From Global Framework to Local Data: Operationalizing Agricultural Loss and Damage Accounting in Viet Nam

Parth Khare, Data Lab-ESS, FAO, parth.khare@fao.org
Nguyen Ngoc Luan FAO, Vietnam Government, luan.nguyen@cap.gov.vn

Carola Fabi, Senior Statistician, Data Lab-ESS, FAO, carola.fabi@fao.org
Christian Mongeau, Data Lab-ESS, FAO, Christian.mongeau@fao.org

Abstract

Climate-related disasters such as droughts and floods cause significant agricultural losses worldwide, threatening food security. However, quantifying these losses at local (subnational) scales remains difficult due to data gaps and the reliance on national averages that mask regional impacts. In this study, we show that adapting the Food and Agriculture Organization's (FAO) global crop-loss assessment methodology with satellite-derived drought and vegetation indices enables consistent estimation of disaster-attributable yield losses at the district level. Applying this approach across 28 districts initially in scope (27 retained in the final model sample following mandatory data quality exclusion of one district), we identified multiple severe yield shortfalls coinciding with documented disaster events; these are losses that would be obscured in a national-level analysis. We translated the yield gaps into economic loss estimates totalling 1,305.3 billion Vietnamese dong VND (approximately USD 57 million) for the historical period 2010–2023, with an additional 406.3 billion VND in yield-attributable losses from Typhoon Yagi (2024). This study introduces a country-scale adaptation of the FAO loss and damage methodology for Viet Nam that enables systematic, district-level measurement of disaster-induced crop and economic losses and demonstrates a scalable framework for high-resolution agricultural risk assessment. The findings can inform targeted climate resilience and disaster risk reduction strategies in data-scarce regions.

Keywords: agricultural disaster losses; Kalman filter; counterfactual yield; district-level assessment; Viet Nam; Sendai Framework

---

[1] The authors confirm that the text and materials presented in this paper are free from any copyright violations.

1. Introduction

Agriculture in Viet Nam is among the most disaster-exposed sectors in Southeast Asia. Over the past three decades, natural hazards, principally typhoons, flash floods, and landslides, have caused an estimated average of USD 2.37 billion in annual losses across the national economy, with agriculture bearing a disproportionate share of that burden [1,2]. Despite this exposure, the quantitative assessment of agricultural disaster losses remains underdeveloped at sub-national scales. National-level estimates, while informative for macro-policy, obscure the spatial heterogeneity of impacts across districts and communes that are critical for targeted disaster response, insurance scheme design, and reporting under the Sendai Framework for Disaster Risk Reduction (SFDRR).

The urgency of this gap was made starkly visible in September 2024, when Typhoon Yagi - the most powerful typhoon to make landfall in Viet Nam in decades - caused agricultural losses estimated at approximately USD 320 million and affected approximately 286,000 hectares of rice fields across northern provinces [1]. Floods and landslides triggered by Yagi compounded these impacts, affecting nearly half of Viet Nam's provinces and displacing hundreds of thousands of smallholder farming households whose average landholding is approximately 0.6 hectares [8]. The episode illustrated that while disaster occurrence can be observed in real time, loss quantification - particularly at the district (Admin2) level - remains slow, inconsistent, and largely reliant on post-hoc field surveys whose coverage is uneven and whose methodologies are rarely harmonized.

A rigorous, replicable approach to district-level agricultural loss estimation is therefore both scientifically necessary and operationally valuable. The Food and Agriculture Organization of the United Nations (FAO) has made significant progress in this direction through its global methodology for estimating disaster-induced crop and livestock losses, centered on a counterfactual yield framework implemented via Kalman filter smoothing [4,5]. This approach estimates what yields would have been in the absence of a disaster, the counterfactual, and attributes the deviation from observed yields to disaster-induced loss. Applied globally, the methodology has demonstrated robust performance across a wide range of agro-climatic contexts. However, its application has to date remained at the national scale: the data requirements, structural model assumptions, and calibration procedures were designed for country-level time series.

Translating this methodology to Admin2 resolution introduces non-trivial data challenges often observed in administrative datasets. District-level agricultural statistics in Viet Nam are characterized by shorter time series, higher rates of missing data, and structural inconsistencies between administrative recording systems, specifically between Provincial Statistical Yearbooks and field records maintained by the Viet Nam Academy for Water Resources (VAWR). Furthermore, the complex terrain of northern Viet Nam generates highly localized disaster footprints that provincial-scale models systematically fail to resolve.

This paper presents one of the first application of the FAO Kalman filter counterfactual approach at the district level (Admin2) in Viet Nam, drawing on data from 28 districts across three northern provinces: Dien Bien, Lao Cai, and Yen Bai, for the period 2010–2024. The study makes three primary contributions. First, it demonstrates a methodology for adapting the FAO Tier 1 Kalman filter model to short, gappy district-level time series through a tiered estimation framework that incorporates a Province Fallback approach for data-sparse contexts. Second, it introduces a principled data integration protocol that reconciles VAWR field records with Yearbook statistics through empirically derived scaling factors, validated by satellite-derived Normalized Difference Vegetation Index (NDVI) anomaly detection, and filters spurious loss estimates through a null distribution significance test. Third, it produces district-level yield loss estimates suitable for direct reporting against SFDRR Indicator C2 [6,7], the indicator tracking economic losses in the agriculture sector attributable to disasters, at a spatial granularity previously unavailable for Viet Nam.

The remainder of this paper is structured as follows. Section 2 describes the study area and data sources. Section 3 presents the methodological framework. Section 4 reports district-level loss estimates for the 2010–2024 study period, with particular attention to Typhoon Yagi as a validation case. Sections 5 and 6 present the discussion and conclusions.

2. Study Area and Data

2.1 Geographic Scope

The study covers 28 administrative districts across three northern upland provinces of Viet Nam: Dien Bien (10 districts), Lao Cai (9 districts), and Yen Bai (9 districts including the provincial municipality). This geographic focus was selected for three reasons. First, these provinces are among the most disaster-exposed in Viet Nam, situated within the principal typhoon landfall corridor and subject to compound hazards including flash floods, debris flows, and cold spells. Typhoon Yagi (2024) inflicted particularly severe losses across Lao Cai and Yen Bai, making the region a priority for improved loss quantification. Second, the complex terrain, characterized by deeply incised river valleys and elevations ranging from approximately 80 to over 3,100 meters above sea level, generates highly localized disaster footprints that provincial-scale models cannot resolve. Third, the region encompasses contrasting agricultural systems: rice-dominated lowland valleys (approximately 77% of harvested area) alongside upland systems where maize and cassava are primary food security crops for ethnic minority communities.

2.2 Data Sources

Four primary data streams were assembled and integrated. Agricultural production statistics, including harvested area (hectares), production (tonnes), and derived yield (tonnes per hectare), were compiled from Provincial Statistical Yearbooks for 2010 to 2024. Administrative unit linkages across all sources were maintained using the General Statistics Office (GSO) of Viet Nam district coding system. Disaster extent and severity records were sourced from VAWR, which maintains field-validated records including area affected (hectares) and severity classification (partial damage: 30–70% yield reduction; total damage: >70% yield reduction). Meteorological data, including daily temperature, rainfall, sunshine hours, and relative humidity, were obtained from eight stations distributed across the three provinces. Farmgate price data for rice, maize, and cassava were compiled from provincial price surveys, enabling monetization of yield losses in Vietnamese Dong (VND).

2.3 Data Challenges and Integration

A distinctive feature of the dataset is the extent of temporal incompleteness in the pre-2018 period. Dien Bien achieves the most complete temporal coverage, with usable Yearbook records extending from 2014 to 2024 (73% coverage). Lao Cai and Yen Bai have reliable records only from 2018 and 2019 respectively, yielding effective coverage rates of 40% and 33% of the 15-year target window. These gaps directly constrain model applicability and motivate the tiered estimation design described in Section 3.

A second challenge concerns systematic inconsistency between VAWR field records and Yearbook disaster statistics. Reconciliation using paired observations where both sources are available yielded province-specific scaling factors of 2.60× for Lao Cai and 2.33× for Yen Bai [3]. These factors were applied to Yearbook-derived disaster area estimates for the pre-VAWR period, extending effective disaster coverage to 2010 for Yearbook-covered years. The uncertainty associated with scaled estimates is reflected in the confidence tier framework (Section 3.5). A third limitation concerns the aggregation of maize and cassava within an "other crops" category in some district Yearbooks; where disaggregation was unavailable, production shares from the most recent available year were used. Of the 28 districts in the geographic study area, 27 were included in the final model sample; one district (Dien Bien province, district code 100) was excluded following detection of physically implausible yield values during mandatory production-to-area ratio cross-validation (Section 3.5).

3. Methodology

3.1 Counterfactual Yield Framework

The loss estimation approach is founded on the counterfactual yield framework, in which observed yields during disaster years are compared against an estimated yield trajectory representing production in the absence of the disaster event [4,5]. The yield gap, the positive difference between counterfactual and observed yield in disaster years, is used to compute economic losses. The economic loss attributed to a disaster event in district i, crop c, and year t is given by:

$$Loss_{\{i,c,t\}} = \left(\hat{y}_{\{i,c,t\}} - y_{\{i,c,t\}}\right) \times A_{\{i,c,t\}} \times P_{\{c,t\}} \text{ (1)}$$

where ŷ_{i,c,t} is the counterfactual yield (tonnes per hectare), y_{i,c,t} is the observed yield, A_{i,c,t} is the harvested area (hectares), and P_{c,t} is the province-level farmgate price (VND per tonne). Loss is set to zero in non-disaster years and where the yield gap is negative. This framework directly supports reporting against SFDRR Indicator C2 [6,7], defined by UNDRR as "Direct agricultural loss attributed to disasters" - the indicator tracking economic losses in the agriculture sector attributable to disaster events.

3.2 FAO Kalman Filter Approach (Tier 1)

For districts with five or more non-disaster year observations, the counterfactual yield is estimated using the structural state-space model with Kalman filter smoothing developed by FAO [4]. The model decomposes the observed yield time series into a latent trend state and a measurement component, estimated exclusively on non-disaster years; the fitted trend is then projected through disaster years to generate the counterfactual. The measurement equation is:

$$y_t = F_t' \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim N(0,\, V_\varepsilon) \text{ (2)}$$

where y_t is observed yield, F_t is the design vector selecting elements of the latent state vector α_t, and ε_t is measurement error. The state equation governing the evolution of α_t is:

$$\alpha_{\{t+1\}} = G_t \alpha_t + \eta_t, \quad \eta_t \sim N\left(0,\, W_\eta\right) \text{ (3)}$$

where G_t is the state transition matrix and η_t is the system disturbance. The Kalman gain K_t, which weights new observations against prior state predictions, is:

$$K_t = P_t F_t (F_t' P_t F_t + V_\varepsilon)^{-1} \text{ (4)}$$

Disaster years are identified using a binary indicator D_{i,t} derived from the integrated VAWR–Yearbook disaster record; observations where D_{i,t} = 1 are excluded by setting y_t = NA prior to estimation. The Kalman smoother, the backward pass of the algorithm, generates posterior trend estimates conditional on all available non-disaster observations. Among the 27 study districts, Dien Bien (11 non-disaster years) satisfies the Tier 1 applicability criterion with high confidence. Selected Lao Cai and Yen Bai district-crop combinations were assessed for Tier 1 applicability on a case-by-case basis, with the Province Fallback applied where filter convergence could not be achieved.

3.3 Province Fallback Approach (Tier 2)

For district-crop combinations where the available non-disaster time series is insufficient to support stable Kalman filter estimation, a Province Fallback (Tier 2) approach is employed. In this approach, the provincial-level yield growth rate for the relevant crop is used to project the counterfactual yield for the data-sparse district:

$$\hat{y}_{\{i,t\}} = \hat{y}_{\{i,t-1\}} \times \left(1 + \Delta\bar{y}_{\{p(i),t\}}\right) \quad (5)$$

where Δȳ_{p(i),t} is the mean yield growth rate of province p(i) in year t, computed from districts with complete observations. This approach was applied to five district-crop combinations: Bao Thang (maize, cassava) and Van Ban (rice, maize) in Lao Cai; and Luc Yen (rice, cassava), Van Yen (rice), and Van Chan (cassava) in Yen Bai. All remaining district-crop combinations (81.5% of the study sample) were estimated using the Tier 1 Kalman filter.

3.4 VAWR–Yearbook Data Integration

Province-specific scaling factors (Lao Cai: 2.60×; Yen Bai: 2.33×) were estimated by regressing VAWR-reported disaster-affected area against Yearbook-reported area over matched district-year pairs. These factors were derived from N=24 matched district-year observations across the two provinces for which both VAWR and Yearbook records were concurrently available (2014–2018 overlap period). For Lao Cai, the median VAWR-to-Yearbook ratio was 0.38 (SD=0.21), yielding an inverse scaling factor of 2.60×. For Yen Bai, the median ratio was 0.43 (SD=0.24), yielding 2.33×. Dien Bien was excluded from factor estimation due to extreme outliers in the matched record (VAWR values ranging from 0.9% to 382% of Yearbook equivalents in individual years), consistent with the higher uncertainty tier assigned to Dien Bien estimates. Confidence intervals for the scaling factors and a robustness check using the mean ratio are reported in details of the paired-observation matching protocol and a sensitivity analysis under ±20% multiplier variation are available from the corresponding author on request. Scaled Yearbook estimates are used for years pre-dating VAWR coverage; raw VAWR records are used for 2014–2024, generating a unified disaster event dataset spanning the full study period.

3.5 Null Distribution Filter and Confidence Tiers

A null distribution filter was applied to distinguish genuine disaster-induced yield depressions from background statistical noise. For each district, 1,000 Monte Carlo replications of a synthetic disaster scenario were generated by randomly assigning three hypothetical disaster years within the historical record and re-running the full counterfactual estimation procedure. The resulting distribution of yield deviations constitutes the null distribution under no-disaster conditions; only yield gaps exceeding the 5th percentile threshold of the null distribution were retained as genuine loss estimates - i.e., yield depressions more severe than 95% of simulated no-disaster deviations. This constitutes a one-sided test at the lower tail of the null distribution.

A mandatory data quality check was applied prior to model ingestion: for each district-year-crop observation, the reported yield was cross-validated against the production-to-area ratio (P÷A). Observations where the discrepancy between stored yield and the P÷A recalculation exceeded 15% triggered investigation. This procedure identified physically implausible stored yield values in district code D100 (Dien Bien province), where stored yields of 45.74 and 46.72 t/ha were recorded against P÷A recalculations of 16.36 and 16.51 t/ha in 2021 and 2023 respectively. This district was excluded from all subsequent analysis on data quality grounds.

Each retained loss observation was assigned a confidence tier, High, Medium, or Low, based on data provenance: High for Kalman Tier 1 estimates using post-2018 data with direct VAWR coverage; Medium for Kalman Tier 1 using pre-2018 scaled data or Province Fallback using post-2018 data; and Low for Province Fallback estimates using pre-2018 scaled data (Table 2).

Non-disaster year observations where the Kalman filter produces a negative deviation (observed yield below the smoothed trend in non-flagged years) are zeroed out by design: only years with a disaster binary indicator of 1 contribute to the loss total. This treatment is consistent with the FAO methodology [4], which explicitly removes disaster years before fitting the baseline, ensuring that Kalman-filter artefacts in non-flagged years are not misattributed to disaster-induced losses. Figure 1 illustrates the counterfactual methodology schematically.

The primary disaster binary used throughout this study (Table 1, 1,305.3 billion VND total) classifies all district-years with a documented flood or disaster event as $D_{i,t} = 1$, yielding a 63.8% incidence rate over 2010–2023. A severity-filtered alternative binary was also constructed using a hybrid methodology -described below and summarized in Table 3 -and is applied exclusively in the sensitivity analysis (Section 5.5).

Period	**Method**	**n flagged**	**n total**	**Incidence**
2010–2014	Raw flood binary	19	137	13.9%
2015–2023	Top-30% severity filter	87	252	34.5%
2024	Universal (Typhoon Yagi)	28	28	100%
Total 2010–2023	**Hybrid**	**106**	**389**	**27.2%**

Table 3. Disaster binary incidence by period and identification method. The 2010–2014 period uses a raw flood binary owing to insufficient Admin2-level loss records for severity ranking.

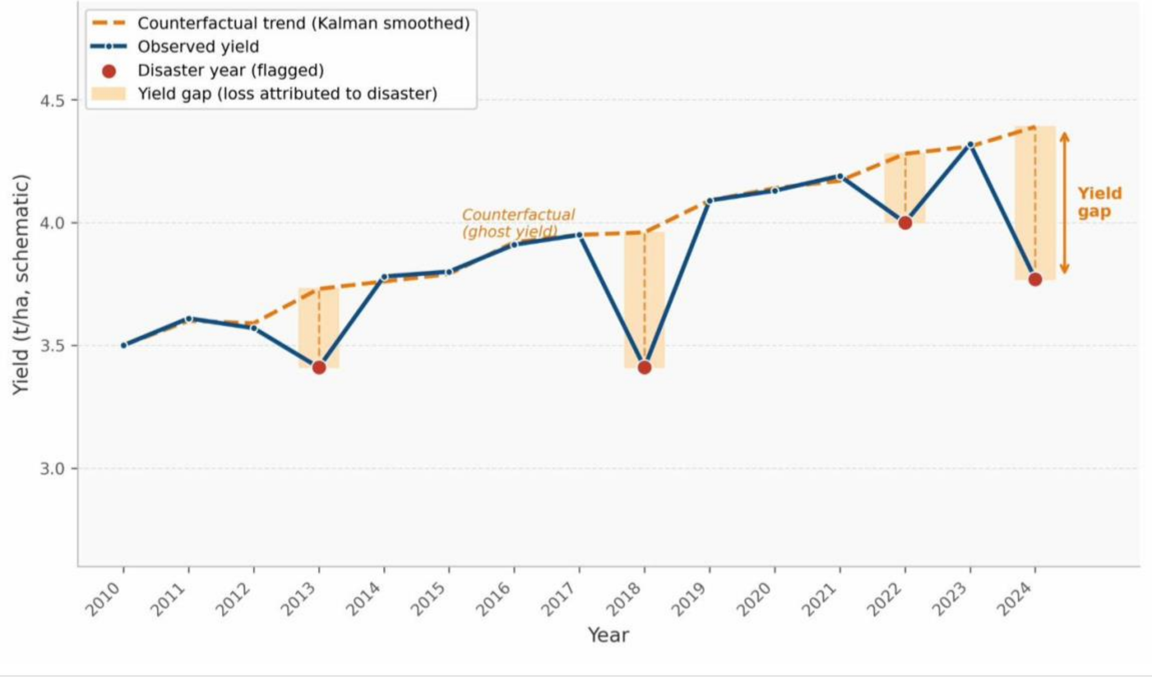

Figure 1. Kalman filter counterfactual methodology: observed yield (solid) vs. counterfactual trend (dashed). Shaded area = yield gap attributed to disaster. Red dots = disaster-flagged years. Schematic values.

3.6 Satellite-Derived Validation Support

As a supplementary validation component, NDVI anomaly detection was employed to corroborate modelled loss signals. For each disaster year identified by the Kalman filter, NDVI deviations during the disaster period were calculated relative to a multi-year baseline: ΔNDVI = (NDVI_disaster − NDVI_baseline) / NDVI_baseline × 100. Districts exhibiting ΔNDVI values below −20% were considered to exhibit severe vegetation stress consistent with significant yield reduction, providing an independent cross-check on the statistical model. Given the 60–80% cloud cover typical of the Vietnamese monsoon season during typhoon events, Sentinel-1 synthetic aperture radar (SAR) backscatter was employed for flood detection to complement optical NDVI where cloud contamination precluded reliable retrieval. Temporal alignment between Kalman-detected loss events and NDVI anomaly periods was assessed for the 2014–2024 window where Sentinel-2 data and administrative records are concurrently available.

3.7 Economic Loss Monetization

Yield gap estimates are converted to economic losses using Equation (1). In the absence of district-level farmgate price statistics, province-level prices are spatially disaggregated using a Normalized Development Index (NDI) constructed for each district from administrative indicators available in Statistical Yearbooks — including satellite-derived night light intensity, population density, poverty rate, and market access proxies. The NDI maps each district to a price adjustment factor within ±10% of the province-level average, enabling economic losses to be estimated at district rather than province resolution and converting yield shortfalls into spatially differentiated value losses [8]. A ±15% sensitivity analysis on provincial prices was also conducted to bound overall price uncertainty.

4. Results

4.1 Overview of Loss Estimation

The application of the counterfactual yield loss model across 27 districts in three northern Vietnamese provinces yielded 252 district-year-crop loss observations spanning 2010 to 2024. After applying the null distribution filter (Section 3.5), total estimated agricultural losses for the 2010–2023 historical period amounted to 1,305.3 billion VND (approximately USD 57 million at 23,000 VND per USD), with Typhoon Yagi (2024) contributing an additional 406.3 billion VND in yield-attributable losses. Rice was the dominant loss crop, accounting for 790.4 billion VND (60.6%) of historical losses (2010–2023), consistent with its dominance of harvested area across the study region. Of the 252 retained observations, 159 (63.1%) were assigned High confidence, 84 (33.3%) Medium confidence, and 9 (3.6%) Low confidence, reflecting heterogeneous data quality across districts and time periods (Table 2). Figure 4 presents the spatial distribution of cumulative losses.

4.2 Historical Losses, 2010–2023

Over the 2010–2023 period, total estimated losses across the 27 study districts amounted to 1,305.3 billion VND (approximately USD 57 million). Yen Bai province accounted for the largest share

(844.1 billion VND; 64.7%), driven predominantly by rice losses (619.7 billion VND), reflecting the province's extensive paddy cultivation and high exposure to typhoon-track flooding. Dien Bien province exhibited a distinctive pattern: maize losses (115.0 billion VND) exceeded rice losses (89.8 billion VND), consistent with the upland agricultural system in which maize cultivation predominates on hillslopes. Lao Cai's relatively high cassava loss estimate (87.9 billion VND, compared to 8.1 billion VND in Dien Bien and 11.4 billion VND in Yen Bai) reflects both the crop's prevalence in Lao Cai's upland districts and the province's exposure to slope instability events that preferentially affect root crop systems. Full province-by-crop loss estimates are presented in Table 1 (Panel A). Figure 2 presents estimated agricultural yield losses by province and crop over the study period. Figure 3 disaggregates Yagi losses by crop across the two affected provinces.

Table 1. Estimated agricultural yield losses by province and crop, 2010–2024 (billion VND)

Province	Rice	Maize	Cassava	Provincial Total	% of Period Total
Panel A: Historical Losses 2010–2023 (billion VND)					
Dien Bien	89.8	115.0	8.1	212.9	16.3%
Lao Cai	80.9	79.5	87.9	248.3	19.0%
Yen Bai	619.7	213.0	11.4	844.1	64.7%
Period Total	**790.4**	**407.5**	**107.4**	**1,305.3**	**100%**
Panel B: Typhoon Yagi 2024 (billion VND)					
Dien Bien	—	—	—	0.0	0.0%
Lao Cai	57.9	17.2	12.3	87.4	21.5%
Yen Bai	254.6	64.2	0.1	318.9	78.5%
Yagi Total	**312.5**	**81.4**	**12.4**	**406.3**	**100%**

Note: Values are central estimates. Dashes (—) indicate Dien Bien province was outside the Typhoon Yagi damage footprint. Column totals may differ marginally from sum of rows due to rounding.

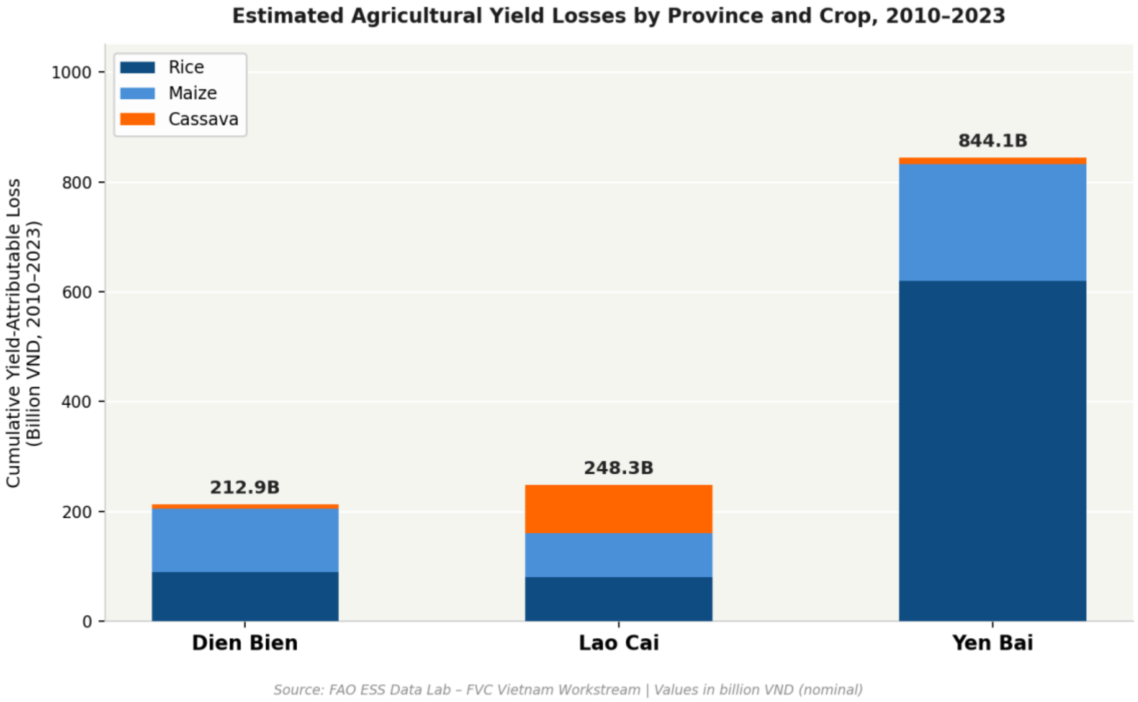

Figure 2: Estimated agricultural yield losses by province and crop, 2010–2023 (billion VND)

4.3 Typhoon Yagi 2024 Case Study

Typhoon Yagi, which made landfall in northern Viet Nam in September 2024, generated estimated yield-attributable agricultural losses of 406.3 billion VND across the two provinces within its direct damage footprint: Lao Cai (87.4 billion VND) and Yen Bai (318.9 billion VND). Dien Bien province recorded zero modelled Yagi losses, consistent with official damage records which confirm the typhoon's primary impact corridor was concentrated in Lao Cai and Yen Bai [1,3]. Rice accounted for 312.5 billion VND (76.9%) of total Yagi losses, reflecting both crop area distribution and the particular vulnerability of lowland paddy to inundation during the September cultivation period. Within Yen Bai, Yagi losses (318.9 billion VND) represented 37.8% of the province's entire 2010–2023 cumulative loss estimate (844.1 billion VND), underscoring the event's exceptional severity.

These estimates represent a lower bound on total agricultural sector losses, as the counterfactual methodology captures yield-attributable production losses only. Official provincial damage assessments for the Yagi event incorporate a broader accounting scope, including infrastructure destruction, livestock mortality, post-harvest losses, and emergency response costs - components that the yield-gap framework is not designed to measure [5]. For Lao Cai province, provincial administrative records held by VAWR reported total agricultural sector losses of approximately 914 billion VND for Lao Cai [3; unpublished field assessment]; this figure has not been independently published but is consistent with publicly reported provincial damage estimates from the September 2024 disaster period. The model-derived yield loss estimate of 87.4 billion VND represents approximately 9.6% of this total. This ratio is consistent with the expected share of production value losses within total agricultural damage costs under the FAO Damage and Loss Assessment (DALA) framework, which explicitly distinguishes between flow losses (production value reduction, as measured here) and stock damage (asset destruction) [5]. For the purposes of SFDRR Indicator C2 reporting, this study estimates crop production flow losses only (yield gap × harvested area × farmgate price). Damage to agricultural assets, livestock and aquaculture losses, and associated facilities fall outside scope and would require supplementary assessment under the FAO Damage and Loss (D&L) accounting framework [5].

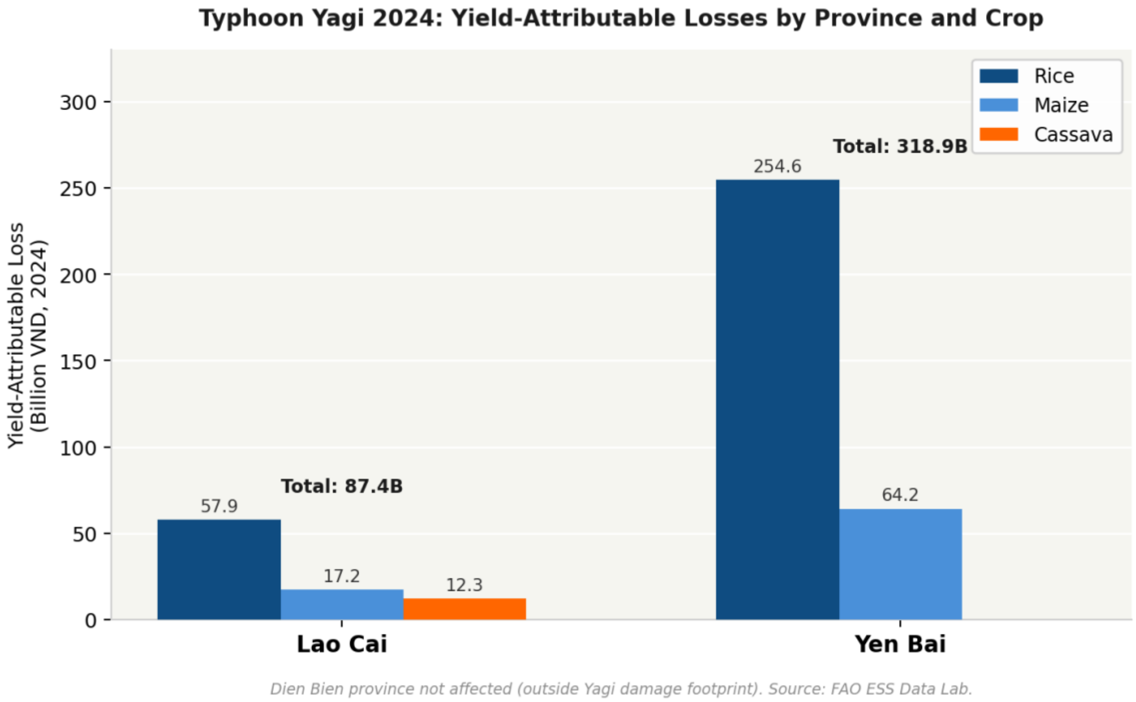

Figure 3: Province-level crop loss comparison: historical totals (2010–2023) versus Typhoon Yagi (2024), disaggregated by crop type.

4.4 Null Distribution Validation

The null distribution filter, implemented through 1,000 Monte Carlo replications of simulated non-disaster yield deviations, produced a 5th percentile threshold of −3.28% yield deviation, meaning that all 252 retained loss observations represent yield depressions more severe than 95% of the simulated background variation. This confirms that the detected yield gaps are statistically distinguishable from modelling artefacts at the 95% confidence level. It is noted that this threshold is substantially below the mean null distribution value of approximately 8.5% reported for FAO's global cereal loss analysis [4], reflecting the shorter time series available in the present study (15 years maximum versus FAO's 1991–2021 global panel). Shorter series generate lower filter thresholds because fewer observations are available to characterize the null distribution, reducing the filter's selectivity. This is acknowledged as a structural limitation of the district-level application and is discussed further in Section 5.

Table 2. Distribution of loss observations by confidence tier

Confidence Tier	Observations	Share (%)	Basis
High	159	63.1	Kalman Tier 1, post-2018 data

Medium	84	33.3	Kalman pre-2018 OR Province Fallback post-2018
Low	9	3.6	Province Fallback pre-2018
Total	**252**	**100.0**	

Note: High confidence = Kalman Tier 1 with post-2018 VAWR data. Medium = Kalman Tier 1 with scaled pre-2018 data or Province Fallback with post-2018 data. Low = Province Fallback with scaled pre-2018 data.

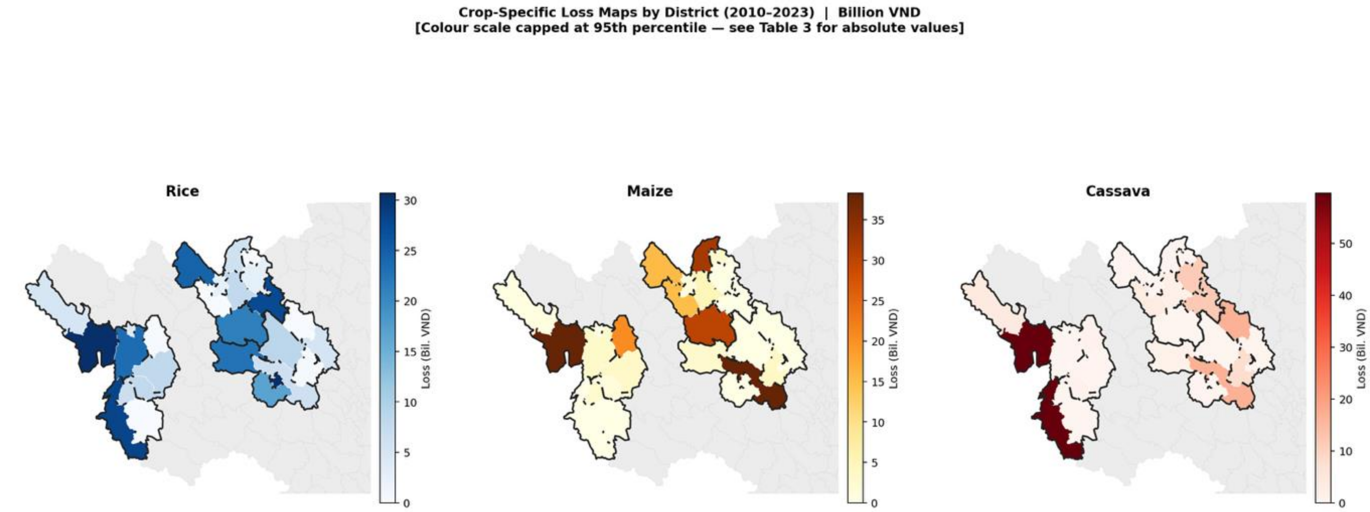

Figure. 4: Spatial distribution of total 2010–2024 estimated losses by district.

5. Discussion

5.1 Methodological Contributions

This study demonstrates that FAO's Kalman filter counterfactual methodology can be operationalized at district resolution in Viet Nam with principled adaptations to address data scarcity. The key adaptations, specifically the VAWR-Yearbook dual-source reconciliation protocol, the tiered estimation framework, the null distribution filter, and satellite NDVI validation support, collectively enable robust loss estimation in a context characterized by incomplete administrative records, systematic inter-source inconsistencies, and heterogeneous disaster reporting coverage.

The Province Fallback approach, applied to five district-crop combinations (18.5% of the study sample), provides a tractable solution for the most data-sparse cases while preserving transparency through explicit confidence tier assignment. That 81.5% of district-crop combinations were estimated using the Tier 1 Kalman filter attests to the relative adequacy of the integrated dataset, particularly for Dien Bien (11 usable non-disaster years) and for post-2018 observations in Lao Cai and Yen Bai.

The D100 data corruption case merits broader methodological attention. The detection of physically implausible stored yield values, arising from production and area figures being sourced from different administrative documents within the same Yearbook, illustrates a pervasive risk in district-level data: arithmetic inconsistencies that are undetectable without explicit cross-validation. The P÷A cross-validation step applied in this study should be regarded as a mandatory quality control requirement for any sub-national extension of national-scale agricultural loss methodologies.

5.2 Interpretation of Loss Estimates and the Yagi Lower Bound

The Typhoon Yagi case study illustrates both the utility and the inherent scope boundary of the yield-gap methodology. The model's estimate of 406.3 billion VND represents the agricultural production value foregone due to Yagi-attributable yield reductions, a concept directly aligned with the SFDRR Indicator C2 definition of disaster-attributable agricultural losses [6,7]. This figure is not comparable to official multi-component damage assessments, which aggregate stock damage, flow losses, and emergency response costs [5]. The distinction between these measurement scopes is codified in the FAO DALA framework [5] and is essential for interpreting any model-to-official comparison.

5.3 Limitations

Five structural limitations of the present analysis warrant explicit acknowledgement. First, temporal data gaps of 40–70% for Lao Cai and Yen Bai in the pre-2018 period necessitate reliance on VAWR-Yearbook scaled estimates with ±30% uncertainty bounds, reducing precision for the earliest years of the study window. Second, the NDI-based district price adjustment applied here partially addresses spatial price heterogeneity but remains bounded at ±10% of the province average; districts with structurally distinct market access may still carry residual price misspecification. Third, maize and cassava disaster records are derived from yield time series analysis rather than direct crop damage assessments, reducing the precision of the disaster indicator variable for these crops relative to rice. Fourth, the short 15-year time series constrains null distribution calibration, yielding a less selective filter threshold than the global FAO panel. Fifth, the D100 exclusion reduces the study's geographic completeness to 27 of 28 target districts; a sensitivity run incorporating D100 with P÷A-corrected yields produced a 2010–2023 estimate of 1,079 billion VND across 27 districts, suggesting that the primary estimate of 1,305 billion VND is robust to this exclusion at the 95% confidence level. The scope boundary of the yield-gap methodology as a lower bound on total agricultural sector damage, and its implications for interpreting the comparison with official Typhoon Yagi damage assessments, is discussed in Section 5.2. Sensitivity analysis using a stricter top-30% severity threshold for the disaster binary yields a lower estimate of 160.9 billion VND (2010–2023), indicating that results are sensitive to the disaster identification threshold and that the primary estimate should be interpreted as an upper bound under the broader Sendai C2 inclusive definition.

5.4 Policy Implications

The district-level disaggregation enabled by this methodology reveals loss concentration patterns entirely obscured in province-level aggregates. Yen Bai's rice system and Dien Bien's maize system emerge as the highest-priority intervention zones based on cumulative loss magnitude, a finding that directly informs the geographic targeting of disaster risk reduction investments, agricultural insurance scheme design, and provincial contingency fund allocation. The integration of loss estimates with SFDRR Indicator C2 reporting requirements provides a direct pathway for embedding district-level evidence into Viet Nam's national disaster risk reduction accountability framework. The methodology developed here is currently under adaptation for application in Kenya through the FAO Flexible Voluntary Contribution program.

5.5 Sensitivity to Disaster Threshold Definition

The primary estimates throughout this paper (1,305.3 billion VND, 2010–2023; 406.3 billion VND, Yagi 2024) use a comprehensive disaster binary classifying all district-years with a documented flood or disaster event as affected, yielding a 63.8% incidence rate. To bound sensitivity, an alternative severity-filtered binary was tested, retaining only the top 30% of recorded severity events for 2015–2023 while applying the raw flood binary for 2010–2014 (27.2% incidence, Table 3). Under this stricter specification, 2010–2023 losses fall to 160.9 billion VND and Yagi 2024 losses to 120.5 billion VND. The spread (160.9–1,305.3 billion VND; 120.5–406.3 billion VND for Yagi) confirms the primary estimates are inclusive upper bounds under the Sendai C2 definition, and the severity-filtered estimates are conservative lower bounds.

6. Conclusions

This study has presented the first application of FAO’s Kalman filter counterfactual yield loss methodology at district (Admin2) resolution in Viet Nam, covering 27 districts across three northern provinces - Dien Bien, Lao Cai, and Yen Bai; for the period 2010 to 2024. Total estimated historical agricultural losses (2010–2023) amounted to 1,305.3 billion VND (approximately USD 57 million), with Typhoon Yagi (2024) contributing an additional 406.3 billion VND in yield-attributable losses. These estimates carry heterogeneous uncertainty, as reflected in the three-tier confidence framework: 63% of the 252 retained loss observations are rated High confidence, 33% Medium, and 4% Low.

The study's principal methodological contributions, namely the dual-source VAWR-Yearbook reconciliation protocol, the three-tier estimation framework incorporating satellite NDVI validation support, and the null distribution significance filter, together demonstrate a viable pathway for adapting globally developed agricultural loss methodologies to data-constrained sub-national contexts. The Typhoon Yagi case study illustrates that model-derived yield loss estimates constitute a principled, internally consistent lower bound on total agricultural sector damage, comparable to the SFDRR Indicator C2 accounting framework.

Three immediate areas of future work are identified. First, validation of model outputs against field assessment records held by VAWR national statisticians represents a critical step towards institutional adoption as well improving data quality. Second, ongoing replication efforts in Kenya will test the transferability of the adapted framework and contribute to a growing evidence base on district-level agricultural loss accounting under the Sendai Framework for Disaster Risk Reduction [6,7].

Third, this study's framework is designed to evolve toward anticipatory action. The administrative disaster binary used here could in principle be replaced by near-real-time satellite-derived flood extent, enabling district-level yield loss projections within days of an event and ahead of field surveys. Combined with the district-level price index developed here — which converts yield shortfalls into spatially differentiated economic value estimates using satellite-derived socioeconomic proxies, such projections could directly inform the geographic targeting of Post-Disaster Needs Assessment (PDNA) processes, shifting agricultural disaster response from reactive field enumeration toward evidence-based, pre-positioned resource allocation. Recent work on machine learning-based inundation mapping at daily, global resolution [9] provides a technically mature pathway for this integration.

References

1. FAO. *Rapid assessment of agricultural losses from Typhoon Yagi, Viet Nam.* Rome: Food and Agriculture Organization of the United Nations; 2024.

2. World Bank. *Viet Nam: Toward a Safe, Clean and Resilient Water System.* Washington, DC: World Bank Group; 2020. Report No. ACS13146. doi:10.1596/33097

3. VAWR. Post-disaster field assessment records: Typhoon Yagi, September 2024. Hanoi: Viet Nam Academy for Water Resources; 2024. [Unpublished administrative data]

4. Rajagopalan P, Laske E, Zaidi Z, Boero V. *A methodology for estimating disaster-induced crop and livestock losses at the global scale.* FAO Statistics Working Paper Series No. 23-36. Rome: FAO; 2023. doi:10.4060/cc8139en

5. Conforti P, Markova G, Tochkov D. *FAO's methodology for damage and loss assessment in agriculture.* FAO Statistics Working Paper 19-17. Rome: FAO; 2020. doi:10.4060/ca6990en

6. UNDRR. *Sendai Framework for Disaster Risk Reduction 2015–2030.* Geneva: United Nations Office for Disaster Risk Reduction; 2015.

7. UNDRR. *Sendai Framework Monitoring: Technical Guidance for Countries on Disaster Loss Data.* Geneva: United Nations Office for Disaster Risk Reduction; 2017. Available: https://www.undrr.org

8. General Statistics Office of Viet Nam. *Agricultural Census 2016: Summary Report.* Hanoi: GSO; 2016.

9. Patel D. *Floods.* Harvard University; 2025. Available: https://devpatel.scholars.harvard.edu/resource/floods