Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

Smart Data and Knowledge Services

On what depends the robustness of multi-source models to missing data in Earth observation?

Francisco Mena (1,2), Diego Arenas (1),
Miro Miranda (1,2), Andreas Dengel (1,2)

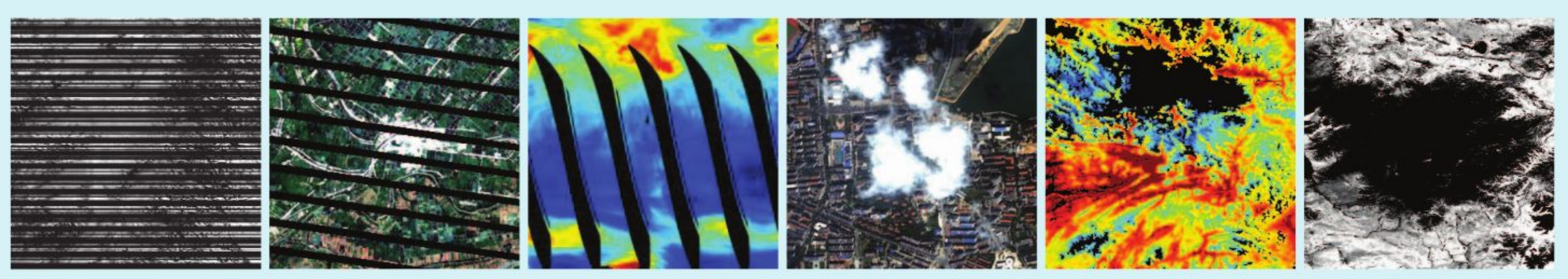
1 University of Kaiserslautern-Landau (RPTU)
2 German Research Center for Artificial Intelligence (DFKI)

IGARSS Brisbane 2025

ai4eo SOLUTION FACTORY

Motivation

- The EO domain suffers from the missing data problem.
- Data collection occurs under operational constrains in real-world environment.



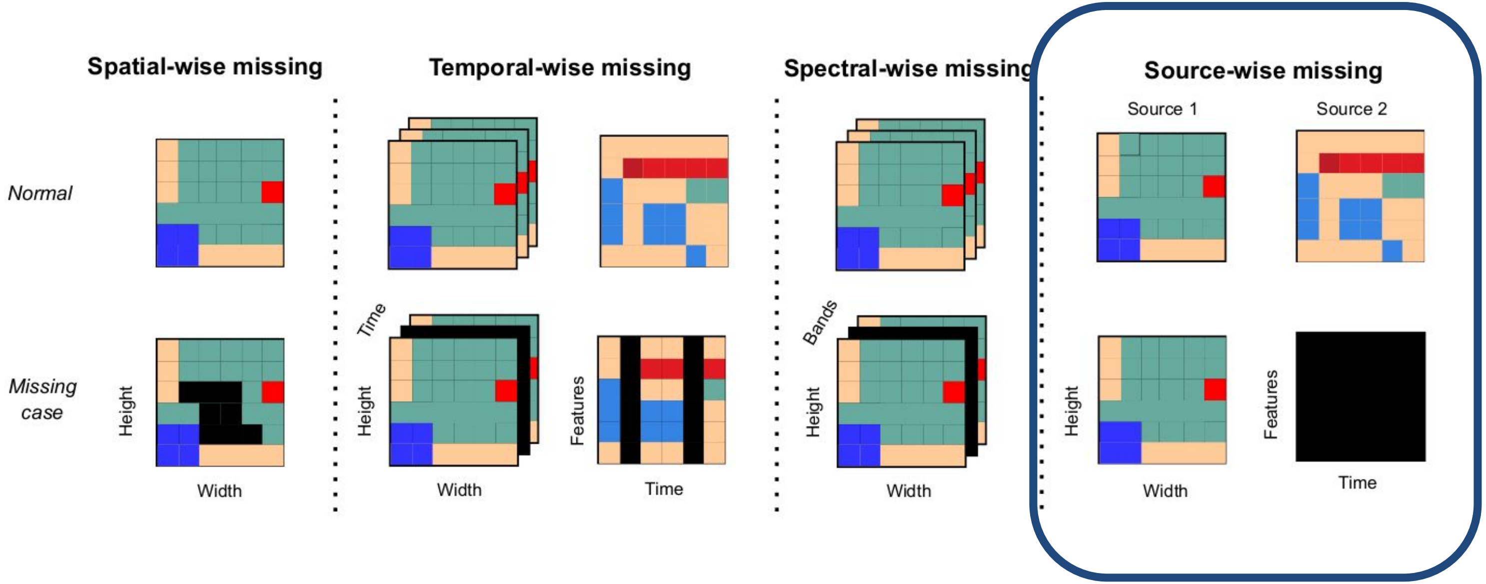
Mission ends for Copernicus Sentinel-1B satellite

03/08/2022 34953 VIEWS 119 LIKES

Shen, et al (2015). Missing information reconstruction of remote sensing data: A technical review. IEEE Geoscience and Remote Sensing Magazine

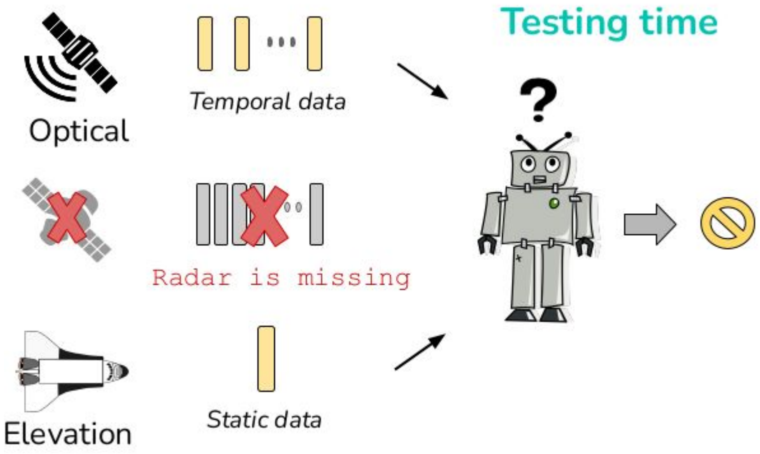
https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1/Mission_ends_for_Copernicus_Sentinel-1B_satellite

Types of missing Earth observation data



Source-wise missing in multi-source models

- How to handle source-wise missing data at inference time?

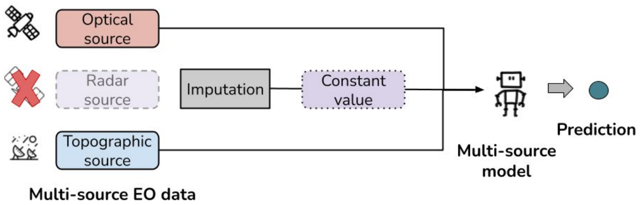


Mena et al. "Missing Data as Augmentation in the Earth Observation Domain: A Multi-View Learning Approach." arXiv preprint arXiv:2501.01132 (2025).

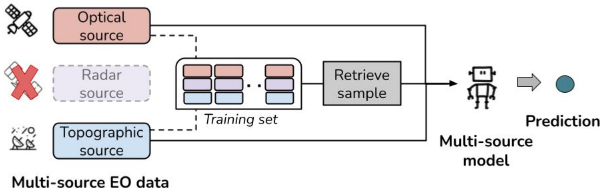
Data processing techniques

- How to handle source-wise missing data at inference time?

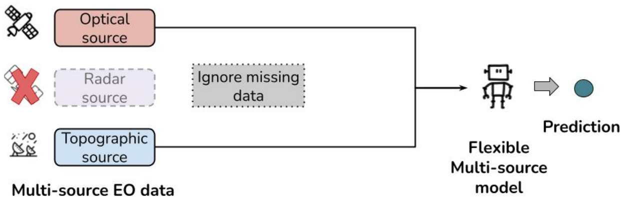
- **Impute**: replace missing data with a numerical value [1].



- **Retrieval**: retrieve a similar training sample based on available data [2].



- **Ignore**: design a model that can ignore the missing data [3].



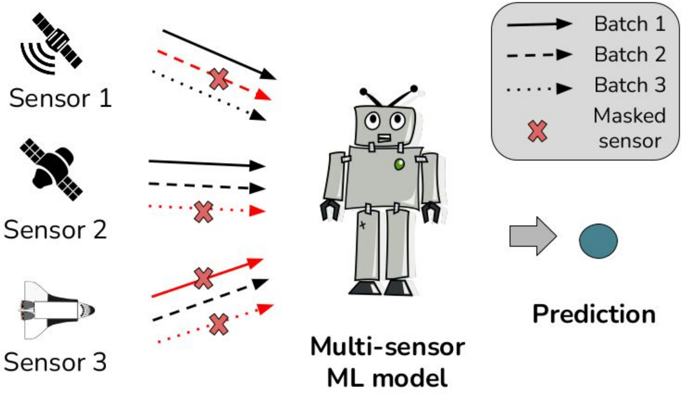
[1] Hong, Danfeng, et al. "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification." *IEEE Transactions on Geoscience and Remote Sensing* (2020)
[2] Srivastava, Shivangi, et al. "Understanding urban landuse from the above and ground perspectives: A deep learning, multimodal solution." *Remote Sensing of Environment* (2019)
[3] Mena, Francisco, et al. "Impact assessment of missing data in model predictions for Earth observation applications." IGARSS 2024-2024 IEEE, 2024.

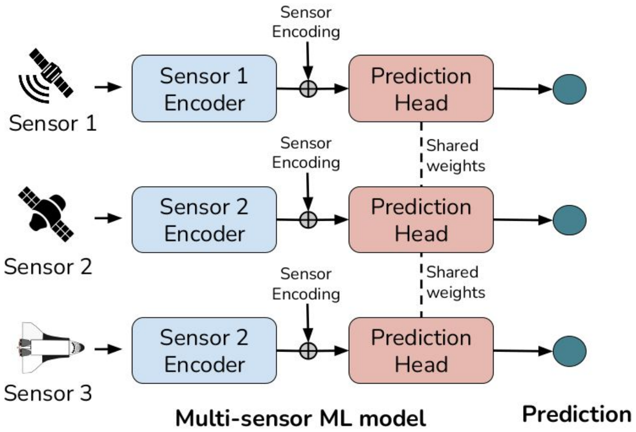
Robustness component for missing data

- How to be robust to source-wise missing data?

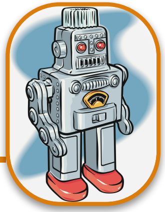
- **Simulate missing data** (during training)
  - e.g. sensor dropout

- **Enhance similarity** (between data sources)
  - e.g. sharing weights in per-source models





Model comparison



- Single-source model as baseline

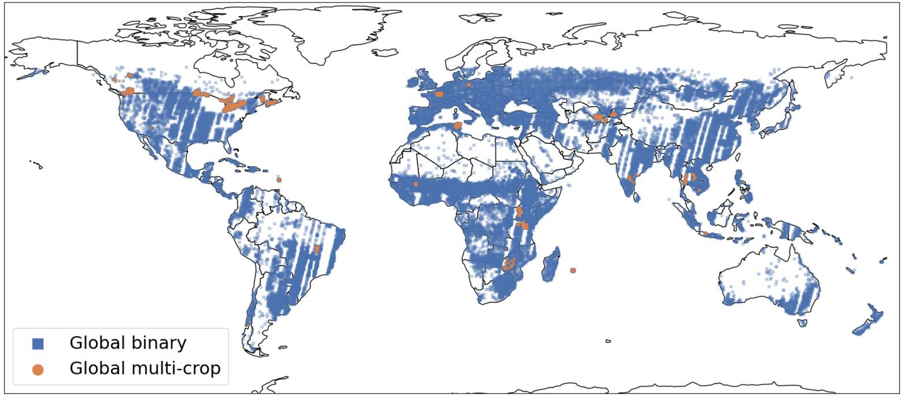
Name	Fusion strategy	Data processing	Robustness component	*Extra*
ISensD [1]	Input-level	Impute (zero)	Sensor dropout	*concat-based fusion*
FSensD [2]	Feature-level	Ignore (mask)	Sensor dropout	*transformer-based fusion*
OOD-f [3]	Feature-level	Impute (zero)	Simulate all missing cases	*concat-based fusion*
TIMML [4]	Feature-level	Ignore (mask)	Sensor dropout	*auxiliary predictions, transformer-based fusion*
ESensI [1]	Ensemble-based	Ignore	Sharing weights	*average-based fusion*
Co-Ens [5]	Ensemble-based	Ignore	Prediction similarity	*average-based fusion*

[1] Mena, Francisco, et al. "Impact assessment of missing data in model predictions for Earth observation applications." IGARSS 2024-2024 IEEE, 2024.
[2] Chen, Yuxing, et al. "A novel approach to incomplete multimodal learning for remote sensing data fusion." IEEE Transactions on Geoscience and Remote Sensing (2024).
[3] Gawlikowski, Jakob, et al. "Handling unexpected inputs: incorporating source-wise out-of-distribution detection into SAR-optical data fusion for scene classification." EURASIP .1 (2023).
[4] Xu, Guozheng, et al. "Transformer-Based Incomplete Multi-Modal Learning for Land Cover Classification." IGARSS 2024-2024 IEEE, 2024.
[5] Xie, Yuxing, et al. "A co-learning method to utilize optical images and photogrammetric point clouds for building extraction." International Journal of Applied EO and Geoinformation 116 (2023)

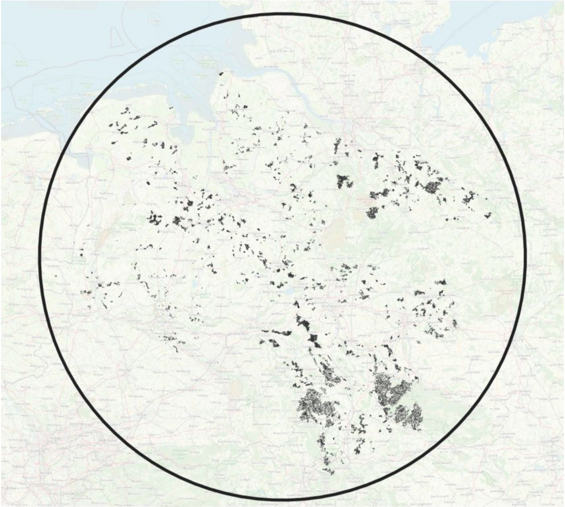
Vegetation datasets

- Datasets with temporal length over 1 year.

CropHarvest [1]: crop recognition



TreeSatAI-TS [2]:
tree-species recognition



[1] Tseng, Gabriel, et al. "Cropharvest: A global dataset for crop-type classification." *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track* (2021)
[2] Astruc, Guillaume, et al. "Omnisat: Self-supervised modality fusion for earth observation." *European Conference on Computer Vision*. Cham: Springer Nature Switzerland, 2024.

Pixel-wise classification datasets

- Datasets with temporal length over 1 year.

Predictive Task	Samples	Where	When	Data sources with temporal features	Data sources with static features
Cropland classification [1]	62,100 /6,900	Global	2016-2021	optical, radar, weather	topography
Crop-type classification [1]	26,678 /2,964	Global	2016-2021	optical, radar, weather	topography
Tree-species classification [2]	38,520 /5,044	Germany	2017-2020	optical, radar	aerial image

[1] Tseng, Gabriel, et al. "Cropharvest: A global dataset for crop-type classification." *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track* (2021)
[2] Astruc, Guillaume, et al. "Omnisat: Self-supervised modality fusion for earth observation." *European Conference on Computer Vision*. Cham: Springer Nature Switzerland, 2024.

(weighted) F1 Results I
Cropland and crop-type recognition

Cropland classification

Data sources: Optical	Radar	Weather	Topography	Single	ISensD	FSensD	OOD-f	TIMML	ESensI	Co-Ens
✓	✓	✓	✓		80.0	82.1	**84.7**	83.1	81.4	79.2
-	✓	✓	✓		72.6 (10.2)	79.0 (3.9)	**82.5** (2.7)	80.1 (3.7)	78.1 (4.2)	76.8 (3.1)
✓	-	✓	✓		79.4 (0.8)	81.8 (0.4)	**84.2** (0.6)	82.8 (0.4)	82.4 (†)	79.3 (†)
✓	✓	-	✓		80.4 (†)	80.0 (2.6)	**81.3** (4.2)	**81.3** (2.2)	79.7 (2.2)	78.4 (1.0)
✓	✓	✓	-		80.0 (0.0)	82.1 (0.0)	**84.2** (0.5)	83.1 (0.0)	81.2 (0.3)	80.0 (†)
✓	-	-	-	81.8	80.0	79.7	79.5	**80.9**	80.8	**80.9**
-	✓	-	-	71.4	44.6	69.5	69.6	70.1	70.7	**72.2**
-	-	✓	-	77.8	60.8	75.8	77.2	76.2	**77.3**	77.2
-	-	-	✓	67.7		56.5	66.9	63.8	65.9	**68.3**

The model with the best robustness depends on the missing data scenario

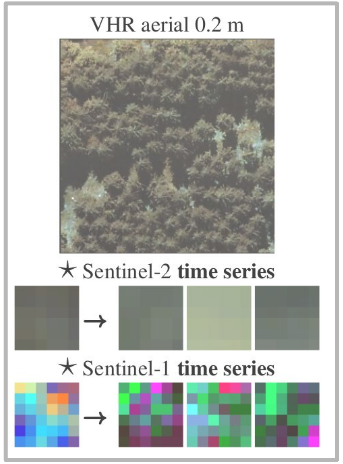
Optical	Radar	Weather	Topography							Ens
✓	✓	✓	✓		71.3	69.5	**73.2**	72.8	71.2	63.6
-	✓	✓	✓		46.1 (54.7)	60.8 (14.3)	**65.2** (12.3)	63.3 (15.0)	59.5 (19.8)	55.5 (14.6)
✓	-	✓	✓		69.8 (2.1)	68.1 (2.0)	69.9 (4.7)	**71.4** (2.0)	70.2 (1.5)	60.0 (6.0)
✓	✓	-	✓		70.5 (1.1)	69.1 (0.7)	72.0 (1.7)	73.0 (†)	**73.1** (†)	66.3 (†)
✓	✓	✓	-		71.3 (0.0)	69.5 (0.0)	**73.1** (0.1)	72.8 (0.0)	71.2 (0.1)	65.5 (†)
✓	-	-	-	72.8	69.1	67.6	69.3	71.8	**72.5**	70.8
-	✓	-	-	55.8	17.4	46.9	51.6	53.7	54.8	**56.5**
-	-	✓	-	46.9	15.7	42.0	42.8	44.8	45.0	**47.4**
-	-	-	✓	28.0		15.3	15.1	3.1	14.8	**30.0**
Average				50.9	53.9	56.5	**62.5**	58.5	59.1	57.3

(weighted) F1 Results I
Tree-species recognition

$$(P_A - P_{m(s)})/P_{m(s)},$$

Tree-species classification

Optical	Radar	Aerial	Single	FSensD	OOD-f	TIMML	ESensI
✓	✓	✓		61.8	**68.0**	65.6	62.6
-	✓	✓		59.3 (4.1)	**65.3** (4.2)	64.7 (1.4)	63.0 (†)
✓	-	✓		61.1 (1.1)	**66.1** (3.0)	65.2 (0.7)	65.6 (†)
✓	✓	-		57.5 (7.4)	27.1 (150.8)	**59.3** (10.7)	56.4 (10.9)
✓	-	-	65.8	53.5	5.5	**55.5**	55.3
-	✓	-	55.0	51.3	10.4	**52.6**	52.0
-	-	✓	63.7	56.0	62.2	**63.3**	62.2
Average			61.5	57.2	43.5	**60.9**	59.6



Differences between the best individual prediction and complementarity

Main findings

- *The individual-source performance* ***does not necessarily*** *relate to their complementary effectiveness in the task.*
  - Depends on the common multi-source information

- *Additional data sources* ***are not always beneficial*** *in the modeling.*
  - Depends on how task-informative the source is, complementarity, and model design

- *A multi-source model is* ***not robust for all cases*** *of source-wise missing data.*
  - Specially in single-source predictions (harder).
  - Depends on the model design.

It depends!

Analysis

Effectiveness of **model robustness** to missing data sources depends on several factors.

A. How **task-informative** is the individual data source.
   a. *Clue: single predictive performance*

B. How **complementary** is the data source regarding others.
   a. *Clue: complementarity score*

C. How **invariant** is the model design to missing data.

Conclusion

- Robustness evaluation of various multi-source models in the literature.

- We identified **three** main factors on which robustness depends on.

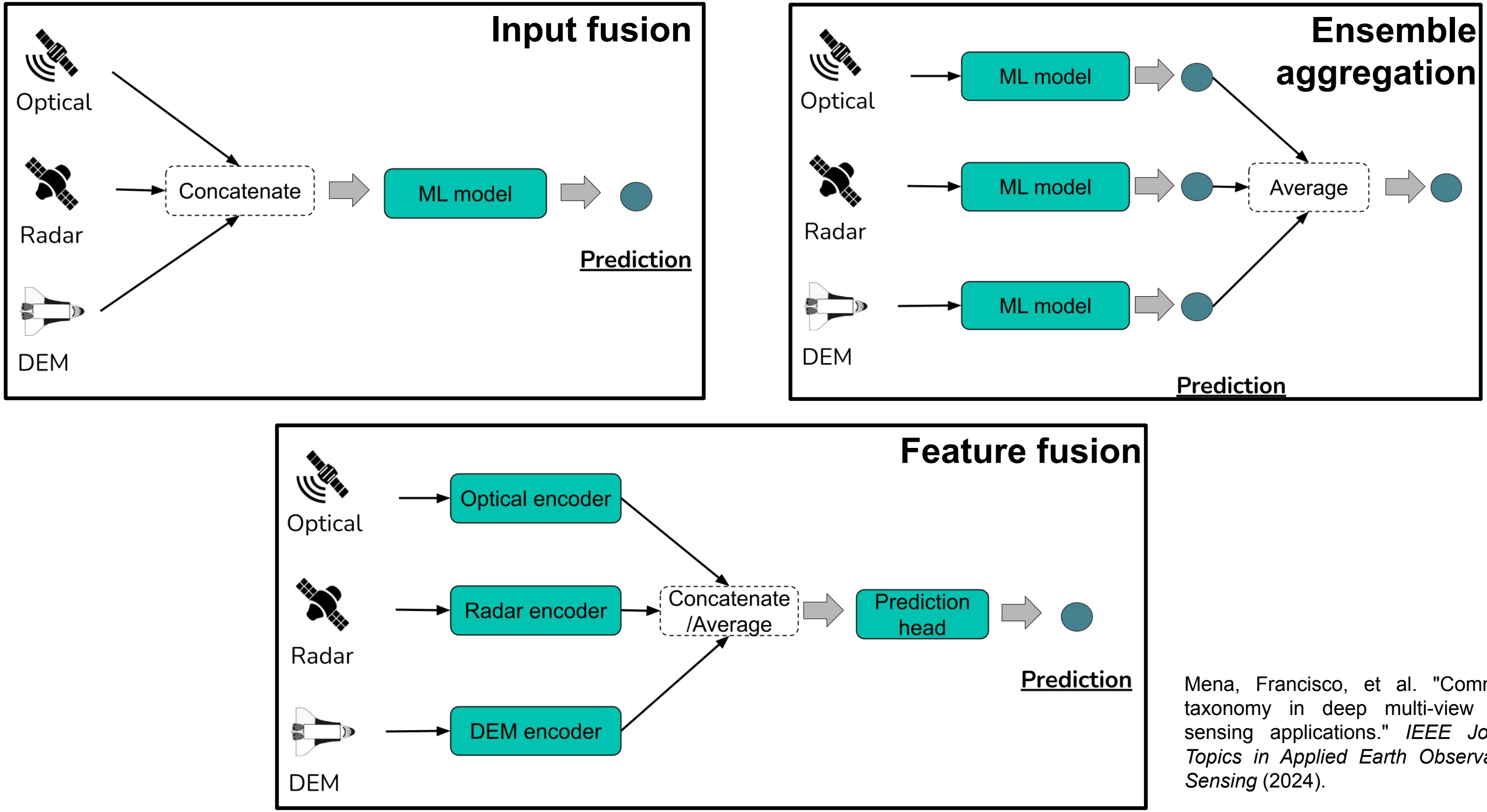
Future work: Explainability and quantifying the contribution/importance of each data source.

On what depends the robustness of multi-source models to missing data in Earth observation?

<u>Francisco Mena</u> (1,2), Diego Arenas (1),
Miro Miranda (1,2), Andreas Dengel (1,2)

1 University of Kaiserslautern-Landau (RPTU)
2 German Research Center for Artificial Intelligence (DFKI)

Fusion strategies



Mena, Francisco, et al. "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* (2024).

Model implementation

- Architecture
  - **Encoders**: Two layers with 128 units
    - Static pixel-wise sources: MLP
    - Static image sources: ResNet-50
    - Temporal views: 1D CNN
  - **Prediction Head**: a MLP with one layer of 128 units.
- Optimization:
  - ADAM
  - 128 batch size
  - early stopping
- Loss function
  - Cross entropy