Multimodal Fusion for Deforestation Detection: Integrating Weather and Satellite Alerts with Deep Learning

Kevin Elezi[1], João Pedro Martinez[2], Felipe Ferrari[3], Raul Queiroz Feitosa[4],

[1] Dept. of Data Science and Engineering, Politecnico di Torino, Italy - kevin.elezi@studenti.polito.it
[2] Dept. of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, Brazil - joao.martinez@aluno.puc-rio.br
[3] Dept. of Cartographic Engineering, Military Institute of Engineering, Brazil – ferrari@ime.eb.br
[4] Dept. of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, Brazil - raul@ele.puc-rio.br

Keywords: deforestation, multimodal fusion, deep learning, weather data, satellite alerts

Abstract

This paper presents a multimodal deep learning framework for deforestation detection that integrates satellite-based deforestation alerts (*DETER*) with weather and atmospheric variables (*WF*). While we hypothesized that *WF* could provide complementary signals for short-term deforestation prediction, our experiments show that fusion models provide only isolated and modest gains, with no consistent improvement over *DETER*-only baselines. The *WF*-only model highlights structurally vulnerable regions but lacks precision in identifying which specific pixels will be deforested—an ability retained by the alert-based models. Our findings confirm the dominant role of satellite alerts for precise deforestation monitoring, with *WF* signals offering limited added value for operational systems.

1. Introduction

1.1 Context

The Brazilian Amazon faces significant deforestation from farming, road construction, and large-scale mining [Giljum et al., 2022], putting critical ecosystems at risk by threatening biodiversity, climate stability, and carbon storage [Lorenz et al., 2021]. If forest loss reaches 40% of the area, it could trigger global temperature increases of up to 4°C by century's end [Pereira et al., 2020], making improved monitoring essential [Lemos and Silva, 2011, Aguiar et al., 2016].

Deep learning, especially convolutional neural networks (CNNs), has become the dominant approach for deforestation monitoring. These models effectively detect forest loss [de Bem et al., 2020], support uncertainty estimation [Martinez et al., 2024], and enhance detection through data fusion [Ferrari et al., 2023], highlighting the value of integrating complementary data—such as SAR, optical, or atmospheric signals—into pipelines for cloud-prone regions like the Amazon.

With advances in environmental sensing, mobile sources (e.g., smartphones, USB, and drone sensors [Mobile Physics, 2024, Sladojevic et al., 2024, Zboralski and Kunz, 2024]) may soon complement or replace satellites for weather and air quality data collection. We hypothesize that such mobile-collected data, here simulated using satellite sources, could enhance deforestation prediction in a cost-effective, scalable way. Nonetheless, we acknowledge the preliminary nature of this integration and the well-known challenges of linking diffuse environmental signals with precise, localized deforestation events such as satellite-detected alerts. This study evaluates whether integrating weather and air quality variables (collectively termed ***WF*** for weather features) can improve satellite alert performance and reduce reliance on costly verification methods such as helicopter surveillance.

We selected *WF* that are scientifically linked to forest loss and openly available on Google Earth Engine (GEE), using satellite datasets as proxies in the absence of mobile-collected data. Precipitation patterns shift in deforested areas [Qin et al., 2025]; water vapor (H_2O) and carbon monoxide (CO) levels reflect biomass removal and atmospheric changes [Lemes et al., 2023]; fire alerts indicate clearing events [Menezes et al., 2021]; and surface temperature signals vegetation stress [Lewis, 1998, Maillard et al., 2022]. These variables were sourced from MODIS (LST) [NASA, 2025b], FIRMS (fire alerts) [NASA, 2025a], CHIRPS (precipitation) [UCSB, 2025], and Copernicus Sentinel-5P (CO and H_2O) [Copernicus, 2025].

1.2 Contributions

This work explores the feasibility of using derived weather and air quality features (*WF*) for deforestation prediction, under the hypothesis that such data could one day be collected via mobile sensors. Our main contributions are:

- Starting from the standard deforestation prediction pipeline based on alerts from the *Detecção do Desmatamento em Tempo Real* system (***DETER***) [Diniz et al., 2015], we evaluate the impact of incorporating simulated *WF* using satellite data into a late fusion model. By keeping the model architecture, loss function, and patch extraction strategy fixed, we aim to isolate and quantify the specific contribution of *WF* signals when combined with *DETER* alerts.
- While overall results were modest, we showed that in certain periods, the model captures similar spatial patterns from *WF* as it does from *DETER*, suggesting these variables hold relevant environmental signals.
- We provide a controlled, high-resolution evaluation of *WF* utility in a realistic, imbalanced setting, offering insights into their potential as complementary inputs in future monitoring systems.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-W4-2025-121-2026

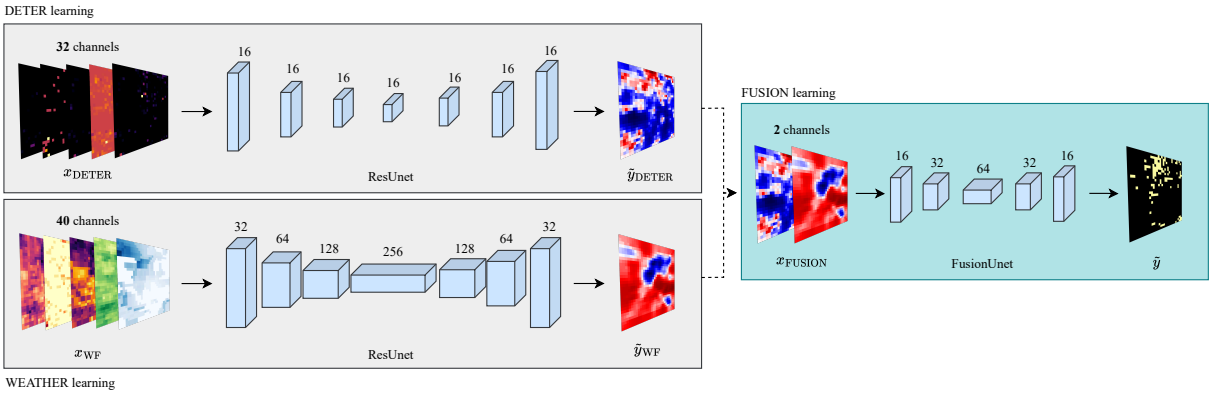

Figure 1. Overview of the proposed dual-stream fusion pipeline. Separate ResUNet models process *DETER* and weather data (x_{DETER}, x_{WF}) to produce prediction maps ($\tilde{y}_{\text{DETER}}, \tilde{y}_{WF}$). These outputs are stacked into a fusion tensor (x_{FUSION}) and then fed into a lightweight CNN (FusionNet) to generate the final deforestation predictions.

2. Methodology

We adopted ResUNet as our core architecture for deforestation prediction due to its proven performance in remote sensing tasks [de Bem et al., 2020]. This choice was informed by preliminary experiments comparing alternative models, as detailed in Section 4.

Following this, we explored strategies for incorporating *WF* signals into the learning pipeline. A first attempt to feed both *DETER* alerts and *WF* channels into a unified model input, denoted as $x_{[DETER,WF]}$, resulted in lower performance compared to the *DETER*-only baseline. This suggested that the *WF* signal—being weaker and more diffuse—was likely overshadowed by the more directly informative *DETER* data.

To address this, we adopted a modular approach by training two separate models: one using only *DETER* data (serving as our baseline) and another using only *WF* inputs. We then applied a **late fusion strategy** [Zhao et al., 2021], combining their respective outputs—$\widetilde{y}_{\text{DETER}}$ and $\widetilde{y}_{\text{WF}}$, obtained after sigmoid activation—through a lightweight convolutional network called *FusionUNet*. This compact encoder-decoder architecture is composed of `Conv2d`, `BatchNorm2d`, and `ReLU` layers in each block, designed to capture potential complementary relationships between the two modalities while preserving the interpretability of each stream. The complete methodology pipeline, including FusionUNet, is depicted in Figure 1.

3. Experimental Analysis

3.1 Dataset

The *DETER* dataset comprises near real-time alerts of forest cover changes in the Amazon, provided by the National Institute for Space Research (INPE) [Diniz et al., 2015]. Although it includes multiple alert classifications, this study focused only on deforestation-related records. In addition to spatial annotations, these records include dates, which were adjusted to a biweekly frequency: dates after the 15th of the month were reassigned to the 16th; others were set to the 1st.

To further work with this data, we built a grid at 5km resolution over the Brazilian Amazon biome area. This grid was used as a spatial reference to create alert maps for each time instant, all stacked at the end (see Figure 2). With these spatio-temporal data structures, called *datacubes*, the following features were derived for each grid cell:

***DETER* features**

- ***area***: deforested area alerted, in km^2.
- ***count_area***: normalized count of deforested spots.
- ***acumul_area***: accumulated deforested area.
- ***diff_area***: difference series of deforested area.
- ***season_area***: seasonality factor of deforested area.

The seasonality factor aims to capture deforestation patterns that potentially repeat over time. To achieve this, a first moving average was computed, pixel-wise, for the deforested area. This average, denoted μ_t, where t indicates the temporal index of the biweeks, targeted the pixel value one year before, i.e. 24 biweeks, and its 3 past and future values. It's worth noting that, in order to compute μ_t, we lost the first 27 biweeks values pixel-wise. From this smoothed series, the seasonality factor, Σ_t, was finally computed in a recursive manner, geometrically weighting every seasonal period (SP) value of this new series. In practice, it gives more importance to the latest values, as presented below:

$$\Sigma_t = \begin{cases} [\mu_t + \Sigma_{t-SP}] \cdot 0.5, & \text{if } t > SP \\ \mu_t, & \text{if } t \leq SP \end{cases} \quad t \in \mathbb{Z}_+ \tag{1}$$

On the other hand, the *WF* were pre-processed to match the biweekly frequency of the deforestation labels, normalized to $[0, 1]$, and mapped onto the same 5 km grid as *DETER* to ensure precise spatial alignment across modalities. Since the data came from GEE and different satellites, some values were occasionally missing. This mainly happened because clouds or their shadows blocked the sensors, or because of gaps caused by the satellites' viewing angle and orbit. To handle these cases, we applied inverse distance weighting (IDW) interpolation, which estimates the value at a target location x_0 as a weighted average of neighboring valid measurements. The interpolated value $\hat{f}(x_0)$ is defined as:

$$\hat{f}(x_0) = \frac{\sum_i w_i f(x_i)}{\sum_i w_i}, \tag{2}$$

where $f(x_i)$ denotes the value at a neighboring valid point. The

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-W4-2025-121-2026

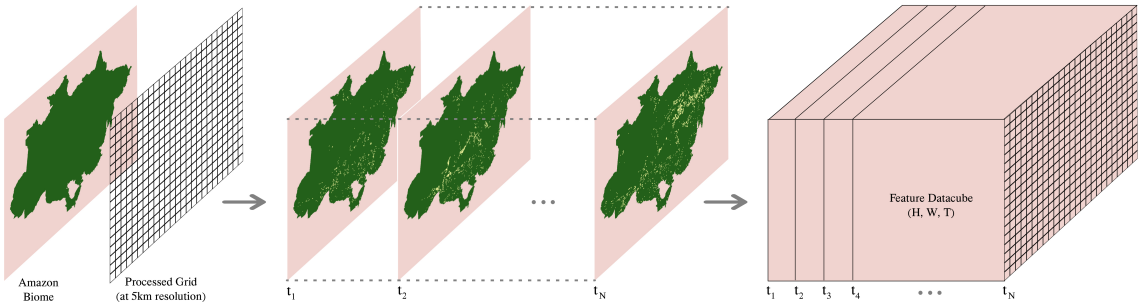

Figure 2. Construction of a *DETER* feature datacube. A 5km grid is overlaid on the Amazon, aggregating deforestation alerts per cell for each biweek. Stacking these maps over time produces a spatio-temporal datacube for model input. The *WF* datacubes are processed similarly, aggregating their geodatapoints into the same grid cells.

weights w_i follow an inverse-square relationship:

$$w_i = \frac{1}{d(x_0, x_i)^2}, \tag{3}$$

with $d(x_0, x_i)$ representing the distance between the target location x_0 and the neighbor x_i.

The final *WF* set comprised precipitation, H_2O, CO, fire alerts, and LST.

***WF* features**

- ***LST***: average daytime land surface temperature
- ***CO***: column density of carbon monoxide
- ***H_2O***: column of water vapor for humidity
- ***fire alerts***: alerts of nearby fire hot spots
- ***precipitation***: rainfall time series

*Biweekly temporal aggregation applies a **mean** for LST, CO, and H_2O, and a **sum** for fire alerts and precipitation.*

In the end, to achieve a fair comparison of *DETER*-only and late fusion methodology including *WF*-only, all the experiments have considered for both datasets the following time span: from *[2018-11-16]* up to *[2024-12-16]*, a bit more than six years of data.

3.2 Experimental Protocol

Initially, the protocol started setting the temporal splits to training and inference phases, common to all the processed feature datacubes. In this study, it was separated a total of 103 predictable biweeks for training, 12 for validation and 24 for testing.

To conduct the experiments—whether following the *DETER* or Fusion methodology—we extracted patches according to the spatio-temporal structure of the processed features. As illustrated in Figure 3, the patches were selected for each biweek based on the presence of at least one deforested pixel in the target reference.

This strategy was mainly an attempt to mitigate the intrinsic imbalance between deforested and non-deforested areas, ensuring relevant training samples. Furthermore, the patches had dimensions of 32×32 pixels and were extracted with 60% overlap to balance spatial coverage. We also defined an inner border within each patch to discard edge pixel losses during training, as these border regions have less contextual information than inner pixels.

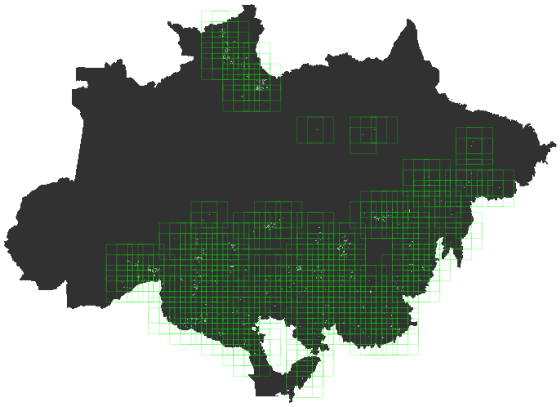

Figure 3. Distribution of training patches (green) over the Amazon for the reference date *[2023-02-16]*. Patches were selected if they contained at least one deforested pixel, focusing training on relevant areas and reducing class imbalance.

Using the spatial reference of each patch, we extracted temporal slices from the feature datacubes to construct the input and target tensors, as shown in Figure 4. The *WF* input tensors concatenated time slices composed by 8 consecutive time layers. For *DETER*, the process was similar, except for the features ***diff_area*** and ***season_area***, which used 7 and 1 time layers, respectively. In both datasets, the goal was to capture short-term temporal patterns to predict deforested areas in the target tensors, one biweek ahead.

To supervise the fitting process, we adopted the Focal Loss $\mathcal{L}_{\text{focal}}$ [Lin et al., 2017] function to also handle the disproportion between deforested and non-deforested areas:

$$\mathcal{L}_{\text{focal}} = -\alpha_t (1 - p_t)^{\gamma} \log(p_t) \tag{4}$$

where p_t is the model's estimated probability for the deforestation class, α_t is a balancing factor, and γ is the focusing parameter.

During loss computation, we systematically disregarded pixels within the patches' inner borders and losses from pixels outside the binary mask delimiting the studied region.

Furthermore, the experiments were conducted using the Adam optimizer, initialized with a learning rate of 1e-3 and a weight

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-W4-2025-121-2026

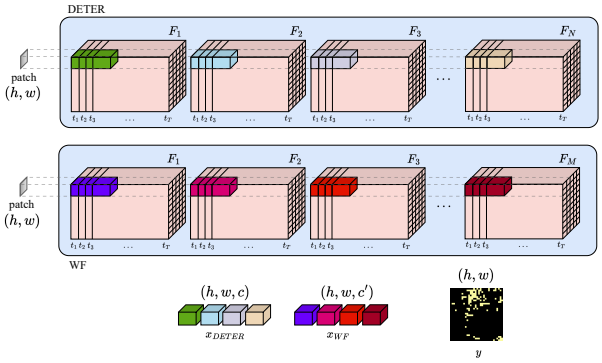

Figure 4. For each selected patch and biweek, the *DETER* input tensor (x_{DETER}) is built by extracting fixed-length temporal sequences for all features within the *DETER* datacubes. The *WF* input tensor (x_{WF}) is built similarly from *WF* datacubes. The target tensor (y) contains the subsequent deforestation alert from *DETER*.

decay of 1e-5. A cosine annealing schedule was applied to adapt the learning rate throughout training too, limited to a minimum value of 5e-5. We set up a maximum of 100 training epochs, with early stopping activated based on validation loss, employing a patience of 10 epochs to prevent overfitting.

All experiments were conducted on a machine equipped with dual NVIDIA GeForce RTX 4090 GPUs (each with 24 GB VRAM) running CUDA 12.9. This also included their testing phase, which was not patch-based, but in fact predicted the entire maps of deforestation for each biweek ahead.

4. Results

4.1 Model Selection via Regression

To identify the most suitable model for deforestation prediction, we initially framed the task as a regression problem. This approach allowed us to evaluate each model not only on accuracy but also on prediction confidence and temporal stability across biweeks. Regression outputs are continuous, providing a more direct measure of model certainty than thresholded classification probabilities, which is particularly useful for selecting a model that is reliable and robust over time. We compared several state-of-the-art architectures, including DeepLabv3+ [Chen et al., 2018], SwinUNet [Cao et al., 2023], SimVP2 [Tan et al., 2025], XGBoost [Cue La Rosa et al., 2024], and ResUNet.

As shown in Figure 5, ResUNet consistently achieved the lowest mean absolute error (MAE) across biweeks, confirming both its accuracy and stability [de Bem et al., 2020]. The consistency of this comparison is further highlighted by the inverse-U shape shared across each model's metric, indicating a similar pattern of performance variation over time and reinforcing the reliability of ResUNet relative to the other architectures. Its strong performance and stability motivated its selection for the subsequent classification and fusion experiments, where confidence in predictions is crucial for interpreting the effects of different input features and ensuring the robustness of downstream deforestation mapping.

4.2 Classification and Fusion Performance

For the main deforestation detection experiments, we transitioned to a classification setting to better quantify predictive performance using F1, precision, and recall — metrics well-suited for imbalanced data where accuracy is misleading. These metrics were reported per biweek and averaged across the test period.

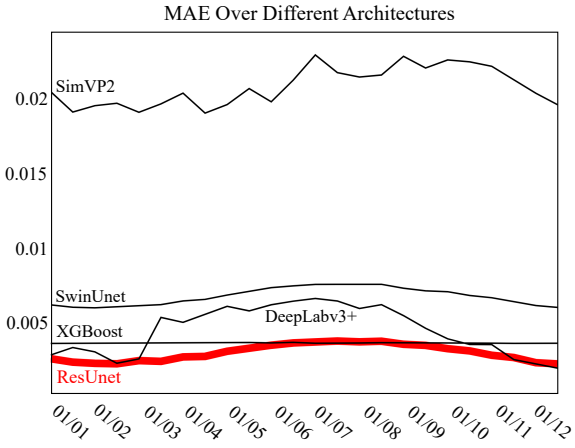

Figure 5. Mean absolute error (MAE) comparison of tested architectures across biweeks. ResUNet consistently achieved the lowest error.

In certain biweeks, marked by stars in Figure 6, the Fusion model shows modest improvements in F1 score and recall compared to the *DETER*-only baseline, suggesting that under specific seasonal conditions, *WF* variables provide limited complementary information. However, these gains were isolated and not systematic, highlighting the limited reliability of *WF* signals for consistent deforestation prediction.

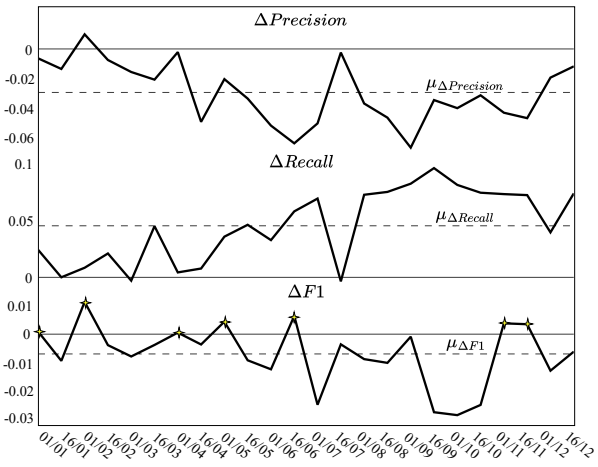

Figure 6. Difference between Fusion and *DETER* models in precision, recall, and F1 score across the entire 2024 biweeks. Positive values indicate Fusion outperforming *DETER*; stars mark biweeks with notable gains.

As illustrated in Figure 7, these rare gains are visible as true positive pixels (in yellow, indicated by arrows) that appear only in the Fusion output $\hat{y}_{\text{FUSION}}$ and not in $\hat{y}_{\text{DETER}}$, while green indicates true positives detected by both models. Here, $\widetilde{y}_{\text{FUSION}}$, $\widetilde{y}_{\text{DETER}}$, and $\widetilde{y}_{WF}$ denote probability maps after sigmoid activation; $\hat{y}$ denotes thresholded binary maps.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-W4-2025-121-2026 | © Author(s) 2026. CC BY 4.0 License.

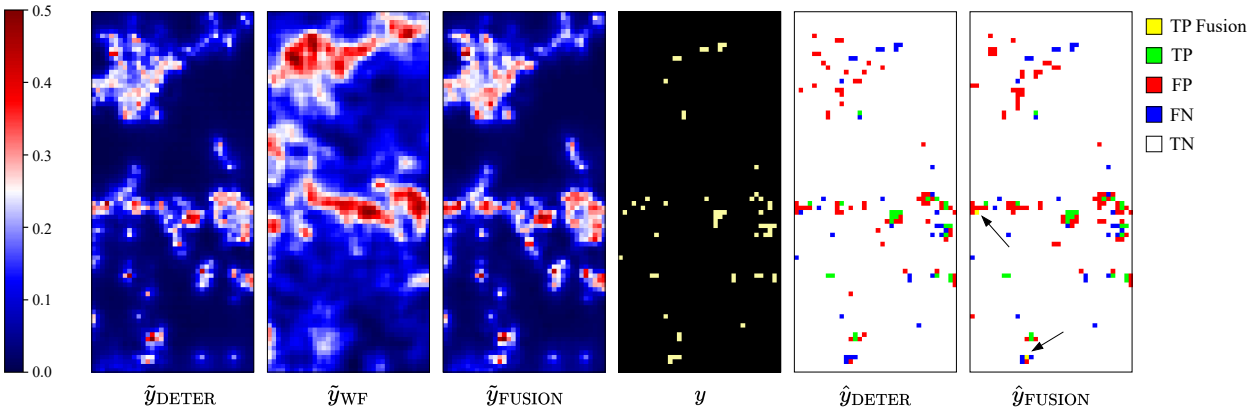

Figure 7. Predicted probability maps (post-sigmoid) and final classifications for biweek *[2024-11-01]*. The left panels show probability maps, while the right panels display classifications where yellow marks true positives captured only by Fusion (indicated by arrows) and green indicates true positives detected by both Fusion and *DETER*.

4.3 Discussion

The relationship between the *DETER* and weather feature (*WF*) outputs reveals a clear pattern: the FusionUNet tends to preserve the *DETER* prediction as the dominant signal. This behavior aligns with our expectation, as *DETER* provides direct satellite-based alerts of deforestation. However, the model demonstrates only limited ability to modulate this signal based on auxiliary *WF* information.

While isolated cases exist where the *WF* stream highlights areas missed by *DETER*, these are outweighed by frequent false positives. Figure 6 shows this trade-off: recall improves slightly in some biweeks, but precision drops significantly due to more false alarms. This imbalance has serious operational implications: higher false positive rates trigger costly field checks [ICMBio, 2021]. Thus, even a small increase in false positives can create a significant operational burden, outweighing the marginal gain in true positives.

We initially hypothesized that a deep convolutional neural network could capture subtle, complex dependencies between weather variables and deforestation—patterns otherwise hard to formalize. Yet, despite architectural flexibility, the network fails to leverage weather features to enhance precision in pinpointing deforested pixels. The fusion process instead introduces noise rather than complementary information, leading to unnecessary over-predictions in regions without deforestation.

We believe this limited contribution is not merely a modeling issue but reflects the intrinsic weakness of the *WF* signals themselves. Compared to the direct and precise nature of *DETER* alerts, the information encoded in *WF* variables appears too diffuse to meaningfully support the prediction of short-term deforestation.

Despite experimenting with various architectural configurations and input combinations to guide the model toward extracting only the most relevant patterns from *WF* data, the performance plateaued. This suggests that much of the predictive signal that the model can learn from *WF* is already implicitly contained in the *DETER* input, rendering the added complexity of fusion minimally beneficial.

However, as shown in Figure 7, the output of $\tilde{y}_{\text{WF}}$ continues to reflect spatial patterns similar to those of the *DETER*-only model ($\tilde{y}_{\text{DETER}}$), suggesting that *WF* signals highlight structurally vulnerable areas but lack the precision needed for accurate deforestation prediction. This explains why the deforested pixels in $\tilde{y}_{\text{FUSION}}$ resemble those in $\tilde{y}_{\text{DETER}}$, yet appear more dispersed: the fusion model tends to broaden the spatial extent of deforestation predictions, building around the core areas flagged by *DETER* using the additional, but less precise, cues from *WF*. While this expansion occasionally captures new deforestation signals, our metrics show that it more often introduces noise, limiting the practical benefits of fusion for improving predictive performance.

5. Conclusion

Despite the general trend in our results, where the weather features (*WF*) do not significantly improve the baseline provided by *DETER*—the Fusion model shows that *WF* inputs occasionally offer complementary information. In certain periods, *WF* slightly enhances recall by helping detect deforestation signals missed by the satellite alert baseline.

Nonetheless, we expected that indicating where fires occurred, where precipitation might support forest regrowth, or where temperature could reflect vegetation resilience would refine predictions. In practice, the deforestation alerts alone seem to capture much of this information. Adjusting predictions with *WF* did not confirm the need—at least for now—for a mobile data collection pipeline. The observed gains are isolated and not consistent enough to justify the added complexity or cost of integrating *WF* into deforestation monitoring.

The *WF*-only model, though much weaker in accuracy, tends to highlight areas also flagged by *DETER* as deforestation-prone. This suggests that *WF* signals may encode environmental conditions linked to vulnerability, but lack the precision to predict which specific pixels will actually be deforested.

References

Aguiar, A. P., Vieira, I. C., Assis, T. O., Dalla-Nora, E. L., Toledo, P. M., Santos-Junior, R. A., Batistella, M., Coelho, A. S., Savaget, E. K., Aragão, L. E. O. C., Nobre, C. A., Ometto, J. P., 2016. Land use change emission scenarios: anticipating a forest transition process in the Brazilian Amazon. *Glob. Change Biol.*, 22(5), 1821–1840.

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., Wang, M., 2023. Swin-unet: Unet-like pure transformer for medical image segmentation. *Proc. Eur. Conf. Comput. Vis. (ECCV)*.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 801–818.

Copernicus, 2025. Sentinel-5p offl l3 co: Carbon monoxide total column, dataset code `COPERNICUS/S5P/OFFL/L3_CO`. Google Earth Engine.

Cue La Rosa, L. E., Ferrari, F., Bezerra, F. G. S., Souza, R. A., Maretto, R. V., Aguiar, A. P. D., Vinhas, L., Nigri Happ, P., Feitosa, R. Q., 2024. XGBOOST and multitemporal DETER data for deforestation forecasting. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, X-3-2024, 193–198.

de Bem, P. P., de Carvalho Junior, O. A., Fontes Guimarães, R., Trancoso Gomes, R. A., 2020. Change detection of deforestation in the Brazilian Amazon using Landsat data and convolutional neural networks. *Remote Sens.*, 12(6), 901.

Diniz, C., Souza, A., Santos, D., Dias, M., Cavalcante da Luz, N., Moraes, D., Maia, J., Gomes, A., Narvaes, I., Valeriano, D., Maurano, L., Adami, M., 2015. DETER-B: The new Amazon near real-time deforestation detection system. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, PP, 1–10.

Ferrari, F., Ferreira, M. P., de Almeida, C. A., Feitosa, R. Q., 2023. Fusing Sentinel-1 and Sentinel-2 images for deforestation detection in the Brazilian Amazon under diverse cloud conditions. *IEEE Geosci. Remote Sens. Lett.*, 20, 3242430.

Giljum, S., Maus, V., Kuschnig, N., Luckeneder, S., Tost, M., Sonter, L. J., Bebbington, A. J., 2022. A pantropical assessment of deforestation caused by industrial mining. *Proc. Natl. Acad. Sci. U.S.A.*, 119(38), e2118273119.

ICMBio, 2021. Termo de referência para contratação de helicópteros para apoio a atividades de fiscalização ambiental. `https://www.gov.br/icmbio/pt-br/acesso-a-informacao/licitacoes-e-contratos/licitacoes/pregao/2021/arquivos/2TERMODEREFERNCIATRanexoIdoedital.pdf`.

Lemes, M. R., Sampaio, G., Garcia-Carreras, L., Fisch, G., Alves, L. M., Bassett, R., Betts, R., Maksic, J., Shimizu, M. H., Torres, R. R., Guatura, M., Basso, L. S., Bispo, P. d. C., 2023. Impacts on South America moisture transport under Amazon deforestation and 2 C global warming. *Sci. Total Environ.*, 905, 167407.

Lemos, A. L. F., Silva, J. d. A., 2011. Desmatamento na Amazônia Legal: Evolução, Causas, Monitoramento e Possibilidades de Mitigação Através do Fundo Amazônia. *Floresta e Ambiente*, 18(1), 98–108.

Lewis, T. J., 1998. The effect of deforestation on ground surface temperatures. *Glob. Planet. Change*, 18(1), 1–13.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017. Focal loss for dense object detection. *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2980–2988.

Lorenz, C., de Oliveira Lage, M., Chiaravalloti-Neto, F., 2021. Deforestation hotspots, climate crisis, and the perfect scenario for the next epidemic: The Amazon time bomb. *Sci. Total Environ.*, 783, 147090.

Maillard, O., Vides-Almonacid, R., Salazar, A., Larrea-Alcázar, D., 2022. Effect of deforestation on land surface temperature in the Chiquitania region, Bolivia. *Land*, 12, 2.

Martinez, J. A. C., da Costa, G. A. O. P., Messias, C. G., Soler, L. d. S., de Almeida, C. A., Feitosa, R. Q., 2024. Enhancing deforestation monitoring in the Brazilian Amazon: A semi-automatic approach leveraging uncertainty estimation. *ISPRS J. Photogramm. Remote Sens.*, 210, 110–127.

Menezes, D., Pucci, R., Mourão, J., Gandour, C., 2021. The relationship between forest fires and deforestation in the amazon: phenomena are more closely related in rural settlements and in occupied public lands. Climate Policy Initiative. [Online].

Mobile Physics, 2024. Mobile physics: Air quality sensing through smartphones (qualcomm snapdragon 8 gen 3 integration). `https://www.mobilephysics.com/`.

NASA, 2025a. Fire information for resource management system (firms), dataset code `FIRMS`. Google Earth Engine.

NASA, 2025b. Modis: Land surface temperature and emissivity 8-day global 0.05deg cmg v061, dataset code `MODIS/061/MOD21C2`. Google Earth Engine.

Pereira, E. J. A. L., de Santana Ribeiro, L. C., da Silva Freitas, L. F., de Barros Pereira, H. B., 2020. Brazilian policy and agribusiness damage the Amazon rainforest. *Land Use Policy*, 92, 104491.

Qin, Y., Wang, D., Ziegler, A. D., Fu, B., Zeng, Z., 2025. Impact of Amazonian deforestation on precipitation reverses between seasons. *Nature*, 639(8053), 102–108.

Sladojevic, S., Arsenovic, M., Nikolić, D., Anderla, A., Stefanović, D., 2024. Advancements in mobile-based air pollution detection: From literature review to practical implementation. *J. Sensors*, 4, 1–17.

Tan, C., Gao, Z., Li, S., Li, S. Z., 2025. SimVPv2: Towards simple yet powerful spatiotemporal predictive learning. *IEEE Trans. Multimedia*, PP, 1–15.

UCSB, 2025. Chirps daily: Climate hazards group infrared precipitation with station data, dataset code `UCSB-CHG/CHIRPS/DAILY`. Google Earth Engine.

Zboralski, D., Kunz, M., 2024. Mobile systems for assessing air quality: available solutions and application examples. *Bull. Geogr. Phys. Geogr. Ser.*, 5–26.

Zhao, J., Wang, Y., Cao, Y., Guo, M., Huang, X., Zhang, R., Dou, X., Niu, X., Cui, Y., Wang, J., 2021. The fusion strategy of 2D and 3D information based on deep learning: A review. *Remote Sens.*, 13(20), 4029.