Using artificial intelligence to monitor live fuel moisture content across France, based on a high-resolution land surface analysis

Referee #1 (Dr. Krishna Rao)

Commentary 1.1: Cross Species LFMC Averaging. While I recognize the constraints imposed by available data on species distribution, it is unclear why LFMC was averaged across species, especially given the authors' own observation that "the species are too varied" (L. 90). Extensive literature demonstrates significant inter species variation in LFMC responses under identical climate conditions (e.g., Martin StPaul et al., 2018; Pimont et al., 2019; Yebra et al., 2024). Could an existing land cover or species map (e.g., Theia or IGN Tree Species Maps) be used to identify the dominant species at a given sampling site and isolate the corresponding LFMC time series? Please clarify the rationale for averaging across species, or discuss the uncertainty this introduces into the downstream model targets.

Response 1.1:

We agree with the reviewer that LFMC responses can differ substantially between species, even when environmental conditions are identical. This has been demonstrated by Martin-StPaul et al. (2018), Pimont et al. (2019) and Yebra et al. (2024). Species-specific LFMC modelling is therefore important. However, the design of this study was to estimate LFMC from the ISBA land surface model, which is based on plant functional types (PFTs) rather than individual species. The operational configuration used here relies on a 12-PFT classification, in which the relevant vegetation type (PFT No. 4 in the model) represents a broad aggregation of temperate broadleaf vegetation and shrublands. Consequently, it would not be consistent with the available information in the predictor variables to establish species-specific targets from the observations. Furthermore, the ONF observation network is heterogeneous with respect to species composition. Some sites contain measurements for a single species, while others include four or five species. Restricting the analysis to one species per site or selecting a dominant species from external maps would reduce the spatial consistency of the reference dataset and introduce assumptions about the selected species' representativeness with respect to the modelled PFT.

In summary, our objective is to provide an operational LFMC product for assessing wildfire danger. Given the current vegetation representation of ISBA, an aggregated LFMC that is representative of the vegetation present within a model grid cell is more relevant than species-specific estimates. To clarify this point, we will expand on this in the manuscript and explicitly acknowledge that averaging across species introduces uncertainty into the target variable. This is likely to contribute to some of the variance that remains unexplained by the model. Additionally, the use of WFC metadata dependent on soil properties in the LAI/WFC predictor could indirectly represent species distribution, given that plant species can vary according to soil properties (Wan et al., 2019). In line with the reviewer's comments, ongoing developments to ISBA are moving towards a more detailed vegetation description based on 20 vegetation types. This will allow LFMC products to be developed that are associated with more specific vegetation functional types, thereby reducing the need for cross-species averaging..

Commentary 1.2: Physical Mechanism Linking LAI/WFC to LFMC. Table 2 shows that the top feature linked to LAI is LAI/WFC (LAI normalized by soil water field capacity), with Section 4.6 providing details of a plausible mechanism where leaf growth under limited soil capacity leads

to competing forces on plant water status, thus lowering live fuel moisture. While this mechanism is plausible for certain conditions, does it hold universally across different plant functional types? How does this relationship perform in drought adapted ecosystems or for species with deep root systems that draw water well below the topsoil layer (i.e., deeper than the depth of 10 cm used here)? For species with internal stem or leaf water storage or lower sensitivity to immediate soil water deficits, is LAI/WFC still an adequate proxy? In Section 4.4, I see mention of other indices including microwave data, but a brief discussion about the potential limitations of relying primarily on canopy leaf volume and shallow soil water parameters would help. I am not familiar with the shrub types in France, so feel free to point me (and the manuscript's readers) to existing literature if such variability does not exist in France.

Response 1.2:

We thank the reviewer for this insightful comment. We agree that the mechanism proposed in Section 4.6 should not be interpreted as universally valid across all plant functional types and ecological contexts. Although the LAI/WFC ratio was empirically identified as one of the most informative predictors for the French operational dataset, it remains a proxy that reflects the balance between canopy development and the water-holding capacity of the local soil, as represented in ISBA. With regard to drought-adapted ecosystems and deep-rooted species, we acknowledge that some Mediterranean shrubs and woody species can access water stored in deeper layers of the soil than those considered in the present analysis. In such cases, a shallow-soil proxy alone may not accurately represent plant water availability. However, the model does not rely solely on LAI/WFC. It also incorporates variables related to soil moisture dynamics, biomass and phenology based on meteorological forcing, which partly account for differences in drought response. Similarly, species with substantial internal water storage or lower short-term sensitivity to soil moisture deficits may exhibit weaker relationships between LFMC and LAI/WFC. Consequently, we do not interpret this variable as a direct physiological descriptor of plant water status, but rather as an integrated indicator that captures a significant proportion of the observed LFMC variability within the vegetation functional types represented by the operational ISBA configuration.

To address this issue, we propose adding a paragraph to the Discussion section to clarify that the predictive importance of LAI/WFC varies among vegetation types. This is particularly pertinent in Mediterranean shrublands and deep-rooted woody ecosystems, where access to deeper soil water reserves and internal water storage mechanisms can decouple LFMC from shallow soil moisture conditions.

Commentary 1.3: Derivation of the "Interest Score". The "Interest Score" concept presented in Figure 10 is creative and practical for future sampling strategy design. However, the methodology behind it is insufficiently detailed. How were the weighting coefficients (alpha, beta, and gamma) determined? I only see in the supporting section 7 that they were "empirically set". What does that mean? What are the exact values assigned to these coefficients? Was any ground truth or validation heuristic used (e.g., constraining the score toward zero at existing, well sampled monitoring stations)?

Response 1.3:

We thank the reviewer for pointing out that the methodology behind the Interest Score was insufficiently described. We will expand Section 7 to explicitly define the score, provide the exact weighting coefficients, and clarify how they were selected.

The Interest Score is a heuristic criterion designed to identify potential locations for future LFMC monitoring sites. Its purpose is to illustrate a practical framework for prioritising these locations, rather than providing a uniquely optimal site-selection criterion. It combines three complementary

normalised indicators: (i) the location's propensity to experience low LFMC values, quantified by the 10^(th) percentile of simulated LFMC over time; (ii) the variability of LFMC over time, represented by its standard deviation; and (iii) the distance to the nearest existing ONF monitoring site, which improves the spatial representativeness of the observation network. Each component is normalised to the interval [0, 1], and the final score is computed as $I = \alpha S + \beta V + \gamma D$, where S is the normalised dryness index, V is the normalised LFMC variability index and D is the normalised distance-to-network index. In the present study, the following coefficients were chosen: $\alpha=0.2$, $\beta=0.4$ and $\gamma=0.4$. The greater importance given to variability and distance reflects the main objective of identifying locations that provide additional information to the existing monitoring network and exhibit informative LFMC dynamics, while still prioritising areas with low LFMC values. As these coefficients were not calibrated against an independent reference dataset, they should be regarded as heuristic rather than optimal. Depending on the intended monitoring strategy, future work could investigate the objective calibration or multi-objective optimisation of these coefficients. No explicit validation dataset or optimisation criterion was employed. Existing ONF stations were only incorporated through the distance component of the score, which discourages the selection of locations that are already well represented by the current monitoring network. For this reason, the Interest Score should be viewed as a heuristic decision-support tool rather than a validated optimisation metric.

Commentary 1.4: Linking LFMC to Ignitions (Figure 9). Figure 9 presents an intriguing relationship between simulated LFMC and ignition timing, but several critical details and controls are missing:
Dataset & Terminology: Please clarify what constitutes a "fire outbreak" versus an "ignition" (L. 380 vs. Figure 9 axis). Where did you obtain the fire outbreak from? Is this based on the Prométhée database? Lack of Counterfactual Control: L. 379 states that this analysis tests whether simulated LFMC aligns with actual ignitions. However, because the evaluated ignitions occur during summer (June to September), LFMC naturally declines during this window. To prove that LFMC carries an ignition specific signal, I think a counterfactual baseline is required: Suggested Test: For each ignition in Canton A on Day T, select a paired control Canton B that had no fire outbreaks on Day T (or earlier in that fire season). Calculate the canton averaged LFMC for T-10 to T+4 days for Canton B. Repeat for each ignition and produce a similar box plot for the counter factual case. Does the non fire control curve look substantially different from Figure 9? If both curves look identical, the trend reflects seasonal drying rather than an ignition specific LFMC threshold. Post Ignition LFMC Trajectory: Figure 9 shows that LFMC remains relatively flat following Day 0. In many burned landscapes, post fire burn scar pixels experience a sharp drop in visible canopy moisture. Is the lack of a post ignition drop an artefact of spatial averaging over an entire canton? Even with spatial averaging, one might expect at least a slight post fire decline relative to pre ignition levels. Spatial Aggregation Level: Why were LFMC values aggregated across entire cantons rather than extracted at the specific 2 km times 2 km ignition pixel or over a fixed spatial radius (e.g., 5 km)? Because cantons vary dramatically in size (from 4 km^2 to 1000 km^2), this introduces spatial heterogeneity. Is the underlying ignition dataset restricted to canton level resolution? As I write this, I acknowledge that this depth of analysis for an experiment that is not the central focus of the paper might be considered too much. But I think half-results are more dangerous that no results. Since there is already plenty of other contributions in the paper, I leave it to the authors to decide whether they include full details of the analysis and make the experiment more robust or leave it out from the paper and follow up with a future study.

<u>Response 1.4:</u>

We agree that the terminology was ambiguous. The term 'fire outbreak' refers to a wildfire event that has been recorded in the French BDIFF (Base de Données sur les Incendies de Forêt en France) database. This database contains details of all reported wildfires, including their occurrence date, administrative location and burned area. When available, it also includes information on the reported cause (e.g. natural or human-induced). We have clarified both the terminology and the data source in the manuscript. We agree that the analysis presented in Figure 9 does not demonstrate that LFMC alone explains fire occurrence. Rather, it illustrates the temporal evolution of simulated LFMC around reported fire outbreaks. The counterfactual analysis proposed by the reviewer would provide a more robust evaluation of whether the observed signal differs from the expected seasonal drying during the summer months. We thank the reviewer for this valuable suggestion. We will consider it and clarify that Figure 9 shows the temporal consistency between the simulated LFMC and reported fire outbreaks, rather than providing evidence of an ignition-specific LFMC threshold. As the LFMC in Figure 9 is simulated using the ISBA land surface model, which does not explicitly represent fire-induced vegetation removal or post-fire canopy changes, a significant decrease in LFMC after Day 0 is not anticipated. Consequently, in the absence of new observations, the model continues to simulate the seasonal evolution of vegetation after the fire event, with no abrupt decrease in LFMC anticipated. However, the ISBA system assimilates satellite-derived LAI observations. This means that a reduction in canopy state variables, and consequently in simulated LFMC, could occur if observations were available during or shortly after the fire event (e.g. within the following 10 days). Therefore, the magnitude of this response depends on the timing and availability of satellite observations relative to the occurrence of the fire. We will add a clarification on this point in the revised manuscript. We chose to aggregate LFMC at canton level because the BDIFF database does not provide precise geographical coordinates for the point of ignition. The only information available is the associated administrative location (municipality), together with event characteristics such as the burned area and reported cause. Aggregating LFMC over the corresponding canton ensured that the reported event was included within the analysed area. We acknowledge that canton sizes vary substantially, which inevitably masks local variability in LFMC. This limitation will be discussed in more detail in the revised manuscript. Analyses based on higher-resolution ignition locations could be considered when such data becomes available.

Minor Comments:

- Figure 12: The current layout makes it difficult to quantitatively compare feature importances. While the larger size of the LAI rhombus implies higher importance, it is hard to discern the exact impact. For e.g., if one were introduced in learning how much performance metrics degrade when LAI/WFC is neutralized, how would they find out? Does the figure show that? I see the 1 - KGE = 1.0 for LAI/WFC. Does that mean upon neutralization, KGE becomes zero? Surely, that's not right?
- L. 401: The text refers to "Supplementary Section 8" for the impact of the day of the year analysis, but this appears to correspond to Supplement 8 in the supplementary material. Please verify and correct the section reference.
- L. 10 (Abstract): "The results demonstrate a strong ability to estimate the range and dynamics of LFMC across most of France" The term "strong ability" is subjective. Please replace or support this statement with concrete performance metrics such as RMSE or NSE.
- L. 71 to 85: I appreciate the authors for including a thorough and clear description of the field sampling protocols. This was very helpful in understanding the labels for the neural

net.

Response 1.5:
Thank you for noting this. It will be corrected in the revised version of the paper.

References

Martin-StPaul, N., Pimont, F., Dupuy, J.L. et al. Live fuel moisture content (LFMC) time series for multiple sites and species in the French Mediterranean area since 1996. Annals of Forest Science 75, 57 (2018). https://doi.org/10.1007/s13595-018-0729-3

Pimont, F., Ruffault, J., Martin-StPaul, N. K., & Dupuy, J. L. (2019). Why is the effect of live fuel moisture content on fire rate of spread underestimated in field experiments in shrublands?. International journal of wildland fire, 28(2), 127-137.

Yebra, M., Scortechini, G., Adeline, K. et al. Globe-LFMC 2.0, an enhanced and updated dataset for live fuel moisture content research. Sci Data 11, 332 (2024). https://doi.org/10.1038/s41597-024-03159-6

Wan, J.-Z., Yu, J.-H., Yin, G.-J., Song, Z.-M., Wei, D.-X., Wang, C.-J.: Effects of soil properties on the spatial distribution of forest vegetation across China, Global Ecology and Conservation, 18, e00635, https://doi.org/10.1016/j.gecco.2019.e00635, 2019.