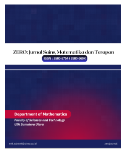
Zero : Jurnal Sains, Matematika, dan Terapan
E-ISSN : 2580-5754; P-ISSN : 2580-569X
Volume 10, Number 2, 2026
DOI: 10.30829/zero.v10i2.30038
Page: 625-637



Measuring Tail Risk in the Farmer Price Index Difference using POT-EVT

[1]Silvina Rosita Yulianti
Department of Statistics, Universitas Sebelas Maret, Surakarta, 57126, Indonesia

[2]Muhammad Bayu Nirwana
Department of Statistics, Universitas Sebelas Maret, Surakarta, 57126, Indonesia

[3]Khrisna Putri Maharani
Department of Statistics, Universitas Sebelas Maret, Surakarta, 57126, Indonesia

[4]Ahmad Zaakiy Hidayat
Department of Statistics, Universitas Sebelas Maret, Surakarta, 57126, Indonesia

[5]Andreas Ronny Wijaya
Department of Statistics, Universitas Sebelas Maret, Surakarta, 57126, Indonesia

Article Info

Article history:

Accepted: 10 July 2026

Keywords:

Extreme Value Theory;
Farmer;
Generalized Pareto Distribution;
Index Difference;
Peaks-Over-Threshold.

ABSTRACT

Farmer welfare suffers when the Farmer Price Received Index does not keep up with the Farmer Price Paid Index, and common ratio-based measures such as the Farmers' Exchange Rate do not capture sharp drops well. This study measures the extreme downside risk of the Index Difference (ID), the difference between the two indices. Because agricultural price data are heavy-tailed and not normally distributed, a Peaks-over-Threshold (POT) model based on the Generalized Pareto Distribution is fitted to the loss variable Loss = −ID, using monthly Indonesian data from January 2006 to December 2023 (216 observations). Risk is measured by Value at Risk (VaR) and Expected Shortfall (ES). At the 95th-percentile threshold, 11 exceedances give a slightly negative, near-zero shape parameter, and the 99% VaR and ES for ID are about −6.44 and −6.81. Applying POT-EVT to a farmer welfare indicator is a new contribution, and the estimates are exploratory, intended to illustrate the potential of ID-based early warning and insurance-trigger design rather than to support immediate application. With only 11 exceedances, however, these uses remain suggested rather than demonstrated, and the tail estimates should be read with caution.

Corresponding Author:

Silvina Rosita Yulianti
Department of Statistics
Universitas Sebelas Maret, Surakarta, Indonesia
Email: silvinarosita@staff.uns.ac.id

1. INTRODUCTION

Smallholder farmers in developing countries are highly vulnerable to covariate risks arising from price and yield uncertainty, which hinders productivity and threatens the long-run stability of their well-being [1], [2]. To monitor these risks, agricultural policy relies on price-based measures such as the Farmer Price Received Index

(It) and the Farmer Price Paid Index (Ib), which respectively track the prices of commodities sold by farmers and the costs of consumption and production inputs they must incur [3]. However, assessing farmer welfare through the conventional ratio of It to Ib, such as the Farmers' Exchange Rate (NTP), often yields biased or ambiguous signals about actual economic conditions [4], [5], motivating alternative measures based on the absolute difference between income and cost indices rather than their ratio [3].

The Index Difference (ID) is defined as the difference between It and Ib. A positive ID indicates that output-price dynamics exceed the dynamics of consumption and input prices, reflecting a surplus and improved purchasing power, whereas a low or negative ID signals an economic deficit in which rising input and consumption costs are not offset by output prices [2], [3]. Persistently low ID values can force farming households into extreme coping strategies, such as selling productive assets or interrupting children's schooling, and several studies indicate that such short-term shocks can develop into a persistent poverty trap [1], [2], [6].

Research on index-based agricultural insurance has focused on indices that monitor not only physical crop losses but also the economic impact of systemic shocks [7]-[15]. The effectiveness of such contracts depends on their ability to trigger payouts during farmers' most critical periods of distress, when each additional unit of income is most valuable for sustaining livelihoods. Most existing tools are built on weather, vegetation, or crop-productivity indices and therefore face basis risk, meaning the chosen indicator does not consistently align with the actual economic losses experienced by households [13], [15]. By treating extremely low or negative ID realizations as tail risk, payouts can be linked directly to the household's net price position rather than to external proxies, allowing an ID-based trigger to complement weather- or crop-based schemes and provide more complete protection against sharp declines in purchasing power [2], [14], [16].

Beyond signalling surplus or deficit, ID can be treated as a risk variable, a sharp decline marks a deterioration in the farmer's net price position. Agricultural commodity-price series typically exhibit heavy-tailed return distributions, implying a non-negligible probability of very large movements that a conventional Gaussian assumption cannot adequately describe [17], [18]. Extreme Value Theory (EVT) offers a principled way to model such rare but consequential events and to quantify the probability and magnitude of extreme losses in agricultural prices and related indices [17].

Within EVT, the Peaks-over-Threshold (POT) method models the distribution tail by fitting the Generalized Pareto Distribution (GPD) to exceedances above a chosen threshold [19]. Tail-oriented measures such as Value-at-Risk (VaR) and Expected Shortfall (ES) are commonly estimated from POT models in financial and commodity markets, and repeated backtesting shows that POT often predicts large losses more reliably than the block-maxima approach [20], [21], [22]. POT-EVT has been applied to estimate VaR for agricultural commodities such as corn, soybeans, and wheat, accounting for seasonal and weather effects and the clustering of exceedances [18].

Despite extensive EVT applications to agricultural commodity prices, weather variables, and financial returns, few studies quantify extreme downside tail risk directly at the level of a farmer welfare indicator. Existing welfare and exchange-rate measures (e.g., NTP) describe average conditions but do not characterize the magnitude or probability of rare, severe deteriorations in farmers' net price position over a long monthly series. This study addresses that gap by applying the POT-EVT framework to the Index Difference (ID), the difference between the Farmer Price Received and Farmer Price Paid indices. The novelty is therefore twofold, it extends EVT-based tail-risk measurement from commodity prices to an agricultural welfare indicator, and it provides a household-level, economically interpretable trigger that improves on conventional ratio-based welfare indicators. Accordingly, the objective of this study is to estimate the extreme downside risk of ID through GPD-based VaR and ES and to assess its robustness to threshold choice, providing an application-oriented contribution to agricultural risk monitoring and index-insurance design.

2. RESEARCH METHOD

2.1 Research Design

This research employs a quantitative methodology, concentrating on the extreme risk analysis of the ID disparity. This quantitative methodology was selected due to its capacity for objective measurement of the statistical properties of the data, especially in discerning tail behavior that standard statistical methods fail to elucidate well. This research is both empirical and computational. The empirical analysis utilized the monthly time series of the Index Difference (ID), whilst the computational component employed the Peaks-Over-Threshold (POT) method within the Extreme Value Theory (EVT) framework [21], [23], [24], [25]. This method allows for estimating the probability of extreme events by modeling values above a given threshold, as previously suggested in studies on extreme risk [23]-[28]. The EVT method, specifically the POT approach, was chosen due to its advantages in modeling tail distributions over the entire distribution based approaches. EVT is widely applied to analyze financial risks, insurance and environmental phenomena with rare but high impact events. This is relevant in the context of agricultural economics for identifying extreme fluctuations that potentially impact the stability of farmer welfare. The aim of this study is not to develop a new statistical method, but to use the EVT framework to measure the features of the extreme risks in the agricultural economic indicators. Similar approaches have been used in research on price index dynamics, climate based agricultural risks and resilience in the agricultural sector.

2.2 Data Source and Variables

The data used is secondary data in the form of the Farmer Price Received Index (It) and the Farmer Price Paid Index (Ib), obtained from official publications of the Central Statistics Agency (BPS). The observation period uses monthly time series data, covers January 2006 - December 2023 period, resulting in 216 monthly observations. Monthly aggregate data are considered appropriate for this study because the Farmer Price Received Index and Farmer Price Paid Index are officially reported on a monthly basis and are commonly used to monitor changes in farmers' economic conditions over time. Although monthly data may not capture short-term daily fluctuations, they are suitable for identifying broader economic pressure patterns and supporting policy-oriented interpretation at the aggregate level. The main research variable, ID, is calculated using the formula in Equation (1) as follows:

$$ID_t = It_t - Ib_t \quad (1)$$

Where ID_t represents the price Index Difference (ID) as an indicator of farmers' economic pressure at time t. The ID variable was chosen because it is more sensitive to capturing price imbalances than aggregate indicators such as the NTP.

2.3 Data Collection Procedure

Data was collected through documentation from the official BPS database. The collection stages included the following:

a. Identification of published agricultural price statistics
b. Extraction of monthly index data
c. Checking for consistency across time periods
d. Handling missing data (if any)

The data cover all 216 months from January 2006 to December 2023. For each month, the two published indices (It and Ib) were read directly from the BPS releases, entered into a single time-ordered table, and cross-checked against the source values to confirm that each month appears once and in the correct order. The Index Difference was then formed by subtracting the two series, ID = It − Ib, as defined in Equation (1), and the loss variable used for the EVT analysis was obtained as Loss = −ID. No seasonal adjustment, smoothing, detrending, or outlier removal was applied to the series, so the raw monthly indices enter the analysis unchanged; this keeps the extreme months, which are the focus of the study, intact rather than dampening them. The ID series is complete, with no missing values, so no imputation was needed. The study period runs through several big economic shocks, including the COVID-19 pandemic in 2020–2021, which may have shifted the level or the volatility of the series. This is less of a problem here because POT-EVT looks only at the values above a high threshold, not at the average, so it still says something useful about the tail when these shifts happen. We still keep these episodes in mind as a possible source of unevenness when reading the exceedances.

2.4 Analytical Methods

Notation is fixed once here and used unchanged throughout. ID_t = It − Ib is the Index Difference in month t; $Loss_t$ = $-ID_t$ is the loss variable on which all tail modelling is done; u is the threshold; N = 216 is the sample size and N_u the number of exceedances of u; $\hat{\xi}$ and $\hat{\beta}$ are the maximum-likelihood estimates of the GPD shape and scale; and VaR_p and ES_p denote the Value at Risk and Expected Shortfall at confidence level p, defined formally in Section 2.4 and reported on the ID scale unless stated otherwise. Every table and every threshold in the paper is interpreted with exactly these definitions.

2.4.1 Threshold Identification

The high threshold (u) was determined using the Mean Residual Life Plot and Parameter Stability Plot approaches. This approach follows standard EVT practices for threshold selection in extreme value analysis [26]-[30]. Since the main concern of this study is the extreme decline in the Index Difference (ID), the left-tail risk of ID is transformed into an upper-tail loss variable by defining $Loss_t$ = $-ID_t$. Through this transformation, a larger value of $Loss_t$ represents a greater decline in ID. For a given threshold u, observations satisfying $Loss_t$ > u are treated as exceedances, and the excess loss above the threshold is defined as shown in Equation (2).

$$Y_t = Loss_t - u \quad (2)$$

Where Y_t denotes the excess loss above the threshold u. This study tested three candidate thresholds, the 90th, 92.5th, and 95th percentiles, to weigh how extreme the tail is against how many exceedances are left. A lower threshold keeps more observations, while a higher one zooms in on the rarest events. Looking at the three side by side shows how stable the GPD parameters, VaR, and Expected Shortfall estimates are.

2.4.2 Model Peaks-Over-Threshold (POT)

Observations that satisfy $Loss_t$ > u, where $Loss_t = -ID_t$, are modeled using the Generalized Pareto Distribution (GPD) as shown in Equation (3) below.

$$G\ (y;\ \hat{\xi}\ ,\beta)\ = 1 - \left(1\ + \frac{\hat{\xi} y}{\beta}\right)^{-\frac{1}{\hat{\xi}}} \tag{3}$$

Where $\hat{\xi}$ represents the shape while β describes the scale parameter.

2.4.3 Parameter Estimation

The GPD parameters are estimated using Maximum Likelihood Estimation (MLE), which is commonly used in extreme risk analysis due to its statistical efficiency.

2.4.4 Extreme Risk Measurement

Extreme risks are evaluated through Value at Risk (VaR) and Expected Shortfall (ES) calculated based on estimated GPD parameters. For a confidence level p, the GPD based VaR for the Loss variable is estimated as shown in Equation (4) below:

$$VaR_p = u + \frac{\beta}{\hat{\xi}}[(\frac{N}{N_u}(1-p)^{-\hat{\xi}} - 1], \hat{\xi} \neq 0 \tag{4}$$

Where N is the total number of observation and N_u is the number of exceedances above the threshold u. Furthermore, the ES is calculated as shown in Equation (5) below:

$$ES_p = \frac{VaR_p + \beta + \xi u}{1-\hat{\xi}}, \hat{\xi} < 1 \tag{5}$$

2.4.5 Robustness Check

To evaluate the sensitivity of the tail risk estimates to threshold selection, robustness checks are conducted using alternative thresholds corresponding to the 90th, 92.5th, and 95th percentiles of the Loss distribution. For each threshold, the GPD parameters, number of exceedances, VaR99, and ES99 are compared. The tail risk estimates are considered robust if VaR99 and ES99 remain relatively stable across different threshold choices.

2.4.6 Treatment of Non-Stationarity and Serial Dependence

The diagnostic tests reported in the Results show that the ID series is non-stationary and strongly autocorrelated, so the i.i.d. sampling model that underlies standard POT inference does not hold exactly. Rather than treating this only as a limitation, the analysis adopts an explicit working interpretation and three concrete adjustments. The working interpretation is that all tail quantities are unconditional sample-tail measures for the 2006–2023 regime: they describe how deep and how often extreme ID declines occurred over the observed period, not the conditional risk of any particular future month. Under this interpretation the empirical exceedance frequencies and the fitted GPD remain meaningful descriptions even when the underlying level shifts. Three adjustments then align the inference with the dependence in the data: (i) the clustering of exceedances is measured with the runs estimator of the extremal index, and the cluster structure is reported instead of a declustered refit when the number of clusters is too small for one; (ii) all resampling uncertainty is taken from a moving-block bootstrap that preserves the serial dependence, and these intervals supersede the i.i.d. bootstrap intervals wherever the two differ; and (iii) conclusions rest only on quantities shown to be stable across thresholds. A fully conditional treatment, in which a time-series filter is fitted first and the POT analysis is applied to its residuals, would place the inference on standard footing but changes the economic object from the level of ID to its innovations; it is therefore identified as the natural next step rather than pursued here.

2.4.7 Software and Tools

This analysis were conducted using R software.

2.4.8 Ethical Considerations

This study used secondary data from official public sources and therefore did not involve human subjects directly. Therefore, ethical approval was not required.

3. RESULT AND ANALYSIS

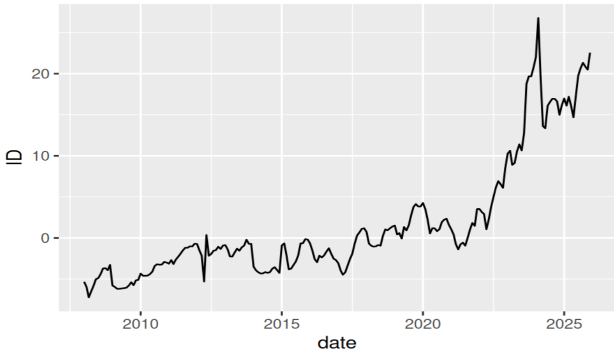


Figure 1. Time Series of the Index Difference (ID)

Based on Figure 1, the result of the descriptive analysis shows that the Index Difference (ID) series has a positive mean value of 1.651, which implies that, in general, the index received is greater than the index paid, but this condition is unstable due to a large level of variation (SD = 7.423) relative to the mean. The wide range of values from −7.24 to 26.78 confirms strong fluctuations and the possibility of extreme periods. The distribution is asymmetric, a positive skewness of 1.500 produces a long right tail that inflates the mean-based summary, while an excess kurtosis of 1.266 (greater than 0) signals tails heavier than the normal distribution. The shape of this distribution is examined next through the histogram and density curve in Figure 2.

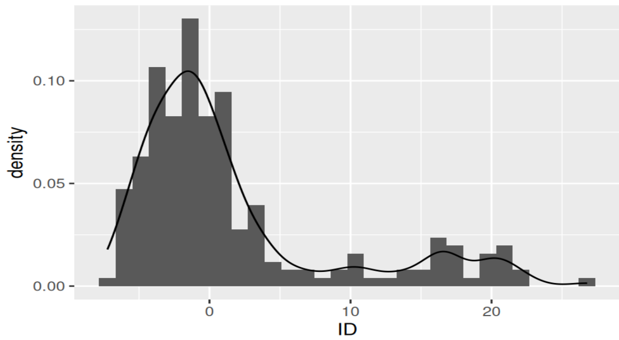


Figure 2. Distribution of the Index Difference (ID)

Figure 2 shows that the ID distribution is not normal. It is skewed to the right and has a long right tail. Tails like this matter because they make extreme values show up more often than a normal model would expect, and that is exactly when Extreme Value Theory (EVT) becomes useful. Figure 3 looks at the same distribution through a boxplot.

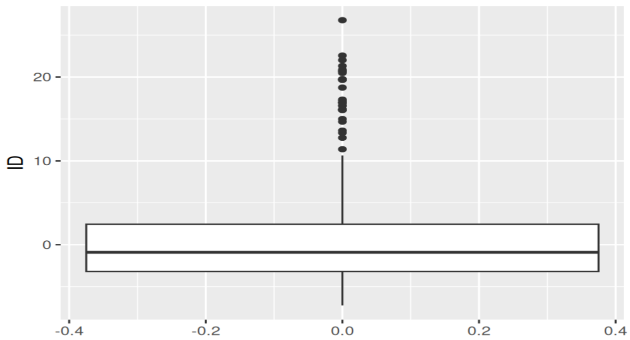


Figure 3. Boxplot of the Index Difference (ID)

Figure 3 shows the boxplot of the Index Difference (ID). The median sits a little below zero or close to it, and the box is narrow compared with the full range of the data. Most values stay in a fairly stable middle band. What stands out is a large group of outliers in the upper tail, with values well above the upper quartile (over 20). The lower tail has only a few outliers, and they sit closer to the box.

One thing worth noting here is where these outliers fall. The clear ones in the boxplot are in the upper tail of ID, which means good periods with a large surplus. The events that matter for farmer welfare are the opposite. They sit in the lower tail of ID, which is the same as the upper tail of the loss = -ID. These low-ID events are rarer and harder to see than the upper outliers. That is exactly why the EVT model works on the loss variable instead it picks out the thin left-tail risk that the plots alone cannot measure. The QQ-plot in Figure 4 looks at this more formally.

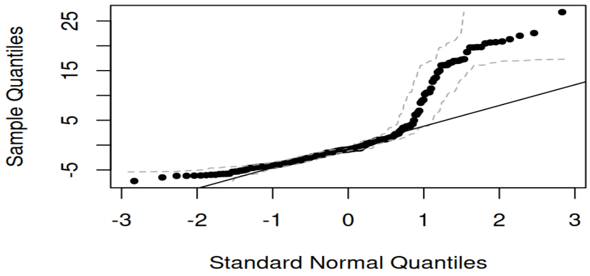


Figure 4. Normal Q-Q Plot of Index Difference (ID)

Based on Figure 4, the results of exploratory data analysis (EDA) shows that the Index Difference (ID) distribution has heavy-tail characteristics and does not meet the assumption of normality, as reflected by positive skewness, excess kurtosis greater than zero, deviations in the QQ-plot, and the presence of outliers in the boxplot. Because the focus of risk in this study is a decline in ID that reflects worsening economic conditions, the analysis is directed at the left tail of the ID distribution.

Before fitting the GPD, the stationarity and temporal-dependence properties of the series were examined, because the standard POT–GPD framework assumes that threshold exceedances behave approximately as independent realizations from a common tail distribution. The Augmented Dickey–Fuller test did not reject the unit-root null (ADF = −1.34, $p = 0.85$), whereas the KPSS test rejected level stationarity (KPSS = 3.09, $p < 0.01$); the ID series is therefore best regarded as non-stationary over the sample period. The series is also strongly autocorrelated (Ljung–Box Q(12) = 1808.6, $p < 0.001$), reflecting the high persistence typical of monthly price indices. Consequently, the independence assumption underlying the GPD asymptotics is only approximately satisfied, and the reported standard errors and confidence intervals should be read as optimistic; the tail estimates are interpreted descriptively rather than as exact probabilities. To address this dependence concretely rather than by acknowledgment alone, two dependence-robust checks complement the standard analysis. First, the temporal clustering of the exceedances was examined through the runs estimator of the extremal index. The estimate is very low ($\hat{\theta} = 0.18$), and the 11 exceedances collapse into only two clusters of consecutive months, February–May 2008 and February–August 2009, so the extreme losses represent two prolonged crisis episodes rather than eleven independent shocks. With only two cluster maxima, a declustered GPD refit is not statistically meaningful and is therefore not reported; instead, the clustering result is read as showing that the effective information in the tail comes from two historical episodes. Second, the resampling uncertainty was re-evaluated with a moving-block bootstrap (block length of 12 months), which preserves the serial dependence of the series. The resulting 95% intervals, [−7.19, −4.57] for the 99% VaR and [−7.97, −5.02] for the 99% ES, are considerably wider than their i.i.d.-bootstrap counterparts and are taken as the more honest measure of uncertainty for this series. This limitation is mitigated in part by the stability of the 99% VaR and ES across alternative thresholds and by the bootstrap confidence intervals reported in the Results. Threshold selection was guided by the Mean Residual Life (mean-excess) plot (Figure 5) and the parameter-stability plot (Figure 6), which show an approximately linear mean-excess region and reasonably stable shape and modified-scale estimates above the 95th-percentile threshold.

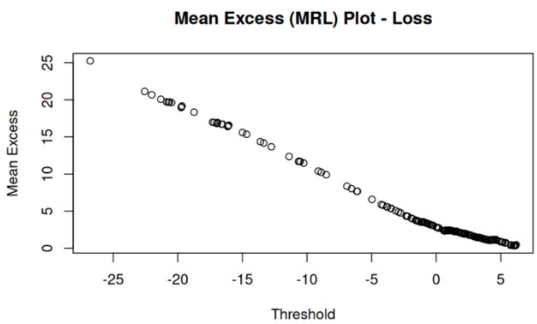


Figure 5. Mean Residual Life (mean-excess) plot of the loss series

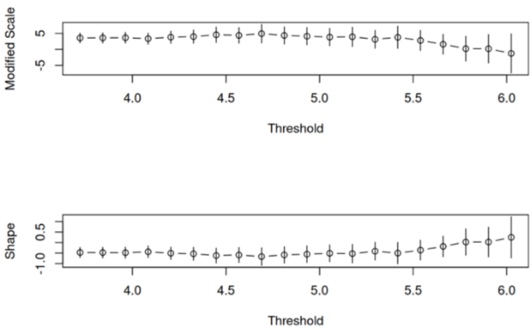


Figure 6. Parameter-stability plot of the GPD shape and modified-scale estimates across candidate thresholds

To facilitate the application of the Extreme Value Theory (EVT) approach, a Loss = $-$ID transformation is performed so that extreme events on the left side of ID are represented as the right tail of the Loss variable. Formally, for a confidence level p, the Value at Risk on the Loss scale is the p-quantile of the loss distribution, $\mathrm{VaR_p}$(Loss) = inf{x : F(x) ≥ p}, and the Expected Shortfall is the conditional mean beyond it, $\mathrm{ES_p}$(Loss) = E[Loss | Loss > $\mathrm{VaR_p}$(Loss)]. Both are positive numbers, and $\mathrm{ES_p} \geq \mathrm{VaR_p}$ by construction. When reported on the ID scale the sign is reversed: $\mathrm{VaR_p}$(ID) = $-\mathrm{VaR_p}$(Loss) and $\mathrm{ES_p}$(ID) = $-\mathrm{ES_p}$(Loss). The interpretation is then single-valued throughout the paper: ID is expected to fall below $\mathrm{VaR_p}$(ID) in a fraction (1$-$p) of months, and $\mathrm{ES_p}$(ID) is the average level of ID in exactly those months. For example, $\mathrm{VaR_p}$(ID) = -6.44 and a Loss-scale value of +6.44 describe the same event. Based on the high quantile of the loss distribution, candidate thresholds are obtained at Table 1.

Table 1. Candidate Thresholds for POT-GPD Modeling

Quantile Level	Threshold u
90%	4.60000
92.5%	5.15375
95%	5.79250

The value in the Table 1 represent the thresholds for defining exceedances in the Peak Over Threshold (POT) approach. For example, using a 95% threshold means that only about 5% of the largest loss observations equivalent to an ID < -5.79 are categorized as extreme events. Choosing a higher threshold allows for a focus on the extreme tail observations, although the number of exceedance observations becomes smaller, resulting in a trade-off between bias and variance in the GPD parameter estimates.

By choosing a threshold of $u = 5.7925$(the 95% quantile of Loss = $-$ID), the POT analysis focuses the estimation only on truly extreme events, namely observations with $loss > u$ which are equivalent to an ID <-5.7925. From a total of N=216 observations, we got $N_u = 11$ exceedances (around 5.09% data), so the empirical probability of the data being below the threshold is $p_{\leq u} = 0.9491$. In practical terms, this suggests that "normal" periods dominate the time series, but there are a small number of extreme episodes when ID falls significantly (reflecting strong economic pressures), and these episodes are relevant for measuring tail risk.

The results of modelling exceedances using the Generalized Pareto Distribution (GPD) with maximum likelihood yielded parameter estimates of shape $\hat{\xi} = -0.0403(\mathrm{SE} = 0.3134)$ and scale $\hat{\beta} = 0.4107(\mathrm{SE} = 0.1786)$. The slightly negative and near-zero $\hat{\xi}$ value indicates that the tail class is weakly identified: the point estimate is near zero, so the data are compatible with a light (exponential-type) tail as well as a moderately heavy one. Throughout the remainder of the paper this single description "weakly identified, near-zero shape" is the one intended wherever the tail is characterized. Due to the relatively large standard error and the limited number of exceedances, this result should be interpreted cautiously and should not be taken as strong evidence of a bounded tail. However, because the standard deviation $\hat{\xi}$ is relatively large compared to its value, inferentially the tail can also be considered close to the case $\xi \approx 0$ (similar to exponential) so that no strong conclusion about a tail limit is drawn. Meanwhile, $\hat{\beta}$ describes the scale of the spread of exceedances above the threshold; the larger β, the greater the variation in extreme severity when $loss$ exceeds u. Model convergence is recorded well (converged indicator = 0), so that parameter estimates can be used as a basis for calculating extreme risk measures such as VaR and expected shortfall at a high confidence level with caution. Based on the reported standard errors, the approximate 95% Wald confidence intervals are [-0.655, 0.574] for the shape parameter and [0.061, 0.761] for the scale parameter. Because the interval for the shape parameter comfortably contains zero, the data are consistent with an exponential-type (light) tail as well as with mildly heavy or mildly bounded tails; the estimate should therefore not be read as firm evidence that the tail is bounded. Indeed, with only 11 exceedances the standard error of the shape estimate (0.31) exceeds the magnitude of the estimate itself, so any statement about the tail class is indicative rather than conclusive. The Wald interval reported here is also known to be unreliable in samples this small; a profile-likelihood interval for the shape parameter, which respects the asymmetry of the likelihood and behaves better in small samples, is the more appropriate summary of this uncertainty. Computed from the same 11 excesses, the profile-likelihood 95% interval for the shape parameter is [−0.645, 1.028]. The interval is markedly asymmetric, extending much further on the heavy-tailed side than the Wald interval suggests, which reinforces that a heavy tail cannot be excluded on the present sample and that the near-zero point estimate should be read as indicative only. To further evaluate the adequacy of the fitted GPD model, diagnostic plots are presented in Figure 7.

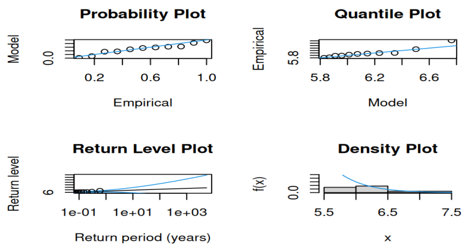


Figure 7. GPD Diagnostic Plots for POT-EVT Exceedances

The Figure 7 displays the GPD (POT–EVT) diagnostic plot, which consists of a Probability Plot, Quantile Plot, Return Level Plot, and Density Plot. Overall, this graph is used to evaluate whether the Generalized Pareto Distribution (GPD) model is adequate in modelling excesses (loss > u).

In the probability plot, the empirical points appear to follow the model's reference line relatively closely. Small deviations at the extremes are reasonable considering the number of exceedances is only 11 observations. The relatively linear pattern indicates that the GPD distribution adequately represents the probability of exceedance. In the Quantile Plot, the empirical quantiles are compared with the model quantiles. Points that are close enough to the diagonal line indicate that the parameter estimates are $\hat{\xi}$ and $\hat{\beta}$ able to capture the quantile tail structure well. No sharp extreme deviations are observed, indicating that the model does not systematically overestimate or underestimate the tail. In the Return Level Plot, the relationship between return period and return level shows a smooth and steady upward curve. This pattern is consistent with the slightly negative and near-zero estimate of $\hat{x}_i$ indicating a a near-zero, weakly identified shape. There is no explosive spike in return level, suggesting that extreme risk grows moderately as the return period increases. In the Density Plot, the GPD model's density curve follows the pattern of the exceedances histogram quite well. The distribution shape does not exhibit very thick tails or large outliers.

Overall, the four diagnostic graphs indicate that the GPD model provides an adequate fit to the exceedances data. There are no indications of serious misfit, so the estimated parameters may be used cautiously to calculate extreme risk measures such as VaR and expected shortfall at a high confidence level shown at Table 2. The adequacy of the fit was also examined through formal goodness-of-fit tests on the threshold excesses. The Kolmogorov–Smirnov test did not reject the fitted GPD (D = 0.18, p = 0.80), whereas the more tail-sensitive Anderson–Darling test indicated a marginal rejection at the 5% level (A^2 = 1.12, p = 0.03). Given that only 11 exceedances are available, these mixed results are interpreted as evidence of a reasonable, though not unambiguous, fit in the tail. To assess the stability of the risk estimates, a nonparametric bootstrap with 1,000 resamples was applied the resulting 95% confidence intervals for the 99% risk measures were approximately [−7.09, −6.13] for VaR and [−7.70, −6.19] for ES. These intervals are comparatively narrow and contain the point estimates of −6.44 and −6.81, indicating that the estimated magnitude of extreme risk is fairly stable despite the imprecision of the individual GPD parameters and the limited number of exceedances. Because the i.i.d. bootstrap ignores the serial dependence documented above, the exercise was repeated with a moving-block bootstrap (block length of 12 months), which yielded wider 95% intervals of [−7.19, −4.57] for VaR99 and [−7.97, −5.02] for ES99. The dependence-preserving intervals are the ones emphasized here: they show that the true uncertainty is larger than the i.i.d. resampling suggests, particularly toward less extreme values, while the point estimates remain well inside both sets of intervals. Because the tail evidence effectively derives from two crisis episodes, all estimates in this section are best read as exploratory rather than confirmatory, and no single-month prediction is implied.

Table 2. VaR and ES estimates for the Index Difference (ID) from the POT–GPD model (index points)

Confidence level	VaR (ID, decline)	ES (ID, decline)	VaR (Loss, +)	ES (Loss, +)
95%	−5.80	−6.19	5.80	6.19
97.5%	−6.08	−6.46	6.08	6.46
99%	−6.44	−6.81	6.44	6.81

The tail risk EVT results in the Table 2 show a measure of the extreme risk of ID decline (left tail) calculated using the POT–GPD approach at three confidence levels. Because the analysis was performed on the variable Loss = −ID, and the sign was reversed when reported, the VaR EVT and ES EVT values are negative meaning that ID is expected to decline to that value under extreme conditions.

Read with the definition above, the three rows of Table 2 have a single, uniform interpretation. At p = 95%, VaR(ID) = −5.80: ID falls below −5.80 in about 5% of months, and in those months its average level is ES(ID) = −6.19. At p = 97.5%, the corresponding values are −6.08 and −6.46, and at p = 99% they are −6.44 and −6.81: in the worst 1% of months, ID falls below −6.44, with an average level of −6.81 once it does. Moving to a higher confidence level therefore shifts both measures further into the tail, and at every level ES(ID) lies below VaR(ID) because it averages only the months beyond the VaR threshold.

Overall, the pattern of ES being constantly more negative than VaR is in line with risk measure theory, ES accounts for the average severity outside the quantiles, making it more conservative in portraying extreme impacts. Increasing the confidence level from 95% to 99% makes VaR and ES more negative, confirming the presence of tail risk in ID although the magnitude of the increase is relatively moderate, consistent with previous findings that the tail parameter $\hat{x}_i$ approaches zero (the tail is moderate rather than extremely heavy).

The results under the normal benchmark scenario show that the extreme risk measure becomes significantly more conservative (more negative) if the loss (or ID) is assumed to follow the normal distribution than if EVT is used. At 95%, VaR Normal = −10.56 and ES Normal = −13.66 are obtained, meaning the Normal model predicts that ID can fall to around −10.56 and an average fall in the worst case tail is −13.66. As the confidence level increases to 97.5% and 99%, the values become more extreme, VaR Normal = −12.90 and −15.62, and ES Normal = −15.70 and −18.13. The Normal benchmark produces a larger (more pessimistic) tail risks estimate when compared to the EVT outcomes (VaR around −5.80 to −6.44; ES approximately −6.19 to −6.81).

This difference shows that the normal assumption does not fit the tail of the trend based on the analyzed ID information, as captured using the POT–GPD based method. From a statistical perspective, a normality will "penalize" the tail in a way not grounded on actual exceedances above the threshold, thus overestimating the magnitude of the risk of extreme ID downside. On the other hand, EVT simulates the tail explicitly with the extreme data (exceedances), which leads to a more tail-oriented estimate and, in this case, a smaller tail-risk magnitude.

Moreover, a Robustness Check needs to be performed to ensure the EVT results are robust to different choices for threshold (e.g., 90%, 92.5% and 95% quantiles). If the parameters ($\hat{x}_i$), ($\hat{\beta}$) and VaR99/ES99 remain relatively consistent when the threshold is changed (and the number of exceedances remains adequate), then the conclusion about the tail risk ID can be considered robust and independent of a specific threshold choice. The robustness check results provided at Table 3.z

Table 3. The Robustness Check Across Alternative Thresholds

Quantile Level	Threshold u	Exceedances	Shape Parameter $\hat{\xi}$	VaR99 for ID	ES99 for ID
90%	4.60000	22	−0.617	−6.70	−6.95
92.5%	5.15375	17	−0.467	−6.62	−6.91
95%	5.79250	11	−0.040	−6.44	−6.81

The Table 3 above indicate that the tail risk estimate is relatively stable across threshold variations at the high quantiles (90%, 92.5%, and 95%). When the threshold is set at u = 4.60 (90%), the number of exceedances increases to 22 observations, and the shape parameter is ($\hat{\xi}$ = −0.617) with VaR99 = −6.70 and ES99 = −6.95. At the threshold u = 5.15375 (92.5%), the number of exceedances decreases to 17, with ($\hat{\xi}$ = −0.467) and estimated VaR99 = −6.62, ES99 = −6.91. Meanwhile, at the most stringent threshold u = 5.79250 (95%), there are only 11 exceedances, with ($\hat{\xi}$ = −0.040) and VaR99 = −6.44, ES99 = −6.81.

In general, although the shape parameter ($\hat{\xi}$) changes significantly when the threshold is lowered (more negative at lower thresholds), the VaR99 and ES99 estimates remain within a relatively narrow range, ranging from approximately −6.44 to −6.70 for VaR99 and −6.81 to −6.95 for ES99. This stability of the extreme risk values indicates that conclusions regarding the magnitude of tail risk are not particularly sensitive to the choice of threshold within the high quantile range. In other words, although the number of exceedances and the parameter ($\hat{\xi}$) vary, the extreme risk estimates at the 99% level remain numerically consistent.

This stability should nevertheless be read with a note of caution rather than as unqualified evidence of robustness. The shape estimate itself swings considerably across thresholds (from −0.617 at the 90th percentile to −0.040 at the 95th), so the near-constancy of VaR99 and ES99 partly reflects the fact that these quantile-based measures are anchored by the threshold and the empirical quantile rather than by the tail shape alone. What the check supports is therefore the stability of the estimated risk magnitude, not a firm identification of the tail class.

With that qualification, the tail risk estimates (ID) in the range of approximately −6.4 to −6.9 are robust, and the results of measuring extreme risk are relatively stable within the examined thresholds.

Based on the EVT estimation results at a 99% confidence level, the Index Difference (ID) has the potential to drop to approximately −6.44 under extreme conditions. This means that in approximately 1% of the worst-case periods, ID is expected to decline by at least 6.44 points. Furthermore, the Expected Shortfall ($ES_{99\%}$) measure of −6.809 indicates that when such extreme conditions actually occur (i.e., ID exceeds the VaR limit), the average decline is even deeper, at approximately −6.81. In other words, ES captures the average severity of the worst-case episodes, thus providing a more conservative risk profile than VaR.

Economically, these findings indicate that while most periods show relatively stable or positive ID conditions, there is a small but non-negligible possibility chance of a sharp decline that could reflect economic stress on farmers. A decline in ID to the range of −6 to −7 points could indicate weakened purchasing power or a significant imbalance between the received and paid indices. Therefore, tail risk measurement is crucial in the context of price stabilization policies and designing necessary schemes.

The findings can be positioned against earlier EVT applications in agriculture. Previous studies have largely applied POT-GPD or related EVT models to agricultural commodity prices and returns, or to premium calculation for crop and weather insurance [17], [18], [24], [25]. The present study differs by targeting the Index Difference (ID), a welfare-oriented indicator at the household level, rather than a single commodity price. The weakly identified, near-zero shape obtained here, with 99% VaR and ES of approximately −6.44 and −6.81, is broadly in line with the near-exponential tails reported for several agricultural price series, while extending the EVT approach to a net price-position measure that is more directly tied to farmers' purchasing power.

The near-zero shape parameter merits careful interpretation. A value close to zero places the fitted GPD near the boundary between the light-tailed (Gumbel-type) and heavy-tailed (Fréchet-type) domains, implying that extreme declines in ID decay at an approximately exponential rate rather than following a heavy power-law tail. However, with only 11 exceedances and a shape confidence interval that spans zero, this should be read as evidence that the tail is not strongly heavy, rather than as a firm conclusion that it is bounded. The robustness check, in which the shape estimate becomes more negative at lower thresholds while VaR99 and ES99 remain within a narrow band, reinforces this cautious reading, the risk magnitude is stable even though the shape parameter itself is imprecisely estimated.

For insurance design, these tail estimates suggest how an ID-based parametric trigger might be calibrated. Because the threshold and the estimated VaR and ES are expressed directly in ID units, a contract could in principle set its payout trigger near the estimated extreme quantile and use the Expected Shortfall as an indicative measure of severity once the trigger is breached. Linking payouts to the net price position in this way could reduce the basis risk that affects weather-, vegetation-, or yield-based indices, since the trigger would respond to the economic pressure actually experienced by households. These possibilities are presented as candidate applications rather than ready-to-deploy contract parameters, as trigger calibration would also require attention to moral hazard, loading, and the cost of cover.

Several limitations should be acknowledged. The analysis relies on monthly, nationally aggregated indices, which smooth short-term fluctuations and mask regional heterogeneity, so the estimated tail risk reflects aggregate rather than locality-specific exposure. The small number of exceedances at the 95th-percentile threshold (11 observations) yields wide parameter uncertainty, a known difficulty in small-sample GPD estimation [31]. In addition, the ID series was found to be non-stationary (ADF $p = 0.85$; KPSS $p < 0.01$) and strongly autocorrelated (Ljung–Box Q(12), $p < 0.001$), so the independence assumption underlying the standard POT–GPD model is only approximately satisfied and the associated confidence intervals are likely optimistic. Future work could strengthen the inference through bootstrap confidence intervals and formal goodness-of-fit tests for the GPD, the use of regional or higher-frequency data, and non-stationary or time-varying EVT models that allow the threshold or scale parameter to evolve with climatic and macroeconomic conditions [32].

4. CONCLUSION

This study used the Peaks-over-Threshold framework of Extreme Value Theory to measure the extreme downside risk of the Index Difference (ID), the gap between farmers' received and paid prices. Fitting the Generalized Pareto Distribution to the loss variable Loss = −ID at the 95th-percentile threshold gave a near-zero shape parameter and, at the 99% level, a Value at Risk and Expected Shortfall for ID of about −6.44 and −6.81. The tail class is therefore not firmly identified: the point estimate is near zero and the data are compatible with light as well as moderately heavy tails, while the 99% estimates stay stable across thresholds, which suggests that ID may fall by about six to seven points in the worst periods, a real but not catastrophic drop in farmers' net price position. These results are exploratory rather than actionable: they illustrate how an ID-based indicator could eventually inform price-stabilization policy and index-based insurance triggers, but they do not yet provide a basis for operational decisions, because they rest on just 11 exceedances that cluster into only two historical crisis episodes (2008 and 2009) and monthly, aggregated data. A further limitation is that the tail-sensitive Anderson–

Darling test gave a marginal rejection of the fitted GPD at the 5% level ($p = 0.03$), so the adequacy of the fit in the extreme tail is not beyond doubt. Future research should extend the analysis with regional or higher-frequency data, block bootstrap methods that account for serial dependence, and non-stationary EVT models for a more detailed picture.

ACKNOWLEDGEMENT

The authors would like to thank Universitas Sebelas Maret and all those who have contributed to this research. This research was supported by RKAT Universitas Sebelas Maret under the Penelitian Penguatan Kapasitas Grup Riset (PKGR- UNS) B scheme, grant number 462/UN27.22/PT.01.03/2026. The authors also acknowledge the use of ChatGPT by OpenAI to assist with language editing and manuscript clarity. All AI-assisted content was reviewed and verified by the authors, who take full responsibility for the final version of this article.

5. REFERENCES

[1] B. Kramer, P. Hazell, H. Alderman, F. Ceballos, N. Kumar, and A. G. Timu, "Is Agricultural Insurance Fulfilling Its Promise for the Developing World ? A Review of Recent Evidence," *Annual Review of Resource Economics*, vol. 14, no. 2, pp. 291–311, 2022, doi: https://doi.org/10.1146/annurev-resource-111220-014147.

[2] D. Gupta, B. Ghosh, S. Neogi, and B. Road, "Measurement ff Riskiness of Agricultural Income - Evidence from India's Agrarian Economy," *Asian Journal of Agriculture and Rural Development*, vol. 15, no. 1, pp. 60–72, 2025, doi: https://doi.org/10.55493/5005.v15i1.5321.

[3] S. R. Yulianti, A. R. Effendie, and N. Susyanto, "Improving the Accuracy of Discrepancies in Farmers ' Purchasing and Selling Index Prediction by Incorporating Weather Factors," *JTAM (Jurnal Teori dan Aplikasi Matematika)*, vol. 8, no. 3, pp. 994–1011, 2024, doi: https://doi.org/10.31764/jtam.v8i3.22584.

[4] C. M. Keumala and Z. Zainuddin, "Indikator Kesejahteraan Petani melalui Nilai Tukar Petani (NTP) dan Pembiayaan Syariah sebagai Solusi," *Economica: Jurnal Ekonomi Islam*, vol. 9, no. 1, pp. 129–149, 2018, doi: http://doi.org/10.21580/economica.2018.9.1.2108.

[5] L. Judijanto, "Beyond Terms of Trade : Challenges in Implementing the Farmers ' Welfare Index (IKP) as a Multidimensional Metric in Indonesia," *Multitech Journal of Science and Technology (MJST)*, vol. 3, no. 2, pp. 107–132, 2026, doi: https://doi.org/10.59890/mjst.v3i2.161.

[6] O. O. Adelesi *et al.*, "The Potential for Index-Based Crop Insurance to Stabilize Smallholder Farmers ' Gross Margins in Northern Ghana," *Agricultural Systems*, vol. 221, no. January, p. 104130, 2024, doi: https://doi.org/10.1016/j.agsy.2024.104130.

[7] B. Kajwang, "Weather Based Index Insurance and Its Role in Agricultural Production," *International Journal of Agriculture*, vol. 7, no. 1, pp. 13–25, 2022, doi: https://doi.org/10.47604/ija.1595.

[8] M. R. Carter, L. Cheng, and A. Sarris, "Where and How Index Insurance can Boost The Adoption of Improved Agricultural Technologies," *Journal of Development Economics*, vol. 118, no. September, pp. 59–71, 2016, doi: https://doi.org/10.1016/j.jdeveco.2015.08.008.

[9] K. Farrin and M. J. Miranda, "A Heterogeneous Agent Model of Credit-Linked Index Insurance and Farm Technology Adoption," *Journal of Development Economics*, vol. 116, no. May, pp. 199–211, 2015, doi: https://doi.org/10.1016/j.jdeveco.2015.05.001.

[10] P. Sulewski and A. Wąs, "Index-Based Insurance of Gross Margin in Agriculture – Key Challenges," *Zagadnienia Ekonomiki Rolnej Problems of Agricultural Economics*, vol. 2, no. 355, pp. 3–26, 2018, doi: https://doi.org/10.30858/zer/92059.

[11] E. Bulte and R. Haagsma, "The Welfare Effects of Index - Based Livestock Insurance : Livestock Herding on Communal Lands," *Environmental and Resource Economics*, vol. 78, no. 4, pp. 587–613, 2021, doi: https://doi.org/10.1007/s10640-021-00545-1.

[12] E. Lichtenberg and E. Iglesias, "Index Insurance and Basis Risk : A Reconsideration," *Journal of Development Economics*, vol. 158, no. November 2020, p. 102883, 2022, doi: https://doi.org/10.1016/j.jdeveco.2022.102883.

[13] B. K. Kenduiywo, M. R. Carter, and R. J. Hijmans, "Evaluating The Quality of Remote Sensing Products for Agricultural Index Insurance," *PLoS ONE*, vol. 16, no. 10, pp. 1–14, 2021, doi: https://doi.org/10.1371/journal.pone.0258215.

[14] M. Jibril *et al.*, "Index-Based Insurance and Hydroclimatic Risk Management in Agriculture : A Systematic Review of Index Selection and Yield-Index Modelling Methods," *International Journal of Disaster Risk Reduction*, vol. 67, no. October 2021, p. 102653, 2022, doi: https://doi.org/10.1016/j.ijdrr.2021.102653.

[15] E. Benami, Z. Jin, M. R. Carter, A. Ghosh, and D. B. Lobell, "Uniting Remote Sensing , Crop Modelling and Economics for Agricultural Risk Management," *Nature Reviews Earth & Environment*, vol. 5, no. December, p. 43017, 2024, doi: http://doi.org/10.1038/s43017-024-00617-y.

[16] H. Assa and M. S. Wang, "Price Index Insurances in the Agriculture Markets Price Index Insurances in the Agriculture Markets," *North American Actuarial Journal*, vol. 25, no. 2, pp. 286–311, 2020, doi: https://doi.org/10.1080/10920277.2020.1755315.

[17] M. R. C. Van Oordt, P. A. Stork, and C. G. de Vries, "On Agricultural Commodities' Extreme Price Risk," *Extremes*, vol. 24, no. Februari, pp. 531–563, 2021, doi: https://doi.org/10.1007/s10687-020-00401-3.

[18] J. Singvejsakul and C. Chaiboonsri, "The Optimization of Bayesian Extreme Value : Empirical Evidence for the Agricultural Commodities in the US," *Economies*, vol. 9, no. 1, pp. 1–10, 2021, doi: https://doi.org/10.3390/economies9010030.

[19] S. M. Bakhtiar, Amran, and Kheruddin, "An Empirical Study for Comparison of Estimation Methods for Value At Risk, Tail Value at Risk, and Adjusted Tail Value at Risk Using Extreme Value Theory Approach to Stock Market Index," *Daya Matematis: Jurnal Inovasi Pendidikan Matematika*, vol. 10, no. 2, pp. 128–137, 2022, doi: https://doi.org/10.26858/jdm.v10i2.35702.

[20] A. Aljadani, "Assessing Financial Risk with Extreme Value Theory: Us Financial Indemnity Loss Data Analysis," *Alexandria Engineering Journal*, vol. 105, no. July, pp. 496–507, 2024, doi: https://doi.org/10.1016/j.aej.2024.08.006.
[21] M. Fałdzińsk, "Peaks Over Threshold Approach with A Time-Varying Scale Parameter and Range-Based Volatility Estimator for Value-At-Risk and Expected Shortfall Estimation," *Przegląd Statystyczny. Statistical Review*, vol. 71, no. 1, pp. 23–77, 2024, doi: https://doi.org/10.59139/ps.2024.01.2.
[22] F. F. Najamuddin, E. T. Herdiani, and A. K. Jaya, "Value at Risk Estimation Using Extreme Value Theory Approach in Indonesia Stock Exchange," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 18, no. 2, pp. 695–706, 2024, doi: https://doi.org/10.30598/barekengvol18iss2pp0695-0706.
[23] G. Lazoglou, C. Anagnostopoulou, K. Tolika, and F. Kolyva-Machera, "A Review of Statistical Methods to Analyze Extreme Precipitation and Temperature Events in The Mediterranean Region," *Theoretical and Applied Climatology*, vol. 136, no. April, pp. 1–19, 2018, doi: https://doi.org/10.1007/s00704-018-2467-8.
[24] Riaman, Sukono, S. Supian, and N. Ismail, "Determining The Premium of Paddy Insurance Using The Extreme Value Theory Method and The Operational Value at Risk Approach Determining The Premium of Paddy Insurance Using The Extreme Value Theory Method and The Operational Value At Risk Approach," *Journal of Physics: Conference Series*, vol. 1722, no. 1, pp. 1–10, 2021, doi: https://doi.org/10.1088/1742-6596/1722/1/012059.
[25] Riaman, Sukono, S. Supian, and N. Ismail, "Analysing The Decision Making for Agricultural Risk Assessment : An Application of Extreme Value Theory," *Decision Science Letters*, vol. 10, no. 3, pp. 351–360, 2021, doi: https://doi.org/10.5267/j.dsl.2021.2.003.
[26] B. Bader, J. Yan, and X. Zhang, "Analysis Via Ordered Goodness-Of-Fit Tests with Adjustment for False Discovery Rate," *The Annals of Applied Statistics*, vol. 12, no. 1, pp. 310–329, 2018, doi: https://doi.org/10.1214/17-AOAS1092.
[27] S. B. Muela, C. López-martín, and M. A. Navarro, "Assessing The Importance of The Choice Threshold in Quantifying Market Risk Under The POT Method (EVT)," *Risk Management*, vol. 25, no. 1, pp. 2341–2356, 2023, doi: https://doi.org/10.1057/s41283-022-00106-w.
[28] C. Laudagé, S. Desmettre, and J. Wenzel, "Insurance : Mathematics and Economics Severity modeling of extreme insurance claims for tariffication," *Insurance: Mathematics and Economics*, vol. 88, pp. 77–92, 2019, doi: 10.1016/j.insmatheco.2019.06.002.
[29] L. F. Schneider, A. Krajina, and T. Krivobokova, "Threshold Selection in Univariate Extreme Value Analysis," *Extremes*, vol. 24, no. April, pp. 881–913, 2021, doi: https://doi.org/10.1007/s10687-021-00405-7.
[30] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*. Springer London, 2001. doi: https://doi.org/10.1007/978-1-4471-3675-0.
[31] W. A. Pels, A. O. Adebanji, S. Twumasi-ankrah, and R. Minkah, "Shrinkage Methods for Estimating the Shape Parameter of the Generalized Pareto Distribution," *Journal of Applied Mathematics*, vol. 1, no. November, pp. 1–11, 2023, doi: https://doi.org/10.1155/2023/9750638.
[32] O. Lee, I. Sim, and S. Kim, "Application of The Non-Stationary Peak-Over-Threshold Methods for Deriving Rainfall Extremes from Temperature Projections," *Journal of Hydrology*, vol. 585, no. November, p. 124318, 2019, doi: https://doi.org/10.1016/j.jhydrol.2019.124318.