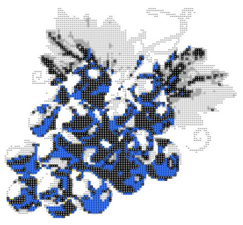
4th Edition

Data Research meetup by MagIC



Validation of Early Vineyard Yield Estimation

Comparing Leave-One-Year-Out and Walk-Forward Back-Testing in the Douro Demarcated Region

Diogo Santos Costa (diogo.costa@novaims.unl.pt)

NOVA Information Management School (NOVA IMS),
Universidade Nova de Lisboa, Campus de Campolide,
1070-312 Lisboa, Portugal

1. INTRODUCTION

Reliable early-season wine-grape yield estimation supports harvest logistics, market planning, quota interpretation and regulatory decision-making in regions with strong interannual variability. This is particularly relevant in the **Douro Demarcated Region (DDR)**, north-eastern Portugal: the world's oldest demarcated wine region, covering about 250,000 ha, with approximately 43,000 ha of productive vineyards across 169 parishes **and three sub-regions: Baixo Corgo, Cima Corgo and Douro Superior**. Its steep, terraced terrain makes automated estimation particularly difficult, while its strong interannual variability makes early estimation especially valuable for planning (Barriguinha et al., 2022).

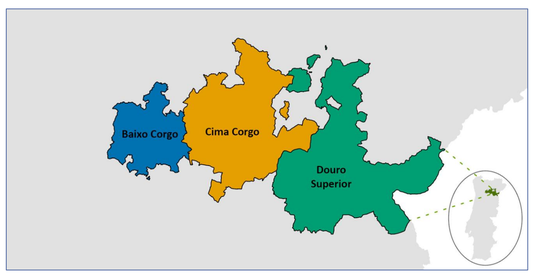


Fig.1: DDR and Sub-Regions.

At regional scale, operational yield-estimation examples remain scarce and have often relied on pollen or climate-based models (Barriguinha et al., 2021). Recent LSTM pipelines combining Sentinel-derived NDVI and climate data can estimate parish-level yield before harvest (Barriguinha et al., 2022), but their operational interpretation depends on temporally realistic validation, a weakness still common in vineyard Remote Sensing studies (Santos Costa et al., 2026a).

This study therefore aims to:

- evaluate an LSTM pipeline using open-access Earth Observation and climate predictors at **flowering (FLO)** and **veraison (VER);**
- compare **Leave-One-Year-Out (LOYO)** and **Walk-Forward (WF)** validation under identical modelling conditions; and
- assess how validation design changes model-readiness interpretation from parish to sub-region and DDR levels.

2. METHODS AND MATERIALS

MATERIALS. Each sample is a parish-year. The target variable is corrected wine-grape yield (kg/ha) for the 169 DDR parishes, obtained from IVDP records. Predictors are five-day sequences aligned with Sentinel-2 observations: NDVI, a missing-data flag, and open-access climate variables (air temperature, relative humidity and wind from AgERA5; precipitation from CHIRPS). These data provide an open alternative to local weather-station inputs while preserving the predictor logic of the published baseline (Barriguinha et al., 2022; Santos Costa et al., 2026b).

MODEL. Preprocessing is leakage-aware: imputation, standardisation and scaling are fitted on training data only, and padded time steps are masked. The network is intentionally compact: *Masking → LSTM(16) → Dense(1)*, trained with Adam, MSE loss and early stopping. Because the model is fixed, performance differences are attributed to validation design rather than architecture.

EVALUATION. Results are reported at parish, sub-region and DDR levels using MAE, bias, AAE (Average Absolute Error) (%) and the share of parishes below 20% AAE.

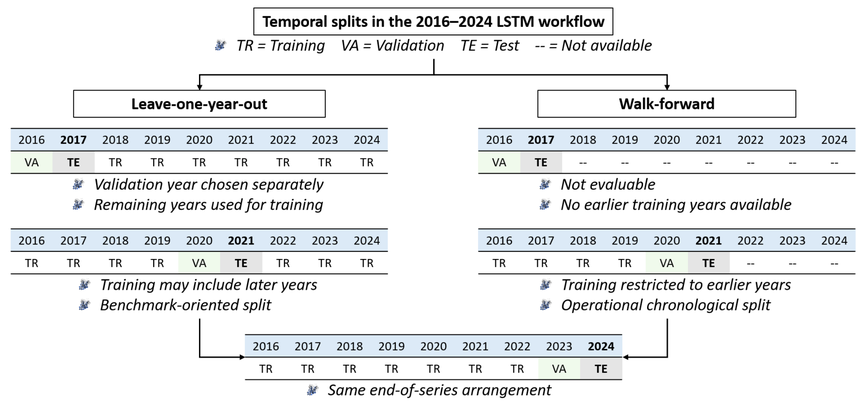


Fig.2: Temporal validation designs. LOYO holds out one test year and may train on later years, supporting benchmark-oriented comparison. WF is stricter: validation uses year t−1, training uses only years before t−1, and testing is performed on year t. To ensure comparability, both designs are evaluated over 2018–2024.

3. RESULTS & DISCUSSION

Research Question *Does a stricter, operationally realistic validation design reduce the measured accuracy of the same early-yield pipeline?*

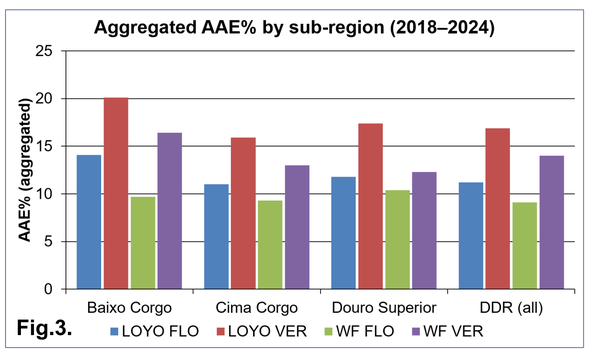


Fig.3.

Tab.1: Summary metrics by design and window (2018–2024).

Design	Window	MAE (kg/ha)	Parish AAE%	<20% AAE	DDR AAE%
LOYO	FLO	896	27.6	40.2%	11.2
LOYO	VER	1030	32.1	23.7%	16.9
WF	FLO	866	25.9	50.9%	9.1
WF	VER	912	25.9	41.4%	14.0

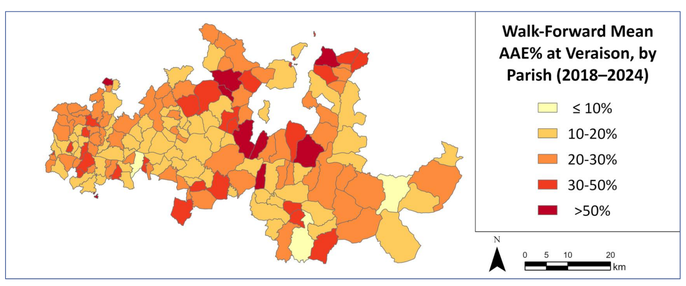


Fig.4: Walk-Forward mean AAE% at veraison, by parish (2018–2024)

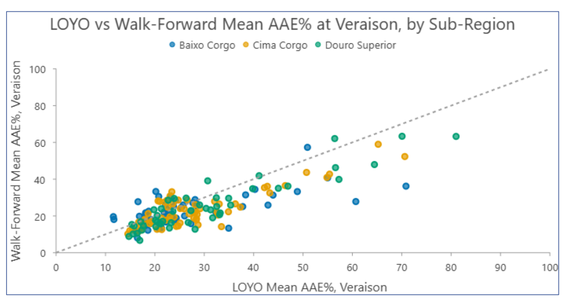


Fig.6: AAE% (LOYO vs Walk-Forward).

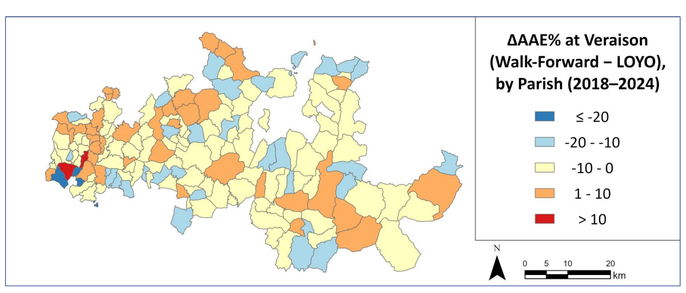


Fig.5: ΔAAE% at veraison (Walk-Forward − LOYO), by parish; negative = lower error under WF.

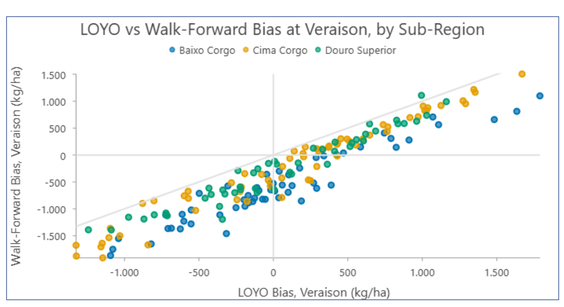


Fig.7: Bias (LOYO vs Walk-Forward).

KEY FINDINGS

- WF did not reduce measured accuracy. Under the matched comparison, WF slightly outperformed LOYO, especially at veraison, with the advantage also holding after aggregation to sub-region and DDR level.
- Parish-level AAE% at VER fell from 32.1% (LOYO) to 25.9% (WF); at DDR level, VER AAE% fell from 16.9% to 14.0%.
- Operationally acceptable parishes (below the 20% AAE threshold) increased under WF, reaching 50.9% at FLO and 41.4% at VER.
- The WF advantage was consistent across the three sub-regions, though parish-level differences were spatially uneven.

INTERPRETATION

- WF does not shift all errors downward; it rearranges performance spatially; some parishes improve while others remain problematic. This suggests the NDVI-plus-climate core already suffices in part of the DDR, but that additional predictors may be needed in more difficult areas.
- Under WF, VER bias became more negative (−337 kg/ha mean parish-year bias, 2018–2024), indicating mild underestimation, relevant for regulatory use, where systematic bias affects interpretation.
- LOYO and WF answer different questions: LOYO suits benchmark comparison, WF is the more credible deployment-oriented test. The result is encouraging but should be read cautiously given the limited number of independent target years.

4. CONCLUSION

Early yield estimates are only as trustworthy as the validation design used to test them. This study held an open-data LSTM pipeline fixed and varied only the temporal validation strategy, asking whether a stricter, operationally realistic scheme would reduce measured accuracy. It did not: over the matched 2018–2024 window, Walk-Forward back-testing matched and slightly improved on the conventional Leave-One-Year-Out benchmark, particularly at veraison and after aggregation from parishes to sub-regions and the whole DDR.

This matters because it suggests that the reported skill is not simply an artefact of allowing the training set to include later seasons. Instead, the pipeline remains credible under the chronological constraints of real forecasting. The finding refines, rather than contradicts, the published DDR baseline (Barriguinha et al., 2022), showing that its performance survives a more demanding validation test.

For regional governance, the results support the use of the pipeline as an aggregated decision-support tool for early production signals, while single-parish estimates should still be interpreted cautiously. The evidence is promising rather than definitive, since the independent target-year horizon remains short. Future work should extend the temporal record and enrich the predictor set in areas where veraison errors remain spatially uneven.

REFERENCES

- Barriguinha, A., de Castro Neto, M., & Gil, A. (2021). Vineyard yield estimation, prediction, and forecasting: A systematic literature review. Agronomy, 11(9), 1789. https://doi.org/10.3390/agronomy11091789
- Barriguinha, A., Jardim, B., de Castro Neto, M., & Gil, A. (2022). Using NDVI, climate data and machine learning to estimate yield in the Douro wine region. International Journal of Applied Earth Observation and Geoinformation, 114, 103069. https://doi.org/10.1016/j.jag.2022.103069
- Santos Costa, D., Barriguinha, A., Areosa, I., Caetano, M., Teixeira, I., & Vanneschi, L. (2026a). Remote sensing-based approaches for automatic vineyard area identification: A systematic review. Smart Agricultural Technology, 13, 101812. https://doi.org/10.1016/j.atech.2026.101812
- Santos Costa, D., Barriguinha, A., Areosa, I., Caetano, M., Teixeira, I., & Vanneschi, L. (2026b). The impact of replacing local meteorological data with open-access grid-format climate products for yield estimation in the Douro Demarcated Region (2016–2021) [Conference poster]. Zenodo. https://doi.org/10.5281/zenodo.19094145

ACKNOWLEDGEMENTS

The author thanks his **PhD supervisors**, **Prof. Miguel de Castro Neto** and **Prof. André Barriguinha** (NOVA IMS), and the **IVDP for data support**. This study is part of an ongoing PhD at NOVA IMS, continuing the concluded AIGODS project (https://aigods.pt/) 2024.07551.IACDC funded by FCT,I.P. (Fundação para a Ciência e a Tecnologia,I.P.) (DOI: 10.54499/2024.07551.IACDC). This work was supported by national funds through FCT,I.P. (Fundação para a Ciência e a Tecnologia,I.P.), under the project - UID/04152/2025 - Centro de Investigação em Gestão de Informação (MagIC)/NOVA IMS (DOI: 10.54499/UID/04152/2025, and by the European Union – NextGenerationEU under the project UID/PRR/04152/2025 (DOI: 10.54499/UID/PRR/04152/2025).

ABOUT THIS POSTER

How to cite? Santos Costa, D. (2026). Validation of Early Vineyard Yield Estimation: Comparing Leave-One-Year-Out and Walk-Forward Back-Testing in the Douro Demarcated Region. Poster session presented at Data Research Meetup by MagIC, 4th ed. Lisboa, Portugal. https://doi.org/10.34619/jlrs-lqw1

Keywords: Vineyard yield estimation; LSTM neural networks; Model evaluation.

Funded by: Fundação para a Ciência e a Tecnologia; PRR Plano de Recuperação e Resiliência; República Portuguesa; Financiado pela União Europeia NextGenerationEU

NOVA Information Management School, 17 July 2026

ONLINE VERSION

ArcGIS STORYMAPS VERSION