

Comments on “Drought Risk Mapping and Parametric Insurance in Agriculture: A Machine Learning-Based Framework” by Belhsen et al., (2026), Journal of Risk Analysis and Crisis Response, <https://doi.org/10.54560/jracr.v16i2.726>

Haidong Li¹

¹ Faculty of Geosciences and Engineering, Southwest Jiaotong University, Chengdu, China

Received: October 16, 2026; Received in revised form: October 25, 2026; Accepted: October 25, 2026; Available online: October 26, 2026

Abstract: Accurate assessment of drought-related crop losses is essential for designing effective agricultural index insurance. An examination of the datasets and methods reported for parametric drought insurance in three Moroccan agricultural regions highlighted three issues: temporal data matching, drought index standardization, and independent evaluation of basis risk reduction. The reported correlation significance levels appear inconsistent with the stated annual sample size, while recurrent extreme drought index values require further diagnostic checks. Basis risk metrics restricted to loss years and premium point estimates also provide an incomplete assessment of contract performance. The framework integrates environmental indicators and actuarial methods, but its operational reliability depends on clarifying these issues and independently comparing it with simpler insurance contracts.

1. Introduction

We would first like to congratulate the authors (Yousra Belhsen, Rim Ouhdouch, and Khalil Said) on the publication of their article entitled “Drought Risk Mapping and Parametric Insurance in Agriculture: A Machine Learning-Based Framework” in the Journal of Risk Analysis and Crisis Response, 16(2), 217–253 (2026) [1]. The study bridges the gap between drought monitoring and the design of locally calibrated agricultural insurance by integrating composite index construction, payout threshold calibration, and premium estimation into a common framework. Applying it to three Moroccan regions advances index insurance development in data-limited agricultural systems.

Here, we focus on three issues affecting the evidence for reduced basis risk: temporal matching and sample independence, drought index standardization, and independent contract evaluation. The numerical checks below use published statistics and do not reproduce the analysis of the underlying observations.

2. Materials and Methods

The temporal unit of analysis needs clarification. Sections 5 and 6 of the original study describe annual observations over fourteen years, whereas Section 3.3 describes a monthly dataset containing yield measurements. To examine the implications for statistical inference, we recalculated the two-sided Pearson significance levels for the Normalized Difference Vegetation Index (NDVI)–yield correlations reported in Table 3 [1], assuming fourteen independent annual pairs per region. For each

published correlation coefficient r , the test statistic was calculated using Eq. (1), with $n=14$ and $n-2=12$ degrees of freedom. The published coefficients were retained at their reported precision. This calculation examines the consequences of an assumed annual sample size; it does not reconstruct annual correlations from the underlying observations.

$$t = r\sqrt{[(n-2)/(1-r^2)]} \quad (1)$$

The seasonal standardization of the Standardized Precipitation Index (SPI) and Standardized Precipitation Evapotranspiration Index (SPEI) also needs a reproducible description. Section 3.2 explains fitting distributions to three-month accumulations but does not specify whether fitting was performed separately for each ending calendar month. Without seasonal conditioning, normal seasonal dryness can be confused with drought anomalies. The reference period and fitting procedure should therefore be reported explicitly [2, 3]. WMO guidance recommends a longer precipitation record, preferably at least thirty years, while recognizing the limitations of shorter records [2]. A fourteen-year yield record does not require restricting the climatic reference to the same period.

Independent evaluation of the complete calibration procedure is not documented. Table 5 [1] reports predictive performance without specifying a training–test separation or cross-validation design. Selecting the index and calibrating payout thresholds on the observations subsequently used for evaluation can produce optimistic estimates [4]. For a prospective application, calibration on earlier years and testing on later years, with a rolling design where feasible, would provide a more relevant assessment. Observations from all regions in each withheld year should remain together to limit dependence between calibration and evaluation samples [5]. Yield-dependent model selection and contract calibration should be repeated within the training data.

3. Results and discussion

Figure 1 compares the published significance levels with the conditional recalculations. For Marrakech-Safi, the reported NDVI–yield correlation ($r = 0.187$) gives $p = 0.5221$ under the assumption of fourteen independent annual pairs, compared with the published $p = 0.0154$. For Oriental, $r = 0.158$ gives $p = 0.5896$ under the same assumption, compared with the reported $p = 0.0403$. Neither coefficient would therefore reach the conventional 0.05 significance threshold under this annual-sample assumption. These differences warrant clarification of the actual number of paired observations and their temporal matching. If annual yields were repeated across monthly records, those records would not supply additional independent yield realizations. The comparison alone does not establish how the original significance levels were obtained.

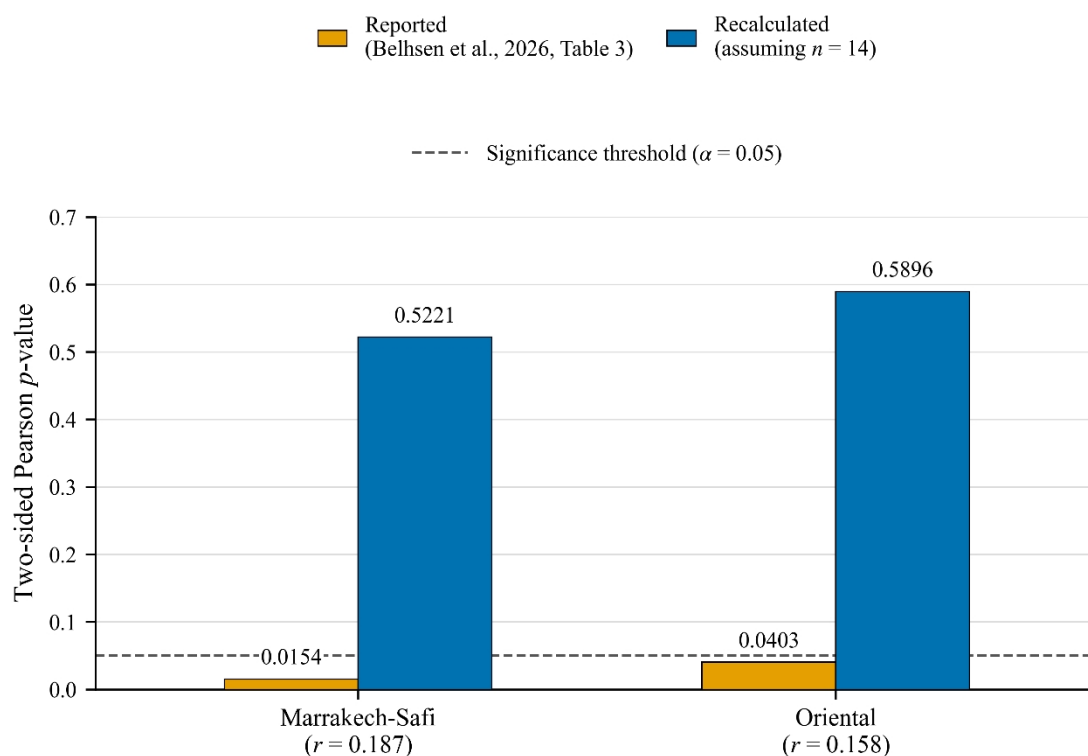


Figure 1. Comparison of reported and conditionally recalculated p-values for NDVI–yield correlations. Reported values are from Table 3 of Belhsen et al. (2026) [1]. Recalculated values use a two-sided Student's t-test, assuming 14 independent annual pairs per region ($df = 12$). The horizontal dashed line indicates the significance threshold $\alpha = 0.05$.

Interpretation of the extreme SPEI values illustrates why index standardization requires verification. Figure 5 [1] contains recurrent values below -8, which Section 3.3.6 attributes to acute drought stress. The same section describes SPI and SPEI as theoretically bounded by -3 and +3, although the inverse-normal transformations in Eqs. (2) and (4) [1] have no finite theoretical bounds. Under a correctly calibrated standard normal interpretation, a value of -8 corresponds to a lower-tail probability of approximately 6.2×10^{-16} . Repeated occurrences within a short record warrant checks of seasonal conditioning, distribution support, and the numerical treatment of near-zero probabilities. A nonsignificant Kolmogorov–Smirnov test and an overall cumulative distribution plot cannot establish accuracy in such an extreme tail. Reporting the corresponding water-balance observations and fitted probabilities would help determine whether these values represent climatic anomalies or sensitivity in the standardization procedure.

The claimed reduction in basis risk also requires a direct comparison. Eqs. (25) and (26) [1] condition on positive agricultural losses and consequently exclude payouts in years with no defined loss. These metrics describe discrepancies during loss years, but an assessment of contract performance should also include all-year payout–loss errors, missed losses, and payouts without losses. Let $L \geq 0$ denote agricultural loss and $I \geq 0$ the corresponding insurance indemnity, expressed on the same monetary basis. The unconditional joint probabilities $P(I > 0, L=0)$ and $P(I < L, L > 0)$ characterize false-claim (Type I) risk and shortfall (Type II) risk, respectively. The latter includes both complete nonpayment and partial undercompensation in loss years. These probabilities should be evaluated across all years and supplemented by measures of the magnitude of payout–loss

discrepancies. Table 7 [1] reports residual relative basis risk values of 45.18%, 17.57%, and 37.61% for the selected contracts. These values alone do not establish a reduction relative to simpler designs. Contracts based on individual indices or the fixed-weight benchmark should be evaluated under comparable coverage conditions and on the same withheld years. Such a comparison would show whether the integrated framework improves protection beyond its fit to the calibration data.

Premium sensitivity forms part of the same validation question. In Oriental, the kernel density estimation (KDE) premium of 176.18 MAD/ha is approximately 37.4% higher than the normal Monte Carlo estimate of 128.27 MAD/ha. This difference warrants qualification of the claimed robustness to distributional assumptions. Increasing the number of simulated draws does not increase the information contained in fourteen observed crop seasons. Confidence intervals reflecting uncertainty in both the historical sample and the fitted index and thresholds would provide a more informative assessment of premium stability than point estimates alone.

4. Conclusions

The proposed framework offers a useful connection between drought monitoring and agricultural insurance design. Its central claim of reduced basis risk nevertheless requires clearer temporal matching, verification of drought index standardization, and independent evaluation of the resulting contracts. Clarifying the sample units, diagnosing extreme SPEI values, and comparing alternative contracts on withheld crop seasons would directly address these concerns. Until these checks are available, the reported results support an exploratory insurance design whose predictive performance and risk reduction require further validation.

CRedit authorship contribution statement

Haidong Li: Writing – original draft, Methodology, Investigation, Conceptualization.

Declaration of Competing Interest

The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Belhsen, Y., Ouhdouch, R. & Said, K. Drought Risk Mapping and Parametric Insurance in Agriculture: A Machine Learning-Based Framework. *Journal of Risk Analysis and Crisis Response* **16**, 217-253, doi:10.54560/jracr.v16i2.726 (2026).
- [2] World Meteorological Organization. (2012). Standardized Precipitation Index User Guide (M. Svoboda, M. Hayes, & D. Wood, Eds.; WMO-No. 1090). Geneva: WMO. <https://library.wmo.int/idurl/4/39629>
- [3] Beguería, S., Vicente-Serrano, S. M., Reig, F., & Latorre, B. (2014). Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, *34*, 3001–3023.<http://doi.org/10.1002/joc.3887>.
- [4] Cawley, G. C., & Talbot, N. L. C. (2010). On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. *Journal of Machine Learning Research*, *11*, 2079–2107. <https://www.jmlr.org/papers/v11/cawley10a.html>.
- [5] Roberts, D. R., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*, 913–929.<http://doi.org/10.1111/ecog.02881>.