IMPACT OF MISSING VIEWS IN MULTI-VIEW MODEL PREDICTIONS FOR VEGETATION APPLICATIONS

Francisco Mena*, **Andreas Dengel**
University of Kaiserslautern-Landau & DFKI
Kaiserslautern, Germany

Diego Arenas, Marlon Nuske
DFKI
Kaiserslautern, Germany

ABSTRACT

Earth observation (EO) applications involving complex and heterogeneous data sources are commonly approached with machine learning models. However, there is a common assumption that data sources will be persistently available. Different situations could affect the availability of EO sources, like noise, clouds, or satellite mission failures. In this work, we assess the impact of missing temporal and static EO sources in trained models across two datasets involving classification and regression tasks. We compare the predictive quality of different methods and find that some are naturally more robust to missing data. The Ensemble strategy, in particular, achieves a prediction robustness up to 99%. We evidence that missing scenarios are more challenging in regression than classification task. Finally, we find that the optical view is the most critical view when it is missing individually.

1 INTRODUCTION

Many data-driven solutions in Earth Observation (EO) leverage data from multiple data sources (Garnot et al., 2022; Mena et al., 2024). The objective is to corroborate and complement the information on individual observations for the particular task. The literature provides evidence that the inclusion of additional data, such as Remote Sensing (RS) based, is crucial to enrich the modeling and improve the predictive quality (Garnot et al., 2022; Hong et al., 2020; Mena et al., 2023). However, the assumption that EO sources are persistently available may not hold.

There are different situations in which EO data sources may not be available. Specific RS instruments have a finite lifetime (e.g. based on the fuel usage), and may be affected by noise (Hong et al., 2020) or clouds in the case of optical sensors (Garnot et al., 2022). Besides, unexpected errors can terminate the operation earlier, such as the failure of the Sentinel-1B satellite in 2021.

Despite the research focus on more complex Multi-View Learning (MVL) models (Mena et al., 2024), few works have explored the challenge of missing views. Here, we refer to a *view* as a data source, and a *missing view* as missing an entire data source. Srivastava et al. (2019) proposed a technique to retrieve a similar sample when one view is missing. Hong et al. (2020) showed that the predictions of specific MVL models worsen less when views are missing. Gawlikowski et al. (2023) showed that a missing optical view affects the predictions more than a missing radar view.

Unlike recent works, we present a study on datasets involving time series and static EO sources. Our research question is: what is the impact of missing views in trained MVL models? The prediction robustness results allow us to formulate advice on model selection based on the task type and input views. Furthermore, this study can serve as a way to understand the impact of specific views in trained models, as well as the sensitivity of these models to missing data.

2 MULTI-VIEW LEARNING AND MISSING VIEWS

A MVL setting consists of having multiple *views* as input data to a machine learning model to improve predictive quality (Mena et al., 2024). A view can be any set of features expressing a

*F. Mena acknowledges support through a scholarship of the University of Kaiserslautern-Landau. Available code at `https://github.com/fmenat/missingviews-study-EO`

Table 1: Datasets description. The last column is the spatial resolution of the target pixel.

Dataset	Task	Samples	Years	Where	Pixel
CH-M Tseng et al. (2021)	Multi-class classification	29642	2016-2022	Global	10 m
LFMC Rao et al. (2020)	Regression	2578	2015-2019	USA	250 m

different perspective of each sample, such as entire data sources, optical or radar images, vegetation indices, terrain information or metadata.

Several works in the literature have explored MVL models with neural networks to achieve an optimal data fusion (Garnot et al., 2022; Mena et al., 2023). Some standard MVL models use *Input*, *Feature*, or *Decision* fusion strategies, where the name suggests where the fusion is placed in the model architecture (first, middle, or last layer respectively). Additionally, in the *Ensemble* strategy (Mena et al., 2023), the predictions from view-dedicated models (previously trained) are aggregated.

During inference, the occurrence of missing views can be seen as a special case of domain shift (Gawlikowski et al., 2023). By lacking views, the input data deviates from the training distribution, leading to a scenario for which the model is unprepared. However, there are some techniques applied to trained models that can mitigate this effect, which we describe below.

Impute. A straightforward technique is to fill in the missing views Hong et al. (2020). However, since the entire data source is missing, there are no available features for interpolation or in-painting techniques. Therefore, we use the average of each view in the training data as the imputation value, as it brings more information than an arbitrary value (e.g. null).

Exemplar. Based on information retrieval, the missing view can be replaced with a similar sample via a training-set lookup. We consider the technique in Srivastava et al. (2019) that searches for the missing view using the available views in a shared space. The space is obtained with a linear projection (via CCA) from features learned by a separated MVL model.

Ignore. Some MVL models are adapted naturally to missing views by a dynamic fusion. In the Ensemble strategy, the predictions of the view-dedicated model associated to the missing view are omitted in the aggregation. Similarly, with Feature fusion, the features of the missing views are ignored when using the average as the merge function.

3 Evaluation

3.1 Datasets

In the following we describe two datasets used in this study, while the Table 1 presents an overview.

CropHarvest Multi-crop (CH-M): We use a multi-crop version of the CropHarvest data for multi-view crop recognition Tseng et al. (2021). This is a classification task in which a given location during a particular season has to be classified among 10 crop-type groups. The input views are optical (from S2), radar (from S1), and weather time series. These temporal views were re-sampled monthly for 1 year. An additional static view is the topographic information.

Live Fuel Moisture Content (LFMC): We use a dataset for multi-view moisture content estimation Rao et al. (2020). This is a regression task in which the vegetation water per dry biomass (in percentage) in a given location at a specific moment is predicted. The input views are optical (from L8) and radar (from S1) time series. These views were re-sampled monthly along 4 months. Additional static views are the topographic information, soil features, canopy height, and land-cover class.

3.2 Experiment Settings

We apply a z-score normalization to the input data. The categorical and ordinal views (land-cover and canopy height) are one-hot-vector encoded. We use an MLP as the encoder for the static views. For the temporal views, we use a 1D CNN encoder in the CH-M, and a GRU in the LFMC. We use two layers with 128 dimensions in the encoders, and an MLP with one hidden layer as prediction head. An ADAM optimizer is used with batch-size 128 and early stopping. The loss function is cross-entropy in classification and mean squared error in regression task.

Table 2: Predictive quality (AA) of MVL models for different missing views scenarios in the CH-M. Columns to the right have more missing views. We highlight the **best** and **second best** value.

Method	Technique	No Miss	Radar	Optical	Weather + static	Radar+weather + static	Optical + weather+static
Input-concat	Impute	**0.738**	0.641	0.296	0.534	**0.534**	0.142
Feature-concat	Impute	**0.727**	0.624	0.290	0.558	0.390	0.159
Feature-cca	Exemplar	**0.727**	0.285	0.384	0.094	0.107	0.100
Feature-avg	Ignore	0.726	**0.674**	**0.542**	**0.582**	0.529	**0.306**
Ensemble-avg	Ignore	0.715	**0.708**	**0.613**	**0.711**	**0.715**	**0.523**

Table 3: Predictive quality (R^2) of MVL models for different missing views scenarios in the LFMC. Columns to the right have more missing views. † is a value lower than -100. We highlight the **best** and **second best** value.

Method	Technique	No Miss	Radar	Optical	Static	Radar+static	Optical+static
Input-concat	Impute	**0.717**	**0.650**	0.060	0.185	0.165	-0.047
Feature-concat	Impute	0.667	0.599	**0.274**	**0.352**	**0.290**	**0.081**
Feature-cca	Exemplar	0.667	†	-0.260	†	†	†
Feature-avg	Ignore	**0.683**	**0.618**	0.142	†	†	†
Ensemble-avg	Ignore	0.312	0.292	**0.243**	**0.407**	**0.392**	**0.239**

We evaluate using the 10-fold cross-validation. The predictive quality is measured with the Average Accuracy (AA) in classification, and the coefficient of determination (R^2) in regression. We include the Performance Robustness Score (PRS) presented by Heinrich et al. (2023) that is based on the error in the predictions with missing views relative to the same error when all views are available.

The missing views scenario consists of making predictions on the validation fold with fewer views available than during training. We experiment with a moderate degree of missingness, as the case when only radar is missing, or when optical is missing; an intermediate missingness, when all other views except radar and optical are missing; and an extreme degree of missingness, when all the views are missing except one (single-view inference with only radar or optical). We compare the techniques described in Sec. 2. Two MVL models with the Impute technique: Input and Feature with concatenation (Input-concat, Feature-concat). Two MVL models with ignoring techniques: Feature and Ensemble with averaging (Feature-avg, Ensemble-avg). Lastly, one MVL model based on the Feature fusion with the Exemplar technique (Feature-cca), see Sec. 2 for details.

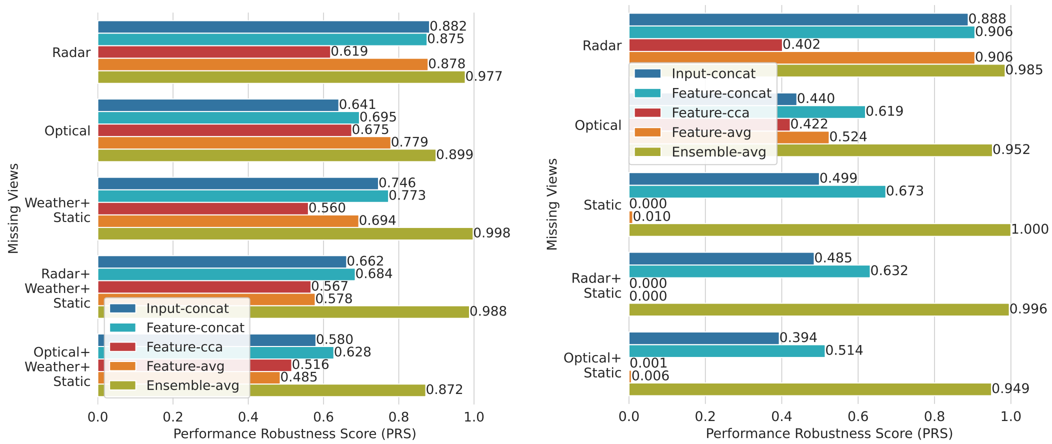
3.3 Experiment Results

In Table 2 we show the predictive quality in the classification task. The results of the Input-concat method decreases significantly when views are missing. We observe that, when using Feature fusion with ignoring techniques (Feature-avg) the impact of missing views is mitigated more than with the Impute or Exemplar (Feature-concat, Feature-cca). However, these do not achieve the values of the Ensemble-avg, which is the method least affected by missing views.

We observe similar results in the regression task, shown in Table 3, except when using the Ignore technique. The predictions of Feature fusion-based models with the ignore techniques become worse, up to negative R^2 values, when views are missing. Besides, the results of the Ensemble-avg method in the LFMC data is relatively bad (≈ 0.3) with no or a moderate degree of missing views.

For the robustness results in Fig. 1, we corroborate the lower impact of missing views in the MVL models when using ignoring techniques. The Ensemble-avg method got a PRS close to one in some cases, which means that the error of predictions with missing views is lower or the same as the error in predictions without missing views. However, this is on average, as there can still be negative changes in the predictive quality, such as when radar is missing in the CH-M data (Table 2), or a bad predictive quality in itself, such as in the LFMC data (Table 3). We notice that the Feature-concat method has higher robustness than Feature-avg in regression. Besides, the Feature-cca method has a fairly low robustness, especially in the regression task, reaching a PRS of 0 in some scenarios.

Overall, we note that the impact of missing views depends on the MVL model along with how to treat missing views, as previous works have shown (Hong et al., 2020; Garnot et al., 2022; Gawlikowski



(a) Classification in the CH-M data. (b) Regression in the LFMC data.

Figure 1: Prediction robustness of MVL models for different missing views scenarios.

et al., 2023). The negative effect of missing view increases from the moderate, intermediate to the extreme degree. In addition, we observe that the impact of missing optical view is stronger than when missing radar. This means that the optical view is more difficult to supplement than others, reflecting its greater importance for RS-based applications. Ancillary data, like static and weather views, still provide valuable information to MVL models. For instance, the predictive quality of some methods in the LFMC data is worse when the static views are missing, and, in CH-M, some are worse when weather and static views are missing compared to the radar view.

In Table 4 we present our ongoing work where we modify the training by applying a sensor dropout technique, i.e. we randomly drop some EO sources in each *batch*. Comparing the results with Table 2, we observe that when there are no missing views, the results decrease. However, for the missing view scenarios, the predictive quality increases significantly. This highlights the potential of techniques that explicit modify the model learning to increase the prediction robustness.

Table 4: AA of MVL models with sensor dropout for different missing views scenarios in CH-M.

Method	Technique	No Miss	Radar	Optical	Weather + static	Radar+weather + static	Optical + weather+static
Input-concat	Impute	0.687	0.665	0.508	0.683	0.655	0.277
Feature-concat	Impute	0.659	0.612	0.510	0.591	0.515	0.292
Feature-avg	Ignore	0.731	0.705	0.610	0.720	0.698	0.455

Table 5: R^2 of MVL models with sensor dropout for different missing views scenarios in the LFMC.

Method	Technique	No Miss	Radar	Optical	Static	Radar+static	Optical+static
Input-concat	Impute	0.539	0.543	0.335	0.337	0.299	0.108
Feature-concat	Impute	0.469	0.447	0.272	0.344	0.273	0.094
Feature-avg	Ignore	0.507	0.474	0.289	-1.173	-21.061	-9.590

4 Conclusion

We evaluated the impact of missing views in MVL models across two tasks with time series and static EO data. We showed that missing specific views (such as optical) significantly affects the predictive quality. Nevertheless, the prediction robustness can be improved by designing a method adjustable for the missing views. In addition, due to the differences in predicting a continuous value to a categorical one, the impact of missing views is more severe in regression than in classification tasks. Based on the results, we provide the following advice for model selection in missing view scenarios: if views are sufficiently discriminative to allow individual predictions of the task, use the Ensemble strategy that ignores the missing predictions, otherwise use the Feature fusion strategy ignoring missing views in classification, or imputing missing views in regression tasks.

REFERENCES

Vivien Sainte Fare Garnot, Loic Landrieu, and Nesrine Chehata. Multi-modal temporal attention models for crop mapping from satellite time series. *ISPRS Journal of Photogrammetry and Remote Sensing*, 187:294–305, 2022. doi: 10.1016/j.isprsjprs.2022.03.012.

Jakob Gawlikowski, Sudipan Saha, Julia Niebling, and Xiao Xiang Zhu. Handling unexpected inputs: incorporating source-wise out-of-distribution detection into sar-optical data fusion for scene classification. *EURASIP Journal on Advances in Signal Processing*, 2023(1):47, 2023. doi: 10.1186/s13634-023-01008-z.

René Heinrich, Christoph Scholz, Stephan Vogt, and Malte Lehna. Targeted adversarial attacks on wind power forecasts. *Machine Learning*, 2023. doi: 10.1007/s10994-023-06396-9.

Danfeng Hong, Lianru Gao, Naoto Yokoya, Jing Yao, Jocelyn Chanussot, Qian Du, and Bing Zhang. More diverse means better: Multimodal deep learning meets remote-sensing imagery classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(5):4340–4354, 2020. doi: 10.1109/TGRS.2020.3016820.

Francisco Mena, Diego Arenas, Marlon Nuske, and Andreas Dengel. A comparative assessment of multi-view fusion learning for crop classification. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 5631–5634. IEEE, 2023. doi: 10.1109/IGARSS52108.2023.10282138.

Francisco Mena, Diego Arenas, Marlon Nuske, and Andreas Dengel. Common practices and taxonomy in deep multi-view fusion for remote sensing applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 4797–4818, 2024. doi: 10.1109/JSTARS.2024.3361556.

Krishna Rao, A Park Williams, Jacqueline Fortin Flefil, and Alexandra G Konings. Sar-enhanced mapping of live fuel moisture content. *Remote Sensing of Environment*, 245:111797, 2020. doi: 10.1016/j.rse.2020.111797.

Shivangi Srivastava, John E Vargas-Munoz, and Devis Tuia. Understanding urban landuse from the above and ground perspectives: A deep learning, multimodal solution. *Remote sensing of environment*, 228:129–143, 2019. doi: 10.1016/j.rse.2019.04.014.

Gabriel Tseng, Ivan Zvonkov, Catherine Lilian Nakalembe, and Hannah Kerner. Cropharvest: A global dataset for crop-type classification. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021.