the dataset, the reference observations for rainfall were measured using a TBRG located on the rooftop of the DICCA department [20], which is about 500 m away from the SRS sensors and below the link.

The dataset, available at SRS Github and denoted as $\mathcal{D}$, comprises $N_{\text{tot}} = 88$ days of labeled 1-min observations spanning from May 2017 to April 2019. The dataset consists of 50 days of nonrainy observations and 38 days containing rain events. It can be formalized as

$$\mathcal{D} = \left\{(\mathcal{X}, \boldsymbol{y})_i,\ \mathcal{X}_i \in \mathbb{R}^{N_o \times 1},\ \boldsymbol{y}_i \in \{0,1\}^{N_o}\right\} \tag{1}$$

with $i = 1, \ldots, N_{\text{tot}}$, where $\mathcal{X}_i$ is the set of SRS measurements (i.e., the raw data at the LNB measured in mV) of the ith day consisting of $N_o = 1440$ observations, and $\boldsymbol{y}_i$ is the corresponding set of no-rain/rain labels. Every minute SRS signals were transmitted to an Artys server via the internet connection of the DITEN department. To ensure a fair comparison, the two-year dataset has been cleaned from events with incomplete time series. The principal cause of data exclusion is related to the occurrence of interrupted time series related to the sensor operation conditions, often due to temporary power interruption or connectivity problems of the data transmission chain. The authors are currently improving the acquisition system to minimize these errors and ensure continuous monitoring.

III. Methodology

The algorithms, entitled to real-time monitoring of rainfall events, require offline training before providing the rain/no-rain classification, and they must be updated with new observations. The learning procedures and the differences between the algorithms are described in Sections III-A and III-B.

A. Learning Procedure

The learning strategy schema that applies to all the algorithms is shown in Procedure 1. In the following, the subscript $\alpha \in \{1, 2, 3\}$ will identify the specific algorithm. In the beginning, N_α days containing N_o observations each are extracted from the raw dataset $\mathcal{D}$ (1) using an initial part of observation history. The resulting dataset is

$$\mathcal{T}_\alpha = \left\{(\mathcal{X}, \boldsymbol{y})_j,\ \mathcal{X}_j \in \mathbb{R}^{N_o \times 1},\ \boldsymbol{y}_j \in \{0,1\}^{N_o}\right\} \tag{2}$$

with $j = 1, \ldots, N_\alpha$. A preprocessing of the data in $\mathcal{T}_\alpha$ is performed with an algorithm-specific procedure (step 1.2) generating a new dataset $\widetilde{\mathcal{T}}_\alpha$

$$\widetilde{\mathcal{T}}_\alpha = \left\{(\widetilde{\mathcal{X}}_\alpha, \boldsymbol{y})_j,\ \widetilde{\mathcal{X}}_{j,\alpha} \in \mathbb{R}^{N_o \times F_\alpha},\ \boldsymbol{y}_j \in \{0,1\}^{N_o}\right\} \tag{3}$$

where F_α represents the number of extracted features for the αth algorithm. Such procedures will be described in Section III-B. The algorithms are then trained offline with this initial $\widetilde{\mathcal{T}}_\alpha$ (step 1.3) to fix a starting point for an online procedure.

Afterward, the online phase is carried out: it consists of a sequence of two operations. First, the real-time rain condition $\hat{y}$ is assessed by the algorithm for each minute of observation (steps 2.3–2.6). Second, at the end of each day, the classification algorithm is updated using the SRS signals, as well as the actual rain condition provided by the TBRG (steps 2.10–2.11). However, to prevent both $\mathcal{T}_\alpha$ and $\widetilde{\mathcal{T}}_\alpha$ from becoming too large, a forgetting mechanism is applied beforehand (steps 2.7–2.9). This mechanism removes the measures corresponding to the oldest day of observations from $\mathcal{T}_\alpha$ while adding the new day. By doing so, the number of data in the set remains constant during training, which helps mitigate the bias introduced by the seasonality of the data [21].

B. Algorithms

1) Anomaly Detection Algorithm: It is currently the operating algorithm in the SRS. It functions as an anomaly detection method, where we define nonrainy observations as normal data and rainy observations as anomalies. Consequently, the training dataset exclusively comprises nonrainy data, which is used to establish thresholds for normal data, aiding in the identification of anomalies. In particular, anomaly detection algorithm (ADA) relies on the heuristic evaluation of three parameters: the minimum signal power $P_R^{\min}$, the maximum difference between two consecutive observations ΔP_R, and the standard deviation stdev, which are computed over a window

Procedure 1 All Algorithms: Learning Procedure

Input: Number of days for offline training N_α

1. **Offline Training**

1: Collect N_α days made of N_o observations in $\mathcal{T}_\alpha$ from $\mathcal{D}$
2: Preprocessing of data in $\mathcal{T} \rightarrow \widetilde{\mathcal{T}}_\alpha$
3: Train the algorithm with $\widetilde{\mathcal{T}}_\alpha$

2. **Online Procedure**

1: **for** a new day d **do**
2: $\quad \mathcal{X}_d \leftarrow empty$
3: $\quad$ **for** each new observation x_t in d **do**
4: $\quad\quad$ Classify $x_t \Rightarrow \hat{y}_t \in \{0, 1\}$
5: $\quad\quad$ Append x_t to $\mathcal{X}_d$
6: $\quad$ **end for**
7: $\quad$ Remove $(\mathcal{X}, \boldsymbol{y})_1$ from $\mathcal{T}_\alpha$
8: $\quad$ Collect the true labels $\boldsymbol{y}_d$ from the TBRG
9: $\quad$ Append $(\mathcal{X}, \boldsymbol{y})_d$ to $\mathcal{T}_\alpha$
10: $\quad$ Apply the preprocessing of step 1.2 on $\mathcal{T} \rightarrow \widetilde{\mathcal{T}}_\alpha$
11: $\quad$ Retrain the algorithm with $\widetilde{\mathcal{T}}_\alpha$
12: **end for**

of previous observations. The offline training of the algorithm occurs during step 1.3 of Procedure 1 and involves analyzing the first N_1 days (being $\alpha = 1$) of $\mathcal{D}$, which do not include any rainfall observation. As result, the training set $\widetilde{\mathcal{T}}_1$ with $F_1 = 1$ is generated. During step 1.3, the algorithm determines the three statistic parameters whose computations are detailed in SRS Github. ADA identifies the starting and end points of the rain by the evaluation of the three parameters. The criteria are described in SRS Github.

The algorithm updates its parameters (following step 2.11 of Procedure 1) after one day of observations only if the TBRG has not detected any rainfall on that day.

2) ML Algorithms: The ADA is evaluated alongside two MLAs: an ANN with only one hidden layer and a convolutional neural network (CNN). Unlike the ADA, the MLAs require a training set that includes both rainy and nonrainy observations. As a result, $N_\alpha/2$ rainy days and $N_\alpha/2$ days without rainfall are collected from $\mathcal{D}$ in $\mathcal{T}_\alpha$. In addition, the ANN requires feature extraction from the dataset $\mathcal{T}_\alpha$. Thus, a moving window with a fixed length l (being $l = 30$ as in [19]) is applied to the SRS signals in $\mathcal{T}_\alpha$. For each observation $x_t \in \mathcal{T}_\alpha$, an array is created from which the features are computed. This array contains the previous $l - 1$ observations and the current one and is denoted by $\boldsymbol{w}_t = [x_{t-l+1}, x_{t-l+2}, \ldots, x_t]$. In detail, the extracted features from each $\boldsymbol{w}_t$ are: mean($\boldsymbol{w}_t$), std($\boldsymbol{w}_t$), mean($\Delta\boldsymbol{w}_t$), std($\Delta\boldsymbol{w}_t$), min($\boldsymbol{w}_t$), max($\boldsymbol{w}_t$), min($\Delta\boldsymbol{w}_t$), and max($\Delta\boldsymbol{w}_t$) (where Δ indicates the difference between two consecutive measurements). The features selected for the MLAs were based on a previous study [19] while considering the feasibility of implementing processing techniques on a resource-constrained device meeting energy-efficient algorithm requirements. As a result, the ANN is trained with $\widetilde{\mathcal{T}}_2$ in (3) with $F_2 = 8$ (being $\alpha = 2$).

The CNN is fed directly with the arrays $\boldsymbol{w}_t$ as time series, resulting in $\widetilde{\mathcal{T}}_3$ with $F_3 = 30$ (being $\alpha = 3$).

Before training the MLAs, the $\widetilde{\mathcal{T}}_\alpha$ datasets are balanced to address the lower number of rainfall observations compared with the nonrainy ones. Since training the MLAs requires both rain and nonrain observations, the forgetting mechanism (steps 2.7–2.9 of Procedure 1) is applied by keeping two buffers of events, one with rainfall events and one without, consisting of a maximum of $N_\alpha/2$ events both. Each day of new observations is appended to the appropriate buffer based on whether the TBRG detects rain. The oldest event is removed from the buffer to ensure that it does not exceed the maximum capacity. In this way, the dataset $\mathcal{T}_\alpha$ used to compute $\widetilde{\mathcal{T}}_\alpha$ for updating the MLAs in steps 2.10–2.11 always includes observations related to rainfall conditions.

C. Experimental Setup

The ADA was trained by setting $N_1 = 5$, which represents a tradeoff choice between the need to consider a sufficient number of days to compute robust signal statistics and the need to take into account recent trends in signal behavior. As a consequence, the first five days of dataset $\mathcal{D}$ without any rain precipitation were collected in step 1.1 of Procedure 1.

The MLAs were trained offline by setting $N_\alpha = 10$ corresponding to five events per buffer, including the same five events collected for ADA, and adding the first five events in $\mathcal{D}$ that contain rain values.

The hyperparameters for the training procedure of the algorithms are defined as follows. For the ANN, we set the possible number of neurons in the hidden layer as $[25, 50, 100, 200]$ and the L2 regularizer $= [10^i$, with $i = -4, -3, \ldots, 4]$. For CNN, we choose two convolutional layers with the possible configuration of filters as $[(4, 4), (4, 8), (8, 8), (8, 16), (16, 32)]$ and kernel size as $[3, 5, 7]$. Each convolutional layer has been followed by a pooling layer with a size of 2. After the last pooling layer, a fully connected layer consisting of ten neurons was added. The convolutional and fully connected layers adopted the ReLU function as activation introducing nonlinearity in the network. For the MLAs, during the offline step, the best configuration of the hyperparameters was chosen by randomly splitting the training dataset into train and validation sets, using 80% of data for training and 20% for the model evaluation, keeping the configuration of hyperparameters that led to the highest classification accuracy on the validation set. The model selection procedure was not involved during the online step where the best models have just been tuned with the new SRS observations.

IV. Results and Discussion

A. Metrics Evaluation

The following notation is used: true positives (TPs) are rainy observations correctly classified, false positives (FPs) are nonrainy observations incorrectly classified as rainy, true negatives (TNs) are nonrainy observations correctly classified, and false negatives (FNs) are rainy observations misclassified as nonrainy.

For the evaluation and comparison of the performance of the algorithms, three metrics were computed.

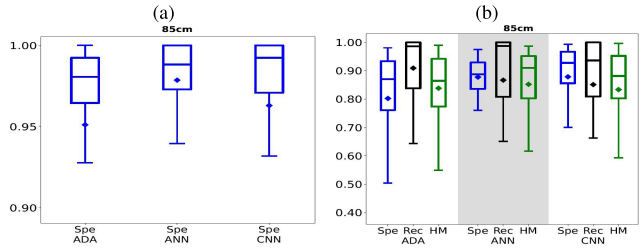

Fig. 1. (a) Specificity boxplots for each algorithm on nonrainy days. (b) Specificity (blue boxes), recall (black boxes), and HM (green boxes) for rainy days.

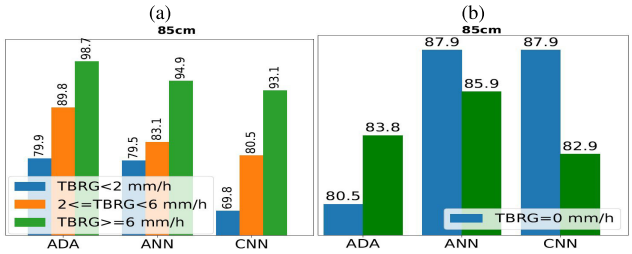

Fig. 2. Metrics computed on the 33 rainy days for each algorithm. (a) Recall (Rec), divided into three intervals. (b) Specificity (Spe) and the HM for both the dishes.

1) Recall (Rec) or sensitivity, Rec $=$ TP/(TP $+$ FN), indicates how much a model is accurate to predict the rainfall;
2) Specificity (Spe), Spe $=$ TN/(TN $+$ FP), indicates how much a model is accurate to predict the nonrainy condition;
3) Harmonic mean (HM) between Spe and Rec, computed as HM $=$ 2*Spe*Rec/(Spe $+$ Rec), trade-offs between the two metrics indicating how much they are balanced.

The metrics were evaluated on events not used during the offline training procedure of any algorithms by adopting only the data measured with the 85-cm dish since algorithms presented similar results concerning the 60-cm one (more details are given in SRS Github). It is worth noting that only specificity was computed for nonrainy days because the number of TP samples is 0, making Rec not computable.

Fig. 1 reports the results for the three algorithms on nonrainy and rainy days, divided into two subfigures. Fig. 1(a) displays the boxplots representing the Spe metric computed on nonrainy days. Because five out of 50 nonrainy days have been used for offline training, 45 nonrainy days from $\mathcal{D}$ were tested. Fig. 1(b) shows the boxplots concerning the three metrics (Spe represented by blue boxes, Rec by black boxes, and HM by green boxes) computed on rainy days. Straightforwardly, 33 rainy days from $\mathcal{D}$ were tested. In the figures, the lower and upper parts of the boxes delimit the first ($Q1$) and third ($Q3$) quartiles, while the whiskers (computed as $wh_{\text{upper}} = Q3 + 1.5 * (Q3 - Q1)$ and $wh_{\text{lower}} = Q1 - 1.5 * (Q3 - Q1)$) denote the variability outside the quartiles. A diamond symbol represents the average value of the metrics, while horizontal lines refer to the median values (i.e., $Q2$) for metrics.

In Fig. 1(a), CNN has the highest median score, while ANN has the highest mean, indicating that ADA and CNN misclassify more nonrainy data on a few days, which are considered outliers in the boxplot and are not visible in the figure. In Fig. 1(b), the CNN has the highest median and mean scores in terms of the Spe metric, but it exhibits the lowest median and mean scores for the Rec metric compared with the ANN and ADA models. As a result, ANN presents the best HM metric with the highest median and mean scores and the lowest variability. In general, ANN can be considered the most robust classifier for rainy/nonrainy periods.

In Fig. 2, all the three metrics are computed on observations from the 33 tested rainy days. Fig. 2(a) shows the Rec metric computed for three intervals of rainfall based on the TBRG measurements: low-level rain (TBRG $<$ 2 mm/h), intermediate-level rain (2 $\leq$ TBRG $<$ 6 mm/h), and high-level rain (TBRG $>$ 6 mm/h). Examining the results, we can observe that all the classifiers have a better recall in the high-level rain interval, indicating that heavier rainfall is detected more easily. The ADA outperforms all others in all the three rainfall intervals but also had the worst specificity [blue bars in Fig. 2(b)]. CNN shows the worst HM [green bars in Fig. 2(b)] since it has the worst Rec metric scores. As an eventual result, ANN presents the most balanced scores even in this experiment, confirming the previous results.

B. Classification Results

Fig. 3 provides three examples of SRS signals classified by the ANN using the 85-cm dish showing the best performance in the previous experiments. Each panel reports 1440 observations, with the y-axis representing the SRS signal in dBm and the RI measured by TBRG in mm/h, while the x-axis depicts time. The blue dots represent nonrainy observations correctly classified, the purple ones are nonrainy observations classified as rain, the green ones are rainy observations correctly classified, and the red dots represent rainy observations classified as not rain. The cyan bars represent TBRG measurements higher than 0 mm/h, i.e., associated with the rain observations.

Fig. 3(a) refers to a day that contains 328 rain observations with a maximum rain intensity of 146.8 mm/h. Most of the precipitation data are correctly classified, achieving a high Rec score, while the classifier presents a higher misclassification rate (i.e., lower Spe) for nonrain data. It is worth noting that two dips in the SRS time series, after 00:00, are classified as rain but the TBRG did not measure any precipitation. One possible explanation for such behavior is that a small cluster of rain crossing the microwave link was located between the sensors and the satellite, but it did not fall over the TBRG, and consequently, it was not observed by the latter. In addition, another group of misclassified data as rain (FP), before 16:00, is between two groups of rain observations, and hence the data could have similar features to the rainy ones or they are associated with a rain condition not measured by the TBRG as mentioned before. Fig. 3(b) presents an example where both rain and nonrain data are well-classified, achieving high Spe and Rec. The last example, in Fig. 3(c), shows a high misclassification rate for rain observations (i.e., a low Rec score), where 112 observations were labeled as rain with a maximum rain intensity of 13.3 mm/h. The misclassified data concern the beginning of the rain phenomenon, i.e., at the beginning of the two SRS dips (before 8:00 and 16:00). Moreover, the ANN classifies most of the data of the two peaks as rain, presenting some FPs. As before, these data may

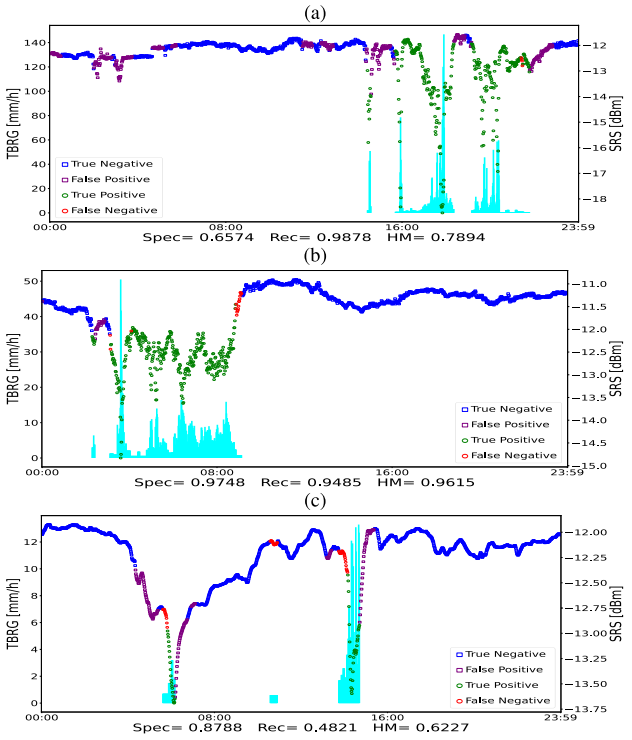

Fig. 3. Examples of classification results of ANN classifier on three rainy days, reporting the Spe, Rec, and HM scores. (a) ANN classification on 2017-09-09 85-cm dish. (b) ANN classification on 2018-01-01 85-cm dish. (c) ANN classification on 2019-04-07 85-cm dish.

be related to rain not observed by the TBRG, possibly due to the actual position and extent of the rainfall shower. It is worth noting that the ANN presents a poorer Rec on a day when the RI is lighter. The attenuation of the SRS signal due to rain is lower for previous cases, thus hindering the ANN from correctly classifying these observations.

V. Conclusion

This study proposes an effective strategy for updating algorithms that recognize rainfall conditions in real-time from SML signals, overcoming the issue of data seasonality. Each algorithm was tested with a new 1-min observation, which was then stored for the updating step at the end of each observed day. ANNs were found to have the best tradeoff in classifying rainy and nonrainy observations. These networks demonstrated a high recall metric, particularly for intense rain events. Future work will focus on implementing the algorithms on embedded systems to enable data processing next to the sensors, thereby avoiding continuous transmission to a central server. In addition, the study will enhance the performance of the ADA since both the inference and parameters update on a resource-constrained device could be faster than MLAs. One potential approach could be combining the ADA with ANNs to identify the onset and conclusion of rainfall events, increasing the overall performance. Furthermore, to address data imbalance during training, it could be beneficial to implement a weighting mechanism. This approach would allow us to leverage all the collected observations, preventing the loss of potentially valuable information that might occur when attempting to balance the dataset.

References

[1] B. B. Mirus, R. E. Becker, R. L. Baum, and J. B. Smith, "Integrating real-time subsurface hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning," *Landslides*, vol. 15, no. 10, pp. 1909–1919, Oct. 2018.

[2] A. Fortelli, N. Scafetta, and A. Mazzarella, "Nowcasting and real-time monitoring of heavy rainfall events inducing flash-floods: An application to Phlegraean area (Central-Southern Italy)," *Natural Hazards*, vol. 97, no. 2, pp. 861–889, Jun. 2019.

[3] Y. Zhao et al., "AI-based rainfall prediction model for debris flows," *Eng. Geol.*, vol. 296, Jan. 2022, Art. no. 106456.

[4] A. Cauteruccio, M. Colli, M. Stagnaro, L. G. Lanza, and E. Vuerich, "In-situ precipitation measurements," in *Springer Handbook of Atmospheric Measurements*. Cham, Switzerland: Springer, 2021, pp. 359–400.

[5] X. Zheng, H. Messer, Q. Wang, T. Xu, Y. Qin, and T. Yang, "On the potential of commercial microwave link networks for high spatial resolution rainfall monitoring in urban areas," *Atmos. Res.*, vol. 277, Oct. 2022, Art. no. 106289.

[6] F. Giannetti and R. Reggiannini, "Opportunistic rain rate estimation from measurements of satellite downlink attenuation: A survey," *Sensors*, vol. 21, no. 17, p. 5872, Aug. 2021.

[7] M. Colli et al., "A field assessment of a rain estimation system based on satellite-to-Earth microwave links," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 2864–2875, May 2019.

[8] M. Colli et al., "Rainfall fields monitoring based on satellite microwave down-links and traditional techniques in the city of Genoa," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 6266–6280, Sep. 2020.

[9] R. Uijlenhoet, A. Overeem, and H. Leijnse, "Opportunistic remote sensing of rainfall using microwave links from cellular communication networks," *Wiley Interdiscipl. Rev., Water*, vol. 5, no. 4, p. e1289, 2018.

[10] C. Chwala and H. Kunstmann, "Commercial microwave link networks for rainfall observation: Assessment of the current status and future challenges," *Wiley Interdiscipl. Rev., Water*, vol. 6, no. 2, p. e1337, 2019.

[11] S. Alam, S. T. Siddiqui, A. Ahmad, R. Ahmad, and M. Shuaib, "Internet of Things (IoT) enabling technologies, requirements, and security challenges," in *Proc. Adv. Data Inf. Sci. (ICDIS)*. Singapore: Springer, 2020, pp. 119–126.

[12] J. Ostrometzky, G. Rafalovich, B. Kagan, and H. Messer, "Stand-alone, affordable IoT satellite terminals and their opportunistic use for rain monitoring," *IEEE Internet Things Mag.*, vol. 5, no. 4, pp. 100–105, Dec. 2022.

[13] L. Barthès and C. Mallet, "Rainfall measurement from the opportunistic use of an Earth–space link in the Ku band," *Atmos. Meas. Techn.*, vol. 6, no. 8, pp. 2181–2193, Aug. 2013.

[14] C. H. Arslan, K. Aydin, J. V. Urbina, and L. Dyrud, "Satellite-link attenuation measurement technique for estimating rainfall accumulation," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 681–693, Feb. 2018.

[15] F. Giannetti, M. Moretti, R. Reggiannini, and A. Vaccaro, "The NEFOCAST system for detection and estimation of rainfall fields by the opportunistic use of broadcast satellite signals," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 34, no. 6, pp. 16–27, Jun. 2019.

[16] B. He, X. Liu, S. Hu, K. Song, and T. Gao, "Use of the C-band microwave link to distinguish between rainy and dry periods," *Adv. Meteorol.*, vol. 2019, Jul. 2019, Art. no. 3428786.

[17] M. Xian, X. Liu, M. Yin, K. Song, S. Zhao, and T. Gao, "Rainfall monitoring based on machine learning by Earth-space link in the Ku band," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3656–3668, 2020.

[18] R. A. Giro, L. Luini, C. G. Riva, D. Pimienta-del-Valle, and J. M. R. Salis, "Real-time rainfall estimation using satellite signals: Development and assessment of a new procedure," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–10, 2022.

[19] C. Gianoglio, A. Alyosef, M. Colli, S. Zani, and D. D. Caviglia, "Rain discrimination with machine learning classifiers for opportunistic rain detection system using satellite micro-wave links," *Sensors*, vol. 23, no. 3, p. 1202, Jan. 2023.

[20] M. Colli, L. G. Lanza, and P. W. Chan, "Co-located tipping-bucket and optical drop counter RI measurements and a simulated correction algorithm," *Atmos. Res.*, vol. 119, pp. 3–12, Jan. 2013.

[21] B. Lian, Z. Wei, X. Sun, Z. Li, and J. Zhao, "A review on rainfall measurement based on commercial microwave links in wireless cellular networks," *Sensors*, vol. 22, no. 12, p. 4395, Jun. 2022.

Open Access funding provided by 'Università degli Studi di Genova' within the CRUI CARE Agreement