Agricultural Systems 180 (2020) 102693

Contents lists available at ScienceDirect

Agricultural Systems

journal homepage: www.elsevier.com/locate/agsy

Forecasting yields, prices and net returns for main cereal crops in Tanzania as probability distributions: A multivariate empirical (MVE) approach

Ibrahim L. Kadigi[a,b,*], James W. Richardson[c], Khamaldin D. Mutabazi[a], Damas Philip[a], Jean-Claude Bizimana[c], Sixbert K. Mourice[d], Betty Waized[a]

[a] School of Agricultural Economics and Business Studies, Sokoine University of Agriculture, P.O. Box 3007, Morogoro, Tanzania
[b] Soil-Water Management Research Group, Sokoine University of Agriculture, P.O. Box 3003, Morogoro, Tanzania
[c] Department of Agricultural Economics, Texas A&M University, College Station, TX 77845, United States
[d] Department of Crop Science and Horticulture, Sokoine University of Agriculture, P.O. Box 3005, Morogoro, Tanzania

ARTICLE INFO

Keywords:
Cereal crops
MVE probability distribution
Stochastic simulation
Semi-arid area
Sub-humid area
Simetar

ABSTRACT

Maize (*Zea mays L.*), sorghum (*Sorghum bicolor* L. Moench) and rice (*Oryza sativa*) are essential staple crops to the livelihoods of many Tanzanians. But the future productivity of these crops is highly uncertain due to many factors including overdependence on rain-fed, poor agricultural practices and climate change and variability. Despite the multiple risks and constraints, it is vital to highlight the pathways of cereal production in the country. Understanding the pathways of cereals helps to inform policymakers, so they can make better decisions to improve the viability of the sector and its potential to increase food production and income for the majority population. In this study, we employ a Monte Carlo simulation approach to develop a multivariate empirical (MVE) distribution model to simulate stochastic variables for main cereal crops in Tanzania. Eleven years (2008–2018) of yields and prices data for maize, sorghum and rice were used in the model to simulate and forecast yields and prices in Dodoma and Morogoro regions of Tanzania for a seven-year period, from 2019 to 2025. Dodoma and Morogoro regions represent semi-arid and sub-humid agro-ecological zones, respectively. The simulated yields and prices were used with total costs and total area harvested for each crop to calculate the probable net present value (NPV) for each agro-ecological zone. The results on crop yield show a slightly increasing trend for all three crops in Dodoma region. Likewise, rice yield is expected to marginally increase in Morogoro with a decreasing trend for maize and sorghum, meanwhile, the prices for the three crops all are projected to increase for the two regions. Generally, the results on economic feasibility in terms of NPV revealed a high probability of success for all the crops in Dodoma despite a higher relative risk for rice. The results in Morogoro presented a high probability of success for rice and sorghum with maize indicating the highest relative risk, and a 2.41% probability of negative NPV. This study helps to better understand the outlook of the main cereal crop sub-sectors in two agro-ecological zones of Tanzania over the next seven years. With high dependence on rain-fed agriculture, production of main cereals in Tanzania are likely to face a high degree of risk and uncertainty threatening livelihoods, incomes and food availability to the poor households.

1. Introduction

Maize (*Zea mays L.*), sorghum (*Sorghum bicolor* L. Moench) and rice (*Oryza sativa*) are major staple food crops in Sub Saharan Africa (SSA) consumed by people with varying food preferences and socio-economic backgrounds (Waithaka et al., 2013). The three staple crops are grown in diverse agro-ecological zones and farming systems and account for the largest share of calories and protein consumed in SSA (Macauley and Ramadjita, 2015). However, recent productivity trends and current performance of food crops in SSA are progressively less able to meet the needs of its rapidly increasing population. The low productivity of these crops in SSA is attributed to many constraints including high dependence on rain-fed agriculture, drought, floods, pest and diseases, and inadequate application of improved seed and fertilizers leading to food insecurity in rural areas (Ziervogel et al., 2006; Cooper et al., 2008; Ziervogel and Ericksen, 2010; URT, 2013; Kahimba et al., 2015; Wilson and Lewis, 2015).

As the population of SSA is likely to grow to around 1.7 billion by 2050, demand for food to feed the population also increases (Waithaka et al., 2013). In the African Union (AU), recommitments have been

* Corresponding author at: School of Agricultural Economics and Business Studies, Sokoine University of Agriculture, P.O. Box 3007, Morogoro, Tanzania.
E-mail address: ibrahim.kadigi@sua.ac.tz (I.L. Kadigi).

https://doi.org/10.1016/j.agsy.2019.102693
Received 18 October 2018; Received in revised form 2 August 2019; Accepted 11 September 2019
Available online 20 December 2019
0308-521X/

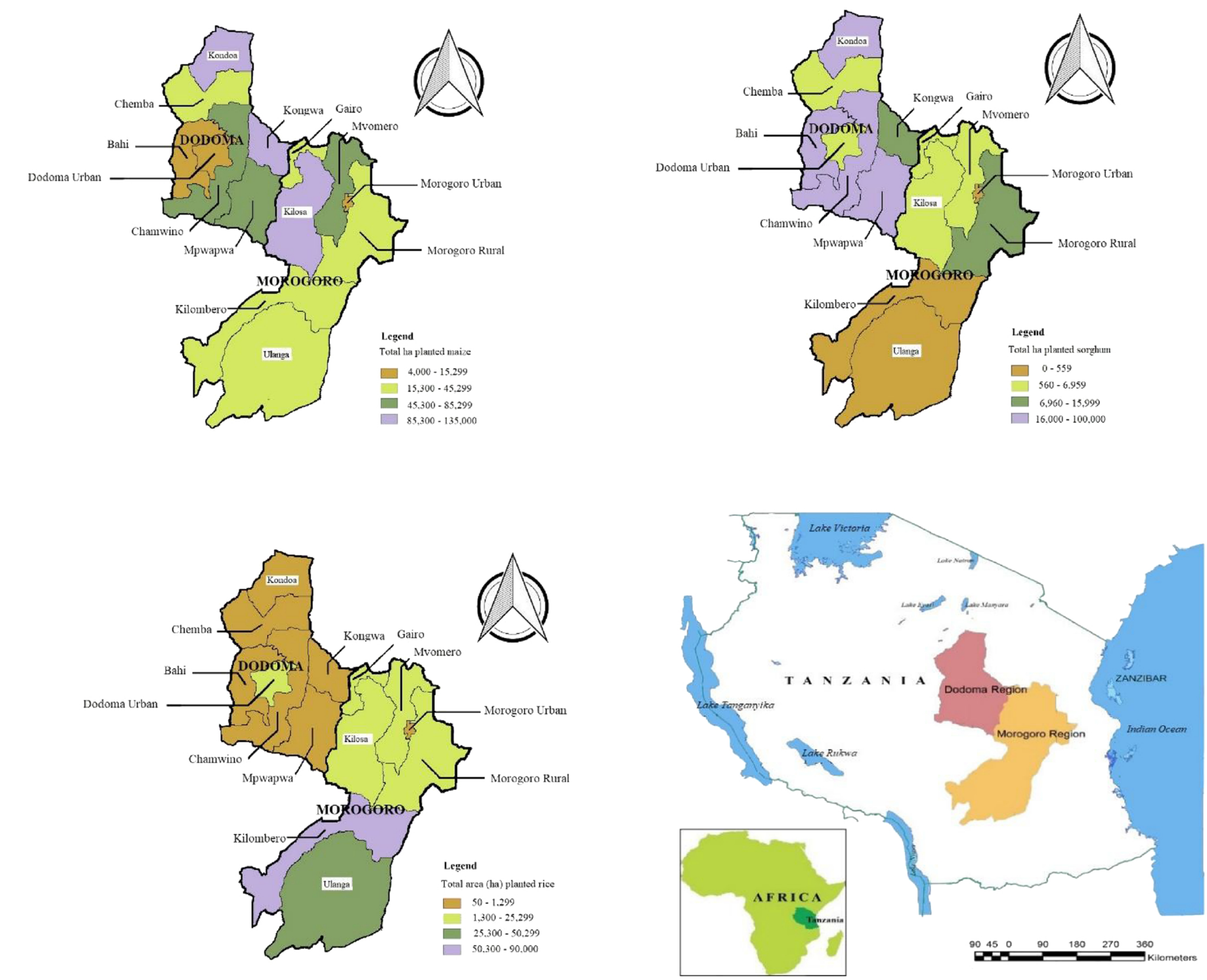

Fig. 1. The study area.

made to transform agricultural productivity in Africa by focusing on vulnerable social groups. One of the examples is the Comprehensive Africa Agriculture Development Programme (CAADP) under the 2014 Malabo Declaration on Accelerated Agricultural Growth and Transformation for Shared Prosperity and Improved Livelihoods (AU, 2014). Addressing the low performance of agriculture in SSA has, therefore, become a focal point towards attaining an agriculture revolution. For instance, under the third commitment of the Malabo Declaration, African leaders agreed to end hunger in Africa by 2025 by at least doubling current agricultural productivity levels (AU, 2014). The same commitment is embraced in the Tanzania National Agriculture Policy, which is an instrument for facilitating the attainment of Tanzania Development Vision 2025 (TDV-2025). Its objective among others is to have modernized agriculture for increased agricultural productivity and profitability by 2025 (URT, 2013a). Additionally, the rationale of the current Agricultural Sector Development Strategy Phase II (ASDP-II) (2015/2016–2014/2025) is to operationalize the transformation of the agricultural sector from low productivity into a semi-industrialized, modernized, highly productive, commercial and more resilient (URT, 2016a).

Nevertheless, some uncertainties remain about the future productivity and profitability of staple food crops in Tanzania. These uncertainties hinder the implementation of different strategies, agricultural policies and plans for achieving an agriculture revolution in the country, hence impacting decisions of investment in agrarian technologies (Ingram et al., 2008; Thornton et al., 2011; Yao et al., 2011; Msongaleli et al., 2017). Lack of timely and accurate information on future yields and prices trends of important crops affect the food security, import and export plans, crop insurance policy and government aid to farmers at national, regional, village and household levels (Kahimba et al., 2015; 2014, Kantanantha, 2007). In general, there is limited information, in terms of understanding future yields and prices of major food crops in Tanzania. This information is essential to ensuring food availability and household income predominantly for the most impoverished population. Therefore, it emphasizes the need for location specific research to estimate the feasibility of staple crops in Tanzania while considering the stochastic nature (uncertainty) of the agricultural sector.

The present study applies a multivariate empirical (MVE) probability distribution model to forecast yields and prices of maize, sorghum, and rice in Dodoma and Morogoro regions of Tanzania. The forecasting analysis performed for seven years through 2025. Dodoma and Morogoro represent the semi-arid and sub-humid agro-ecological zones, respectively. The forecasted values were combined with production costs, harvested area (in ha) and quantities of inputs used in production to estimate the probable net present value (NPV) of each crop for seven years. The MVE distribution model has been used in many studies including Rezende and Richardson (2015), Richardson et al. (2000, 2007), Hardaker et al. (2004), and Richardson et al. (2006, 2008). A seven-year horizon (2019–2025) in this study is in line with the Tanzanian National Agricultural Policy, TDV-2025, ASDP-II, and CAADP. These policy documents have a common goal of modernizing agriculture in Tanzania to a highly productive and profitable sector by 2025. Therefore, this study provides a roadmap of the cereal production from 2019 to 2025, and the findings may help the government especially the National Food Security Division (NFSD) and regional officials to better plan for future production, storage and marketing.

The scope of this study is limited to provide a computational framework for the MVE probability distribution approach by quantifying the possible variability on yield and price of staple crops and the probable economic implications at the agro-ecological zone or regional level. Nonetheless, given the ability of an MVE distribution to simulate multiple correlated random variables concurrently, this study demonstrates a base for similar studies to be conducted including those on food and cash crops possibly covering the entire country.

2. Methods

2.1. Description of the study area

The study area encompassed two agro-ecological zones: (i) the semi-arid represented by the Dodoma region; and (ii) sub-humid represented by the Morogoro region (Fig. 1). The Dodoma region is in the central part of the Tanzania mainland and lies between the GPS coordinates of 6° 9′ 40.2624″ S and 35° 44′ 43.5336″ E. Much of the region is plateau rising gradually from some 830 m in Bahi swamps to 2000 m above sea level in the highlands North of Kondoa (URT, 2012a). The region is characterized by a semi-arid climate, receiving < 800 mm of rainfall with a mean number of rain days between 10 and 20 per annum. It is one of the regions dominated by a long dry season lasting between late April to early December and short single wet season during the remaining months (Schechambo et al., 1999; Kempf, 2007; Yanda et al., 2015). The temperature in the region varies according to altitude, but generally the average maximum and minimum for October to December are 31 °C and 18 °C, respectively. The average rainfall for Dodoma region is low (570 mm on average) and unpredictable in frequency, and the amount. To some extent, the rain is higher in Mpwapwa and Kondoa districts (agriculturally productive parts of the region). Based on 2012 population and housing census, there were about 2.1 million inhabitants in the Dodoma region with the average annual increase of 2.1% (URT, 2013b). Major food crops grown in the Dodoma region are sorghum, maize, paddy (rice), beans, bulrush millet, groundnuts, and finger millet with sunflower and sesame being the main cash crops. Dodoma is one of the regions with numerous livestock including cattle, goats, sheep, poultry, and pigs. Maize is highly grown in Kondoa, and Kongwa districts followed by Chamwino and Mpwapwa districts, while rice is highly produced in Bahi district. Sorghum is dominantly grown in Chamwino, Kondoa, Bahi and Mpwapwa districts (Fig. 1).

The Morogoro region is in the Mid-Eastern part of Tanzania Mainland. It lies between the GPS coordinates of 6° 49′ 49.3428" S and 37° 40′ 14.1204″ E. Morogoro is one of the largest regions in Tanzania having a total area of about 73,000 km^2 much of which are potential areas for agriculture. The region is characterized by a sub-humid climate, with an average temperature of 24 °C, having the minimum of 18 °C in the highland areas and a maximum of 30 °C in the lowland areas (URT, 2012b). The average rainfall for Morogoro region is between 500 mm in lowland areas and 2200 mm in the mountainous zones. The region had about 2.2 million inhabitants, with an average annual increase of 2.5% (URT, 2013b). Major food crops grown in the Morogoro region are maize, rice, sorghum, bulrush millets, and beans, whereas the main cash crops include sugarcane, rice, cotton, sisal, and tobacco. Livestock keeping is also an essential activity in the region. Maize is highly grown in Gairo, Kilosa and Mvomero districts while Ulanga and Morogoro urban districts are the least maize producer in the region. Morogoro rural produces the highest sorghum in the region followed by Kilosa, Gairo and Mvomero districts. Kilombero district is the rice leading producer followed by Ulanga and Kilosa districts (Fig. 1).

Morogoro has approximately 380,000–405,000 ha; with 8000–12,000; and 160,000–222,000 ha of area planted respectively with maize, sorghum and rice, per year, while the areas in Dodoma follows between 370,000–440,000 ha for maize; 160,000–176,000 ha for sorghum; and between 8000 and 14,000 ha for rice (URT, 2012a; 2012b; 2016b; 2017). Morogoro has been named one of the national food basket regions together with Mbeya and Iringa regions in the southern highland leading in rice production by > 12% of the total rice produced in the country and the 8th region in maize production (Cochrane and Souza, 2015). On the other hand, Dodoma has been leading in the production of sorghum having the largest planted area with sorghum by over 31% and producing about 24% of the total sorghum produced in the country but still one of the food deficit regions (Cochrane and Souza, 2015; URT, 2017). Food prices are slightly higher in Dodoma and other deficit regions than the Morogoro region.

2.2. Data

This study uses data from many different sources including a series of focus group discussions (FGDs) with representatives from the National Bureau of Statistics (NBS), Ministry of Industry and Trade (MIT) and Ministry of Agriculture (MoA). Other sources are household surveys under Trans-SEC [http://trans-sec.org/], and Scale-n projects [http://www.scale-n.org] conducted in both Dodoma and Morogoro regions supplemented with grey literature from government agricultural documents. Trends on yield and total area harvested (ha) for each crop were obtained from the National Sample Censuses of Agriculture (NSCA) supplemented by the Annual Agricultural Sample Surveys (AASS). These surveys are carried out by the NBS in collaboration with MoA, MIT, the President's Office, Regional Administration and Local Government (PO-RALG) and the Office of the Chief Government Statistician, (OCGS). The surveys provide data on agricultural production at regional levels and are freely available for public access. The missing data were acquired and compiled from regional agricultural statistics in Dodoma and Morogoro regions where crop yields per district are collected annually and kept in paper-based files. Similarly, the cost of production per unit area for each crop was obtained from different sources. This includes, FGDs with farmers and experts in maize, sorghum and rice supply chains, and supplemented by household surveys conducted under Scale-n and Trans-SEC projects and through reviewing regional agricultural reports (URT, 2012a, 2012b). The costs of production (TZS/ha) are composed of: land preparation, seeds, planting, weeding, fertilizer and pesticide application, harvesting and postharvest handling. Appendix A summarizes the production costs used in this study.

Annual price data for cereals were obtained from the regional agricultural marketing departments, where daily prices of all crops are collected and archived or published. Table 1 presents the historical mean yields (in t/ha) and prices (in TZS/t) for all crops per region. For convenience, we abbreviated our variables for Dodoma region as follows MzY_1 = maize yield, SoY_1 = sorghum yield, RcY_1 = rice yield; MzP_1 = maize price, SoP_1 = sorghum price, and RcP_1 = rice price and variables for Morogoro were subscripted by a number two.

Table 2 provides additional data used in the model. The data include: (i) the approximated area growing maize, sorghum, and rice for each region compiled from NSCA, AASS, NBS, FGDs and regional agricultural offices (RAO); (ii) average production cost for each crop enterprise collected from FGDs, scale-n project, NSCA and RAO; and (iii) inflation rate and the discount rate were obtained from NBS website, Bank of Tanzania (BOT) and the trading economics website [www.tradingeconomics.com]. The website provides a collection of economic indicators, including actual values, historical data charts, time-series, and long-term forecasts. It was assumed that the same areas under all

Table 1
Historical mean yields and prices per crop per region.

Year	Yields for Dodoma			Prices Dodoma (× 100,000)[a]			Yields for Morogoro			Prices Morogoro (× 100,000)		
	MzY_1	SoY_1	RcY_1	MzP_1	SoP_1	RcP_1	MzY_2	SoY_2	RcY_2	MzP_2	SoP_2	RcP_2
	t/ha	t/ha	t/ha	TZS	TZS	TZS	t/ha	t/ha	t/ha	TZS	TZS	TZS
2008	0.60	0.70	0.80	3.487	2.804	10.210	1.30	1.30	1.30	3.551	3.929	9.220
2009	0.30	0.60	0.80	4.111	3.633	11.685	0.80	0.90	1.00	4.173	4.918	10.932
2010	0.63	1.20	1.23	3.904	3.634	11.950	1.33	1.14	1.75	3.748	6.072	10.201
2011	0.90	1.20	1.10	4.196	4.177	14.103	1.60	1.10	1.40	4.094	5.997	11.376
2012	0.50	1.00	0.60	5.748	4.956	19.263	0.80	1.00	1.30	5.202	6.992	16.253
2013	0.60	1.10	0.70	6.969	7.088	16.956	0.80	1.00	1.20	6.033	9.054	13.305
2014	1.00	1.10	0.80	5.125	5.049	13.563	0.99	0.90	2.14	4.243	8.253	11.726
2015	1.01	0.94	0.83	5.560	5.450	17.252	0.99	0.77	2.02	5.338	8.153	16.074
2016	0.98	1.05	1.03	6.809	7.264	16.771	0.93	0.91	2.05	6.944	11.954	16.962
2017	1.07	1.10	1.28	8.042	9.235	18.631	1.09	1.18	2.00	7.590	12.032	16.942
2018	0.99	1.07	1.21	4.959	5.140	17.792	1.04	1.21	2.06	4.357	10.289	17.218

[a] Exchange rate: US$ 1.00/TZS2,340; Source of Data: NBS, MoA, MIT.

the three crop enterprises are the same for the next seven years. The rate of 3.0–3.9 inflated production costs for the future, and the rate of 9.0% discounted the net returns for each crop.

2.3. MVE simulation procedures

The multivariate empirical (MVE) distribution which runs under the Monte Carlo simulation protocols was used in this study to account for the correlation among the stochastic variables Richardson et al. (2000). The MVE model was applied because the yields and prices of maize, sorghum, and rice are correlated and non-normally distributed. Additionally, production of cereal crops in Tanzania, as well as the rest of SSA, is affected by weather leading to a high variability of yields and prices. Given this inconsistency, imposing the MVE distribution to capture the heteroskedastic variability is important (Richardson et al., 2008). Eleven years (2008–2018) of historical yields for maize, sorghum, and rice were used alongside 11 years of local prices to develop the MVE distribution for the three crop sub-sectors at the agro-ecological level. The model was simulated for a period of seven years up to 2025 using stochastic yields and prices from the historical trend to forecast the distribution of the probable yields and prices. The simulated variables were combined with the total area harvested to simulate total revenue for each crop. Next, the inflated production cost was deducted to calculate the stochastic annual net cash income and the net present value (NPV) per crop per agro-ecological zone.

MVE simulation model is a step-wise process that considers changing traditional trend forecasts to stochastic simulations of random variables. The model was programmed in Microsoft Excel using the Simetar© add-in, following a procedure by Richardson et al. (2000). Since we have historical data, de-trending of the random variables (yields and prices) was the first step to estimate the deterministic component of yields and prices. De-trending of historical data helps to was remove possible systematic risk inherent in the random variables. The next steps involved the calculation of the stochastic parts and finally combining the deterministic and the stochastic elements to simulate random values for stochastic modeling.

The steps for simulation procedures are summarized as follows:

Step 1: estimation of deterministic components

The yields and prices are de-trended as the yield and price data are from historical data. Alternative functional forms (linear, quadratic, and cubic) were tested to remove systemic risk and the polynomial function of degree three (a cubic regression) was selected based on the R-Square. The deterministic component of the probability distribution from the trend regression for two equations is expressed without error term ($\hat{e}_{t,i}$) as shown in Eqs. (1) and (2).

$$\widehat{Y}_{t,\mathrm{ij}} = \hat{a} + \hat{b}\, T_{t,\mathrm{ij}} \tag{1}$$

$$\widehat{P}_{t,\mathrm{ij}} = \hat{a} + \hat{b}\, T_{t,\mathrm{ij}} \tag{2}$$

Table 2
Additional data used in the model.

Variable	Units	Assumptions	
		Dodoma[a]	Morogoro
Area under maize	ha	370,000–438,000	380,000–405,000
Area under sorghum	ha	160,000–176,000	8500–12,000
Area under rice	ha	8000–14,000	160,000–222,000
Average production cost for maize	TZS/ha	473,000	561,000
Minimum production cost for maize	TZS/ha	355,000	430,000
Maximum production cost for maize	TZS/ha	597,000	676,000
Average production cost for sorghum	TZS/ha	378,000	375,000
Minimum production cost for sorghum	TZS/ha	328,000	324,000
Maximum production cost for sorghum	TZS/ha	439,000	428,000
Average production cost for rice	TZS/ha	1567,000	1725,000
Minimum production cost for rice	TZS/ha	1212,000	1310,000
Maximum production cost for rice	TZS/ha	2,018,000	2,124,000
Inflation rate for production cost	%	3.0–3.9	3.0–3.9
Discount rate for NPV	%	9.0	9.0

[a] Source of data include FGDs, (URT, 2012a, 2012b, 2016b, 2017); [www.tradingeconomics.com], (details of costs are in Appendix A).

where: $\hat{\alpha}$ = intercept;

$\hat{b}$ = slope;
T = time (Year);
i = crops (maize, sorghum, and rice);
j = regions (Dodoma and Morogoro);
$\hat{Y}$ = average yield for crop i in the year t;
$\hat{P}$ = price for crop i in the year t;

Step 2: estimation of stochastic components

The unexplained variability about the deterministic component or $\hat{e}_{t,ij}$ (in Eqs. (1) and (2)) is the stochastic component for each variable (i) for each year (t). The residuals from the cubic regression forecasts constitute the $\hat{e}_{t,ij}$ and are divided by their respective trend forecasted values for each year to calculate the fractional deviates denoted by $F\hat{e}_{t,ij}$ and sorting of the fractional residuals denoted by $S_{t,ij}$ expressed as follows:

For yield:

$$\hat{Y}_{t,ij} = \hat{\alpha} + \hat{b}\,T_{t,ij} + \hat{e}_{t,ij} \tag{3}$$

$$\hat{e}_{t,ijY} = Y_{t,ij} - \hat{Y}_{t,ij} \tag{4}$$

$$F\hat{e}_{t,ijY} = \hat{e}_{t,ijY}/\hat{Y}_{t,ij} \tag{5}$$

$$S_{t,ijY} = \text{Sorted}\,(F\hat{e}_{t,ijY}) \tag{6}$$

For price:

$$\hat{P}_{t,ij} = \hat{\alpha} + \hat{b}\,T_{t,ij} + \hat{e}_{t,ij} \tag{7}$$

$$\hat{e}_{t,ijP} = P_{t,ij} - \hat{P}_{t,ij} \tag{8}$$

$$F\hat{e}_{t,ijP} = \hat{e}_{t,ijP}/\hat{P}_{t,ij} \tag{9}$$

$$S_{t,ijP} = \text{Sorted}\,(F\hat{e}_{t,ijP}) \tag{10}$$

where: Y and P represent the deterministic component of the Eqs. (3) and (7).

Step 3: setting of the Pseudo minimum ($\text{Pmin}_{\hat{e}}$*) and maximum (*$\text{Pmax}_{\hat{e}}$*)*

The (Pminê) and (Pmaxê) provide the end points for the distribution and calculated by multiplying the minimum and maximum residuals by 1.0001.

Step 4: estimation of the correlated uniform standard deviates (CUSD's)

Estimating the CUSD's is a crucial step in the stochastic simulation as it appropriately correlates the random variables to retain the observed stochastic dependency between variables. The Simetar add-in for Excel generates a correlated uniform standard deviate (CUSD) by calculating the square root of the correlation matrix and multiplying it by a vector of independent standard normal deviates. It then converts the resulting correlated standard normal deviates to CUSDs using the inverse transform of a standard normal distribution (Richardson et al., 2008). The resulting vector of simulated CUSDs is used to simulate random prices and yields that are appropriately correlated. The CUSDs are used to avoid either over or under-stating the variance and mean for cash receipts if price and yield are correlated and the correlation ignored (Richardson et al., 2008; 2000). Stochastic prices and yields for maize, sorghum, and rice for each agro-ecological zone are simulated for seven years using a correlation matrix method. This method ensures that the regions are simulated using local prices and yields. Since we have three cereal crops, and three price sets the correlation matrix is a 6 × 6 dimension for each zone for each agro-ecological zone. Given six random variables and seven years, the model simulated 42 correlated yields and prices per zone using the unsorted deviations from cubic regression. Additional details for step 1, 2, 3 and 4 are in Appendix B.

Step 5: generation of random variables

Step 5 involves the combination of the deterministic forecasts and the stochastic parameters to calculate the random values for a stochastic model. It applies the CUSD to the inverse transform of the empirical distribution defined by the S_i and $p(S_i)$ using the EMP functions demonstrated in Eqs. (11) and (12). The two equations are simulated for 500 iterations using the Latin Hypercube procedure to simulate the random yields and prices for seven years. The Latin Hypercube sampling procedure segments the uniform distribution into N (500) intervals and makes sure that at least one value is randomly selected from each interval. On words, it ensures that all areas of the probability distributions are considered in the simulation (Richardson et al., 2008).

$$\tilde{Y}_{t,ij} = \hat{Y}_{t,ij}{*}(1 + EMP(S_{t,ij},\, p(S_{t,ij}),\, CUSD_{t,ij})) \tag{11}$$

$$\tilde{P}_{t,ij} = \hat{P}_{t,ij}{*}(1 + EMP(S_{t,ij},\, p(S_{t,ij}),\, CUSD_{t,ij})) \tag{12}$$

where tilde (~) represents a stochastic variable; EMP() is the Simetar function that simulates an empirical distribution defined by $S_{t,ij}$ and $p(S_{t,ij})$ using the inverse transform method. $p(S_{t,ij})$ is the frequency distribution for the fractional deviates from the cubic functional form ($S_{t,ij}$), and CUSD defined above.

Step 6: model simulation and evaluation

This step consists of checking the completeness and accuracy of the simulated values. Student's-*t*-test determines if the correlation coefficients for two matrices (historical and simulated) are statistically equal at the indicated confidence level. For example, in Appendix C, the simulated correlation coefficients were statistically equal to the historical correlation coefficients on a critical value of 2.94 at the confidence level of 99.6% (Table C1-a). It also checks if the mean of each simulated and historical variables are statistically equal at a given confidence level (Richardson et al., 2008; Richardson et al., 2000). Hence, the calculated t-test statistics in Table C1-b are all less than critical value of 2.25 so we fail to reject the null hypothesis that the simulated mean of price and yield is statistically equal to the historical mean at the 95.0% confidence level. Additional details describing the evaluation tests are provided in Appendix C.

Step 7: simulation of key output values (KOVs)

After the evaluation of the random variables used in the model, the final step consists in formulating a stochastic simulation model in order to simulate the stochastic KOVs as illustrated in Eqs. (13) to (17).

$$\tilde{\mu}_{ij} = a_{ij} * \tilde{Y}_{ij} \tag{13}$$

$$\tilde{c}_{ij} = \sum\nolimits_{\theta} (a_{t,ij} * \tilde{k}_{t,ij\theta} * (1 + r_{t,\theta})) \tag{14}$$

$$\tilde{C}_{ij} = \tilde{c}_{ij} + \text{FC}_{ij} \tag{15}$$

$$\tilde{V}_{ij} = \tilde{P}_{ij} * \tilde{\mu}_{ij} \tag{16}$$

$$\tilde{\pi}_{ij} = \tilde{V}_{ij} - \tilde{C}_{ij} \tag{17}$$

where tilde (~) indicates a stochastic variable;

$\tilde{Y}_{ij}$ = stochastic yield from Eq. (11)
i = the three crops maize, sorghum, and rice;
j = regions (Dodoma and Morogoro);
$\mu_{t,ij}$ = production for crop i for region j in year t;
$a_{\,t,ij}$ = land area devoted to crop i, for region j (in hectares) in year t;

Table 3
Summary statistics of forecasted maize yield (t/ha) for seven years.

	2019	2020	2021	2022	2023	2024	2025
Dodoma region							
Mean	1.156	1.211	1.278	1.337	1.396	1.463	1.528
STDV	0.258	0.281	0.296	0.312	0.330	0.346	0.358
CV	22.28	23.22	23.13	23.35	23.67	23.68	23.39
Min	0.647	0.682	0.716	0.751	0.786	0.820	0.855
Max	1.576	1.661	1.745	1.830	1.914	1.999	2.083
Morogoro region							
Mean	0.932	0.909	0.887	0.868	0.844	0.823	0.806
STDV	0.194	0.182	0.190	0.180	0.174	0.171	0.163
CV	20.77	20.00	21.43	20.74	20.60	20.79	20.22
Min	0.649	0.634	0.619	0.604	0.589	0.574	0.559
Max	1.350	1.319	1.287	1.256	1.225	1.194	1.162

$k_{t,ij\theta}$ = variable cash cost per ha for every input (θ) applied to crop i,
θ = inputs like land preparation, seeds, fertilizers, weeding, herbicides, transport, labor, storage and marketing cost;
$r_{t,\theta}$ = annual inflation rate in the price per unit of input$_\theta$ for year t;
$c_{t,ij}$ = total variable costs for each crop i in year t,
$FC_{t,ij}$ = fixed cost (for this study FC was set equal to zero) as the crops being mainly cultivated at small-scale level with limited machinery loans, land loans, property taxes, and insurance.
$\widetilde{C}_{t,ij}$= total production cost for each crop i, for each region j in year t;
$\widetilde{V}_{t,ij}$= total receipts or gross revenue for each crop$_i$, per region in year t;
$\tilde{P}_{t,ij}$ = stochastic price from Eq. (12)
$\widetilde{\pi}_{ij}$ = Net return for each crop i for region j.

Stochastic inflation rates for variable inputs were simulated using a uniform distribution function denoted by U(Min, Max). The uniform function is used in Simetar to return a random number between the specified minimum and maximum where each number between the range has an equal probability of being observed. It is simulated by Simetar with the = UNIFORM (min, max) function. The historical values for each cost (in Appendix A1) were used to parameterize the U (Min, Max) distribution for variable costs per hectare. Likewise, the inflation rate of values between 3.0 and 3.9% was simulated as = UNIFORM(3.0, 3.9) to generate a range of random numbers between 3.0 and 3.9. The generated random numbers were used to inflate the current total cost over the 7th forecasting horizon. Similarly, the area of production per each crop is stochastic and is simulated using a uniform distribution. For example, the total area (in ha) under maize in Dodoma region was estimated between the range of 370,000–438,000 (Table 2), therefore, a uniform function was simulated as = UNIFORM (370,000, 438,000) to generate the area for the region.

The net present value (NPV) for each crop for seven years was calculated using the stochastic net crop returns. An annual discount rate of 9% was used for calculating the present value of net crop returns across seven years. For the two agro-ecological regions, the NPV for each crop was calculated as follows:

$$N\widetilde{P}V_{ij} = \sum_{t=1}^{T}\left(\frac{\widetilde{\pi}_{t,ij)}}{(1+R)^t}\right) \tag{18}$$

where: R = discount rate of 9%

t = number of periods (1, 2, 3, …7 years)

Cumulative distribution functions (CDFs) and the probability density functions (PDFs) were used to verify the simulated variables for the KOVs. Fan graphs were developed to show how the relative variability of a stochastic variable changes over time from the year 2019 through 2025.

2.4. Probabilities of target values

Richardson and Mapp (1976) defined the probability of economic success as the probability of NPV being greater than zero. Therefore, when NPV is positive, the business earns a higher rate of return than the discount rate. In this study, we estimated the probability of NPV being negative for each crop, i.e. the probability of failure.

3. Results and discussion

The first step performed by the MVE distribution model is the simulation of cereal yields and prices for seven years from 2019 to 2025 (step 1–5) followed by an evaluation stage (step 6) before the simulation of KOVs (step 7). The strength of the MVE was not only on its ability to account for non-normally distributed random variables and historical correlations, but also its capability to handle the heteroscedasticity of random variables in two regions, and to produce results that are consistent with historical data based on the statistical tests in Appendix C.

3.1. Crop yields

Results of maize yield are presented in Table 3, which summarizes the statistics for maize yield/ha, and displays the distribution of yield in the 1st and 7th year for the two regions. The mean, standard deviation (STDV), coefficient of variation (CV), and minimum and maximum statistics are presented for each year. The CVs measure the relative risk/variability associated with yield. The relative variability of the average maize yield for each year in Dodoma and Morogoro regions correspond to a range of 22.3–23.7% and 20.0–21.4% respectively. The annual mean yield for maize is projected to increase in Dodoma from 1.156 in the first year to 1.528 t/ha in the seventh year with a minimum yield of 0.647 t/ha to 0.855 t/ha and a maximum yield of 1.576 to 2.083 t/ha through 2025. In Morogoro, the mean yield for maize is expected to decrease from 0.932 to 0.806 t/ha with a drop in the minimum yield from 0.649 to 0.559 t/ha and 1.350 to 1.162 t/ha for maximum yield through 2025.

Fig. 2 presents the probability distribution functions (PDF) for total maize production in tons for the year 2019 and 2025. The mean for each function is represented by a vertical bar at the center while vertical bars at left and right sides represent the confidence intervals at the alpha level equal to 5%. In Dodoma, the production distribution for the year 2025 is illustrated in Fig. 2-a. The average, minimum and maximum total production of maize in 2025 is likely to be 619,184 tons, 318,766 tons and 909,265 tons respectively, which is relatively higher compared to 468,142 tons, 239,686 tons and 691,579 tons in 2019 correspondingly. In Morogoro, the PDF chart for maize production in the year 2025 in terms of average (316,236 tons), minimum (212,720 tons) and maximum (467,990 tons) in Fig. 2-b contrasts with 2019 data with average (366,073 tons), minimum (246,938 tons) and maximum (546,001). This difference implies that, for the next 7 years, total production of maize is likely to increase and decrease in Dodoma and Morogoro respectively.

Rising maize production in the semi-arid area is consistent with Kilembe et al. (2013). The latter authors highlighted that by 2050, maize yield in the semi-arid part of Dodoma would experience an increase of > 25%. As for sub-humid areas, the decrease in maize production could be due to many factors including rising maximum and minimum temperatures, increasing variability of rainfall, and increasing frequency and severity of extreme events (Kahimba et al., 2015). Additionally, the decrease in maize yield could be associated with low adoption of the Conservation Agriculture Program in the country, which provides a viable means for strengthening resilience in agro-ecosystems and livelihoods that advance adaptation goals (2014).

Table 4 presents the results regarding sorghum yield for the two regions. The Dodoma region forecasts show a slight annual increase in

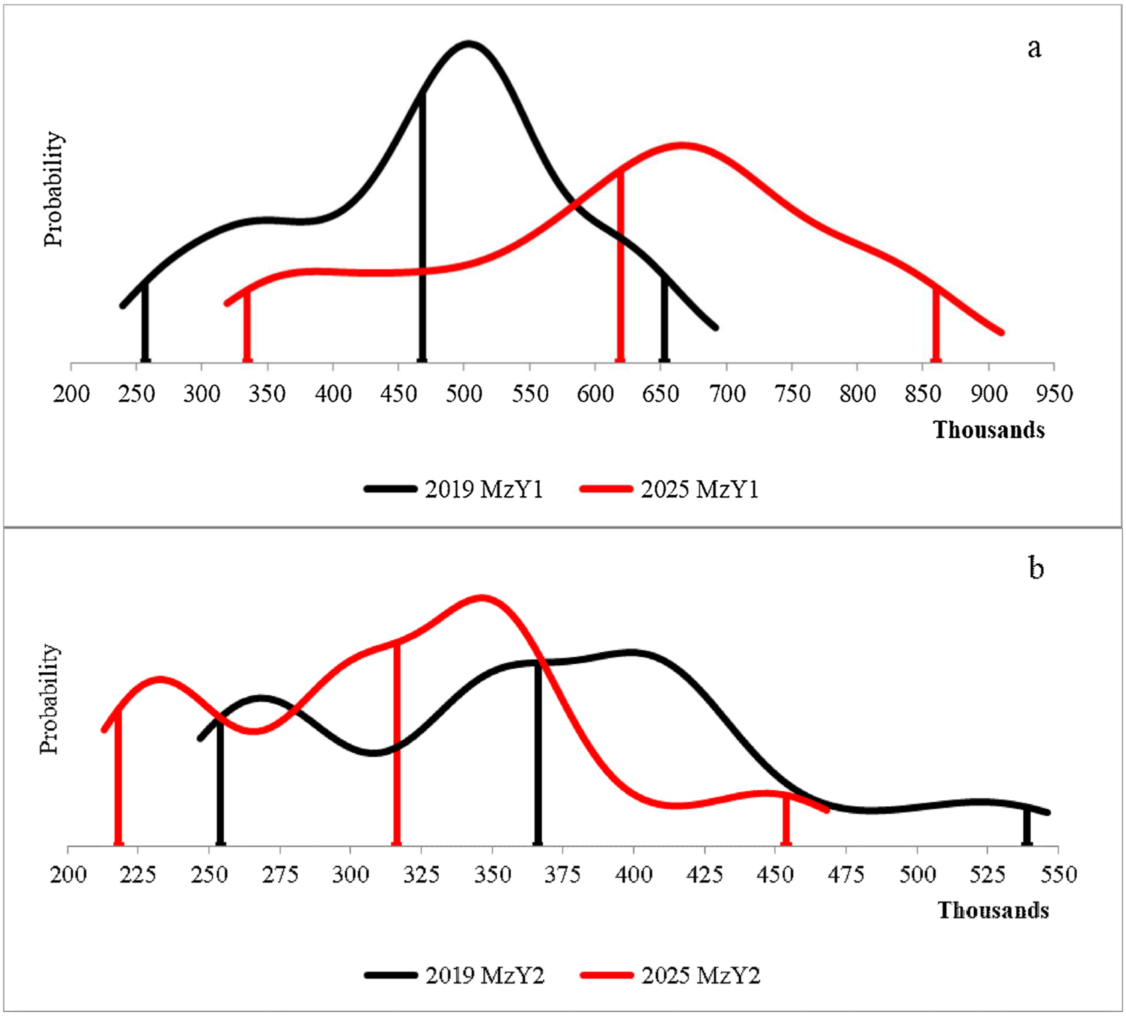

Fig. 2. PDF approximations of total maize production (in 1000 tons) for year 2019 and 2025: a = Dodoma region and b = Morogoro region; MzY1 = tons of maize in Dodoma; MzY2 = tons of maize in Morogoro.

Table 4
Summary statistics of forecasted sorghum t/ha for seven years.

	2019	2020	2021	2022	2023	2024	2025
Dodoma region							
Mean	1.167	1.193	1.222	1.250	1.273	1.301	1.333
STDV	0.197	0.211	0.206	0.212	0.214	0.231	0.236
CV	16.84	17.70	16.90	17.00	16.78	17.74	17.72
Min	0.781	0.799	0.817	0.835	0.853	0.871	0.889
Max	1.516	1.551	1.586	1.622	1.657	1.692	1.727
Morogoro region							
Mean	0.997	0.987	0.981	0.972	0.967	0.960	0.952
STDV	0.147	0.144	0.143	0.141	0.140	0.144	0.139
CV	14.76	14.56	14.53	14.47	14.49	15.02	14.64
Min	0.747	0.741	0.736	0.731	0.725	0.720	0.715
Max	1.205	1.196	1.188	1.179	1.171	1.162	1.153

sorghum yield. The average yield is between 1.167 and 1.333 t/ha, the minimum value is between 0.781 and 0.889 t/ha and the maximum value varies from 1.516 to 1.727 t/ha for the next seven years in Dodoma. Sorghum yield in Morogoro is likely to decrease between 1.997 and 0.952 t/ha on average with the minimum and maximum yield between 0.747 and 0.715 t/ha and 1.205 to 1.153 t/ha respectively. The relative risk associated with the mean yield is higher in Dodoma (16.84–17.74%) than in Morogoro (14.47–15.02%).

Fig. 3 depicts the total sorghum production distribution for the years 2019 and 2025. Dodoma's production PDF for the year 2025 lies to the right of the distribution in the year 2019 (Fig. 3-a), while, Morogoro has a production PDF for the year 2025 which lies slightly to the left of a PDF in 2019. The left shift of a PDF implies a small decrease in total sorghum production in the sub-humid region and an increase in production for the semi-arid region. The results are in agreement with observations by Msongaleli et al. (2015), where more than twenty Global Circulation Models (GCMs) and two crop models (DSSAT and APSIM) were used to assess future production of sorghum in the central Tanzania. They reported that sorghum yield is likely to increase in central Tanzania between 5.4% and 6.9% in the near-term (2010–2040). However, the study didn't develop much about the production of sorghum in sub-humid areas. Elsewhere in Africa, sorghum production is projected to increase in semi-arid areasby a range of 19 to 72% across Eastern and Southern Africa (Turner and Rao, 2013; Zinyengere et al., 2014). The increase in sorghum yields reported by different scholars under different climate change scenarios may be attributed to increases in temperatures and the slight changes in projected rainfall which appear to create conducive conditions for sorghum growth, being more tolerant to heat and water stress (Msongaleli et al., 2015).

Results on rice yield are presented in Table 5. Rice yield is likely to increase for all regions by 2025. A small increase in mean yield between 1.112 and 1.278 t/ha was observed in Dodoma with a minimum of 0.727 to 0.836 t/ha and a maximum yield of 1.584 to 1.821 t/ha. Similarly, the mean yield of rice in Morogoro is expected to increase from

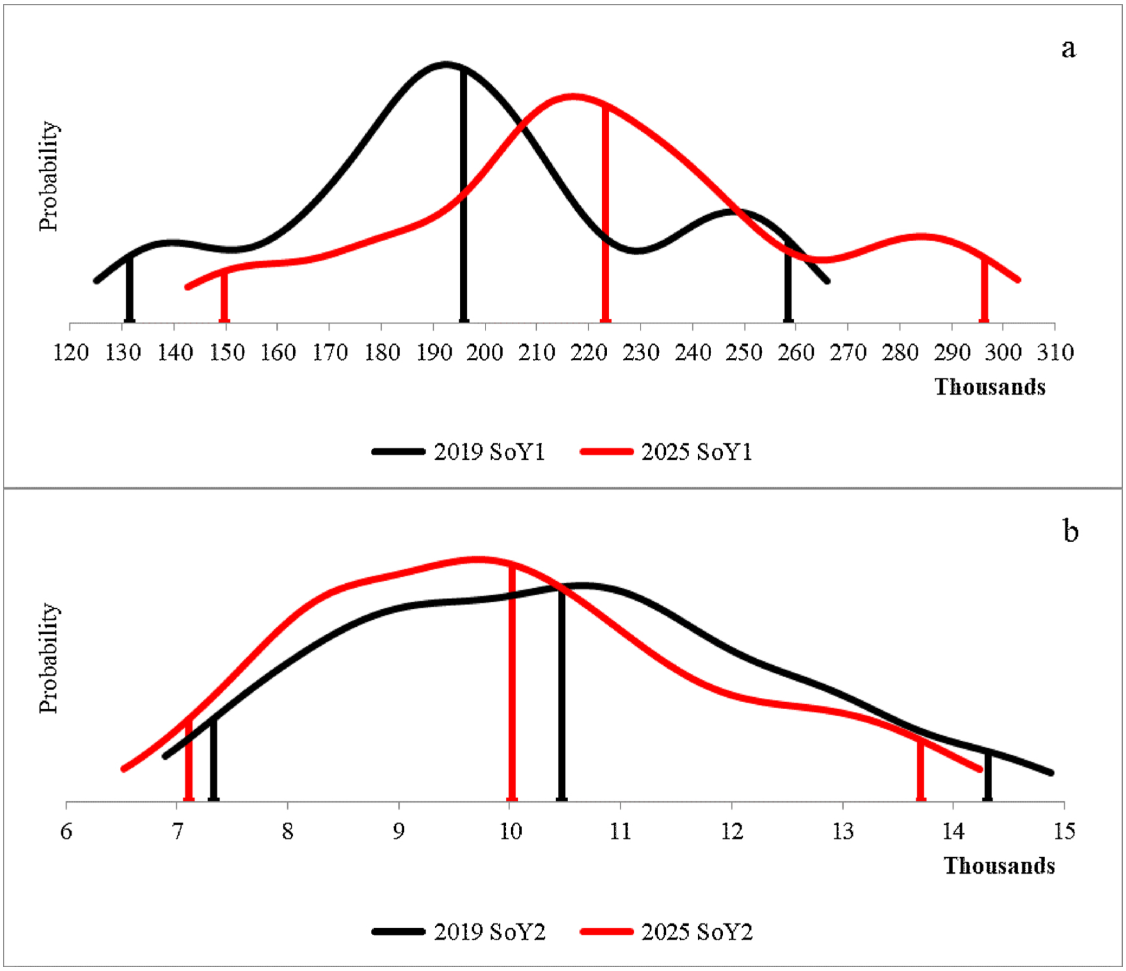

Fig. 3. PDF approximations of total sorghum production (in 1000 tons) for year 2019 and 2025: a = Dodoma region and b = Morogoro region; SoY1 = tons of sorghum in Dodoma; SoY2 = tons of sorghum in Morogoro.

Table 5
Summary statistics of forecasted rice yield (t/ha) for seven years.

	2019	2020	2021	2022	2023	2024	2025
Dodoma region							
Mean	1.112	1.140	1.167	1.195	1.219	1.249	1.278
STDV	0.236	0.248	0.254	0.266	0.264	0.272	0.294
CV	21.21	21.76	21.77	22.31	21.63	21.76	22.97
Min	0.727	0.745	0.763	0.781	0.799	0.818	0.836
Max	1.584	1.623	1.663	1.702	1.742	1.781	1.821
Morogoro region							
Mean	2.252	2.349	2.441	2.544	2.635	2.742	2.839
STDV	0.366	0.391	0.409	0.413	0.442	0.445	0.483
CV	16.26	16.63	16.75	16.24	16.76	16.22	17.01
Min	1.627	1.698	1.769	1.841	1.912	1.983	2.054
Max	2.879	3.005	3.131	3.257	3.382	3.508	3.634

2.252 t/ha in 2019 to 2.839 t/ha in 2025. The maximum yield would attain 3.634 t/ha by 2025.

The PDFs in Fig. 4 portray the total production of rice (in tons) in the two regions. The PDF for the year 2025 in Morogoro lies more to the right of the year 2019 indicating an essential increase for rice in the region (Fig. 4-b). In Dodoma, the distribution of total rice production for 2025 is expected to be 7794 tons, 14,037 tons and 25,429 tons for the minimum, average and maximum respectively (Fig. 4-a). Rice production in Morogoro is forecasted to range between 337,747 tons, 542,220 tons, and 802,842 tons for minimum, average and maximum values respectively. This implies a substantial increase in rice production predominantly in the sub-humid region. However, the rising in rice production in semi-arid areas is consistent with Lamboll et al. (2001). They argue that an increasing trend in low rice-producing areas like Dodoma may have been influenced by many factors including the rice irrigation projects funded mainly by the International Fund for Agricultural Development (IFAD). As rice has become an important cash and food crop, farmers have increased their productivity by adopting technologies such as the application of improved seeds and system of rice intensification (SRI) in addition to having favorable markets and timely payments. Furthermore, since rainfall is unfavorable to semi-arid areas, farmers have learned to collect runoff water and divert it into bunded fields or paddies to facilitate the storage of water for rice growing. This also prevents erosion which may occur when the runoff catchment area becomes too big.

3.2. *Crop prices*

Prices for all three crops for each region are likely to increase throughout the forecasting horizon (Table D1). A higher price of maize is forecasted in Dodoma compared to Morogoro. The results show an increasing average price of TZS 0.719 to 0.904 million/t and TZS 0.659 to 0.817 million/t for the two regions, respectively, over the next seven years. The minimum and maximum prices for Dodoma are forecasted at TZS 0.518–0.651 million/t and TZS 0.938–1.178 million/t, respectively. The minimum maize price for Morogoro is expected to range between TZS 0.454–0.562 million/t with a maximum price ranging

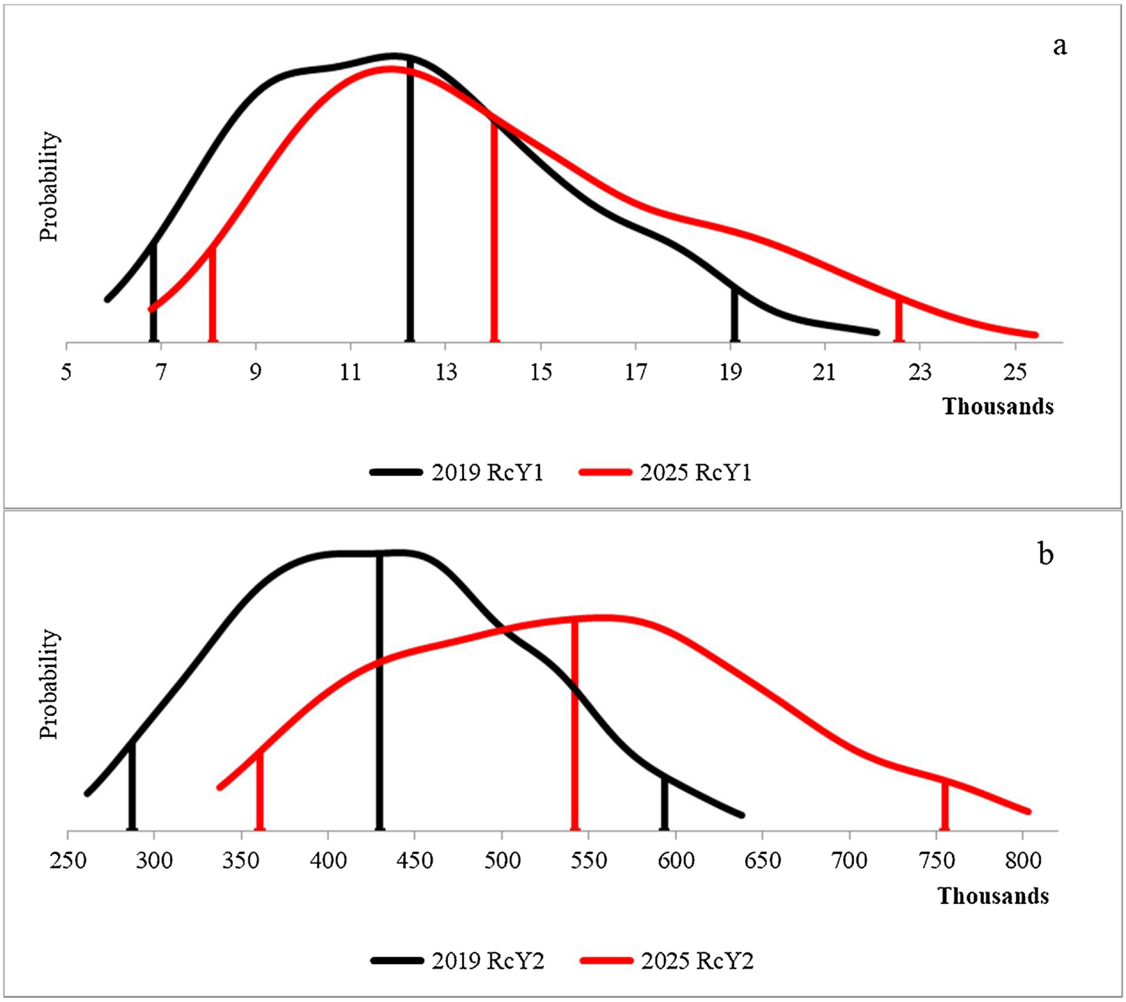

Fig. 4. PDF approximations of rice production (in 1000 tons) for year 2019 and 2025: a = Dodoma region and b = Morogoro region; RcY1 = tons of rice in Dodoma; RcY2 = tons of rice in Morogoro.

between TZS 0.825 to 1.021 million/t. The higher price of maize in Dodoma could be influenced by many factors, including the increasing population in the region. The growing population in Dodoma is likely to be associated with the total shift of the administrative activities of the Tanzanian government from the Dar es Salaam city, accelerating to higher demand for food.

On the other hand, the price for sorghum is generally higher in Morogoro than Dodoma. The forecasting results show that the average price in Morogoro would increase up to TZS 1.706 million/t with a minimum of TZS 1.468 million and the maximum of TZS 1.992 million/t by 2025. This increase approximately represents a 36% increase from the price in 2019 (Table D1). Dodoma would experience an average of TZS 1.047 million/t with a minimum of TZS 0.722 million. The maximum price would go up to TZS 1.402 million/t by 2025, correspondingly to a 33% increase from the 2019 price. The relative variability of sorghum price is twice as high in Dodoma (18.05%) than Morogoro (9.80%). The higher sorghum price in Morogoro could be due to low sorghum production in the region compared to Dodoma. The high price could also be influenced by the redoubled demand for sorghum in different products as a results of educational efforts regarding the nutritional and health benefits. The Sokoine University of Agriculture has been one of the organizations campaigning for sorghum consumption (Nkuba, 2009; Noel, 2015; Kinabo et al., 2016; Mahenge, 2018).

Of all the three crops, the price for rice was the highest, with a slight difference between the two regions as exposed in Table D1. The average annual price reflects an increasing trend of TZS 1.967 to 2.407–million/t in Dodoma and between TZS 1.851–2.338 million/t in Morogoro. The maximum price in the regions is projected to be between TZS 2.607–3.190 million/t and TZS 2.344 to 2.959 million/t for Dodoma and Morogoro, respectively. The stochastic prices are also presented in Fig. D1, and reveal how the relative variability of stochastic crop prices changes from 2019 to 2025. Except for sorghum price in Morogoro, rice price shows a lower relative variability of the average price in comparison with other crops. Generally, the findings on price forecasts are in agreement with observations by Von Braun (2008). He reported that important cereals in most of developing countries would experience a price increase of up to 30% by 2020 mainly due to a recession and reduced investment in agriculture and this could push 16 million more children into malnutrition.

3.3. *Crop profitability*

Table E1 in Appendix E summarizes the statistics for annual net returns per crop per region. The stochastic results for the maize crop in Dodoma demonstrate an increase in the annual mean income from TZS 121 billion in the first year to TZS 192 in the 7th year with a negative minimum net return of TZS -52 billion to TZS -5 billion in the first three years. One observes positive minimum values increasing from TZS 5 to 37 billion for the last four years. The maximum annual net returns would range between TZS 345 to 441 billion by 2025. In Morogoro region, the analysis forecasts a decrease in annual average return from TZS 45 to 36 billion with negative minimum returns oscillating between

TZS - 91 and -51 billion. The maximum values decrease between TZS 225 to 150 billion across the seven-year period. The relative risk associated with annual income for maize in terms of CV is higher in Morogoro (decreasing from 118.4 to 95.4%) than Dodoma (decreasing from 55.7 to 34.9%). These changes suggest that although there will likely be a slight improvement in annual returns for maize crop in the two regions, the risk is higher in Morogoro in 2025. A probability of netative annual net return for maize decreases from 4.6% to 0.2% in the first five years in Dodoma, with a zero probability in the last two years. In Morogoro, a probability of annual net return being less than zero were higher compared to Dodoma. These results imply that Morogoro is more vulnerable in maize profitability than Dodoma.

Likewise, rice crop forecasts an increasing annual net returns for both regions, although the crop has a higher risk associated with the mean return, especially in the Dodoma region. The rice net return CV for Dodoma is 62.1% and highest in the first four years (87.6–60.1%). Meanwhile, Morogoro has a rice net return CV of 22.7% with a zero probability of negative annual net returns throughout the forecasting period. Dodoma is likely to experience a 12.7% to 0.4% probability of negative annual returns from 2019 to 2024 respectively, with a zero probability in 2025. Furthermore, sorghum has positive minimum values for net returns throughout the forecasting period for the two regions. Regardless of high variability of the mean annual return (40.9%), the mean, minimum and maximum annual return is expected to increase in Dodoma region between TZS 82.0 to 92.0 billion, TZS 1.4 to 23.6 billion and 189.1 to TZS 199.6 billion respectively. In Morogoro, the study projects a small decrease for sorghum annual net return. The average, minimum and maximum annual return is expected to decrease between TZS 7.9 to 6.8 billion, TZS 3.4 to 3.0 billion and TZS 14.4 to 12.5 billion, respectively. The CVs for sorghum annual returns are lower in Morogoro (29.1%) than Dodoma (40.9%).

The results regarding the relative variability of annual net returns for seven years for the three crops per region are also presented with fan graphs in Fig. E1. The 5th and 95th percentiles for the stochastic forecast display the lower and upper bounds of a 90% confidence interval for the forecast of the average annual net return. The fan graph for maize annual net shows there is a 5% probability (at 5th percentile) that the net returns will be negative for the year 2019 in Dodoma (Fig. E1-A1) and the first four years in Morogoro (Fig. E1-A2). Likewise, the fan graph for rice indicates a 5% probability of rice annual net turn being negative for the first three years in Dodoma (Fig. E1-C1) with zero probability in Morogoro (Fig. E1-C2).

Table 6 presents the NPV distribution results for each crop, for each region for the seven years. The relative variability of the mean NPV for rice is higher in Dodoma (30.2%) than in Morogoro (13.2%). The CV values for sorghum is 15% for the two regions. The CV for maize NPV in Morogoro is 49.7%, and 18.6% in Dodoma. This difference implies that maize has a higher relative variability (risk) for NPV among the three crops in Morogoro than the Dodoma region. Rice has a higher relative NPV risk in Dodoma (30.2%) than in Morogoro (13.2%). The three crops have positive means for NPV values, with the minimum value being negative only for maize in Morogoro (TZS - 120.45 billion).

Table 6
Summary statistics of NPV (billions TZS).

	Dodoma region			Morogoro region		
	Maize	Sorghum	Rice	Maize	Sorghum	Rice
	(×1000,000,000)			(×1000,000,000)		
Mean	1128.18	620.84	50.56	286.01	51.90	3250.83
STDV	210.15	93.34	15.25	142.05	7.84	428.85
CV	18.63	15.03	30.16	49.67	15.11	13.19
Minimum	520.82	354.73	15.83	(120.45)	33.11	2255.32
Maximum	1687.42	902.64	95.64	650.40	76.62	4494.69
P(< 0)	–	–	–	2.41%	–	–

Fig. 5 presents the CDF charts for NPV of all three crops for the two regions in which Fig. 5-a and -b represent the distribution of NPV for crops in Dodoma region and Morogoro region, respectively. Maize is the highest income generator in the Dodoma region, followed by sorghum with rice being the least (Fig. 5-a). In contrast, rice in Morogoro is the highest income generator of the main cereals in the region (Fig. 5-b). Additionally, all three crops in Dodoma have a zero probability that the NPV will be negative. Meanwhile, the probability of a negative NPV is 2.41% for maize and zero for rice and sorghum in Morogoro (Table 6). In general, the distribution functions indicate there is considerable risk or variability for the three crops. The study reveals that the three crops have a high chance of economic success, however, the sub-sectors call for a tailor-made campaign to invest in technologies that will help reduce the risks and variability associated with yields and hence increase crop net returns.

4. Conclusions

The purpose of this paper was to apply a stochastic model to simulate and forecast stochastic yields, prices and net returns of three food staples in Tanzania from 2019 to 2025. The paper describes the usefulness of using a stochastic simulation approach under a Monte Carlo simulation protocol in delivering accurate and timely forecasts. The study was conducted using 11 years of historical data for maize, sorghum, and rice crops in Dodoma and Morogoro regions of Tanzania. The variables were simulated for seven years using stochastic variables to estimate the distribution of the future yields, prices and net returns per crop by region. The simulation model was run for 500 iterations using Latin Hypercube sampling procedure, and the simulated statistics and correlation matrix were compared to the historical input values. The comparison was done for validation purposes to ensure that the random variables are simulated correctly and to demonstrate the appropriate properties of the parent distribution. The validation statistics showed that the stochastic procedure to simulate the means and historical correlations were met.

Our study enables to predict the evolution of major cereal crops over the next seven years, in terms of grain yield and economic success. The results on yields for all three crops show a slightly increasing trend and the prices are likely to increase for both regions. Moreover, the results showed a high probability of success for all three crops regardless of the small probability of negative net returns for maize and rice. Despite the probability of success for crops, there is a need to increase investment in relation to farm management practices. If no alternative risk management strategies are available, the productivity of main cereal crops will continue to experience a high degree of risk and variability.

The data used in this study is based on aggregates at regional levels, focusing on semi-arid and sub-humid agro-ecological zones. This may generally lead to a downward bias in the estimation of yields and incomes regarding districts and villages specialized in crop production. With the availability of farm-level panel data sets, further analyses should also be directed at district, village or site-specific levels to estimate risks faced by farmers.

The methodology used in this study can be modified to simultaneously simulate an array of random variables using a multivariate empirical distribution. Further studies should also contemplate non-cereal and cash crops like pulses, cassava, banana, potatoes, coffee, cotton, sisal and sugarcane which are important crops in generating domestic and foreign income in the country. One of the most hindering factors in achieving different agricultural development strategies in Tanzania is the lack of an appropriate model to provide a roadmap of what would happen in the year ahead and beyond. Therefore, this study paves a way to a complete stochastic simulation model that comprises the country's essential crops. It provides accurate information to policymakers, particularly the national food security division, which is responsible for ensuring food availability in the country at all levels.

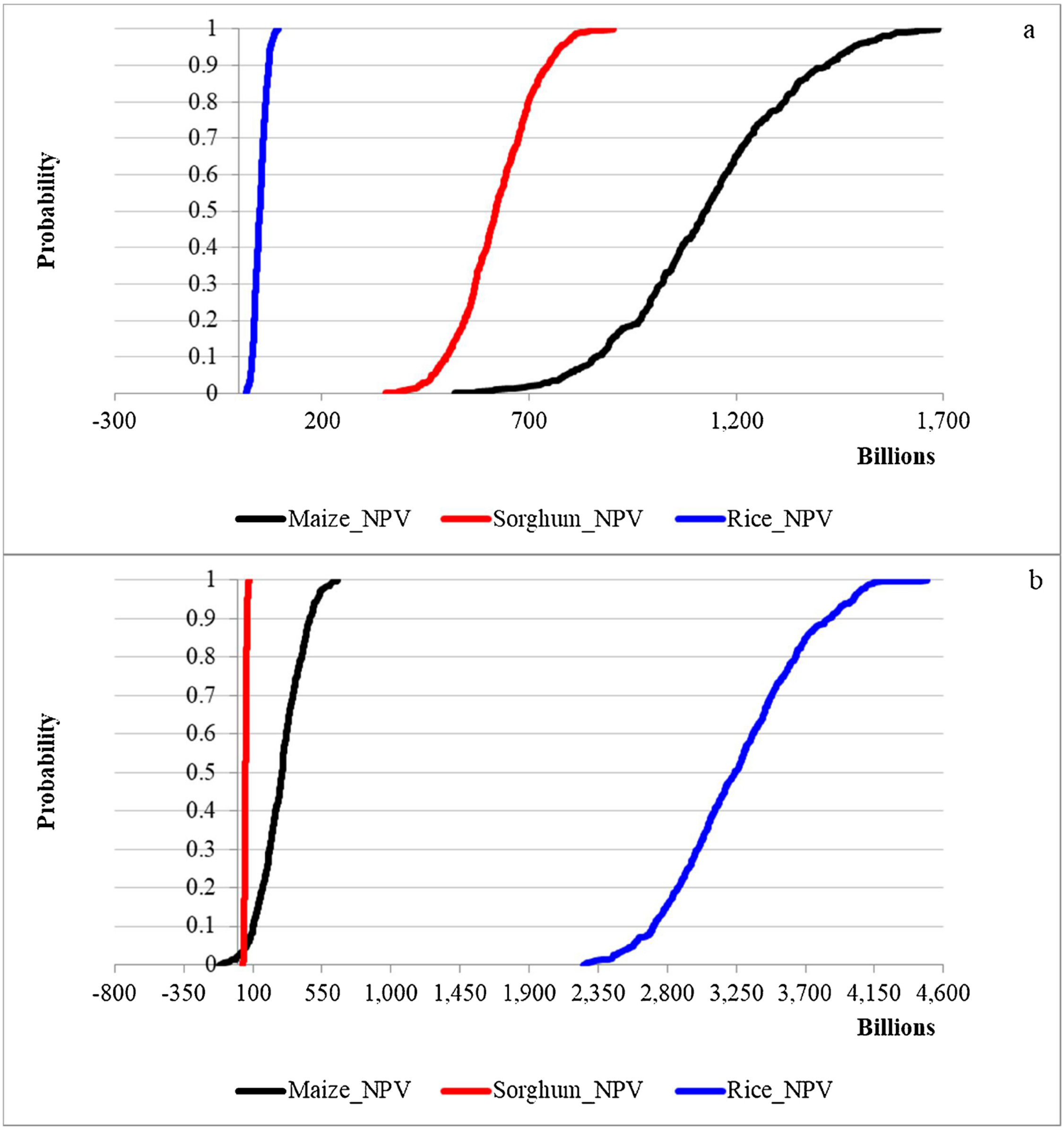

Fig. 5. CDF of NPV for all three crops in Dodoma region (a) and Morogoro region (b) for the seven years up to 2025.

The model could also help agricultural stakeholders to plan for imports and exports, and government aid for farmers. Furthermore, this study utilized historical yields and prices for cereal crops in the regions analyzed. It was assumed that the relative variability of yield would be the same in the future as it has been in the past and that the differences in yields within 11 years represent the effects of weather variability. Further integrated assessments that integrate climatic conditions like precipitation, maximum and minimum temperatures are still needed to provide accurate forecasts beyond a seven-year period.

Acknowledgments

The authors thank the Tanzania National Bureau of Statistics, and the Ministry of Agriculture for providing data for this research. We acknowledge the Feed the Future Innovation Lab for Small-Scale Irrigation (ILSS) Project (https://ilssi.tamu.edu/workshops/tanzania/) funded by the U.S. Agency for International Development (USAID) for conducting training on the methodology described in this publication. The authors would like to sincerely thank the anonymous two reviewers and the two editors, Santiago Dogliotti and Sylvie Lupton for their helpful comments. The content of this article is the sole responsibility of the authors and does not necessarily represent the views of the NBS, ILSS and the USAID.

Appendix A. Production cost distribution per crop per region

Since the three crops are usually cultivated at the local level, usually low technology is used. Of all crops, rice is the costliest in production occupying > 65% of the total costs, followed by maize (20%) the remaining proportion is for sorghum. Farmers do not have the same costs of production for each crop per unit area, some farmers use minimal costs and some use higher costs depending on their income and wiliness to apply the needed inputs. Because, we have a range of costs in terms of minimum and maximum, Uniform Distribution Function U(Min, Max) was used to simulate the distribution of each cost category used. Simetar provides a platform to simulate a uniform distribution using the command = UNIFORM (min, max). For example, land preparation for maize crop per ha in Dodoma has a range of TZS 50,000 to 80,000, in Simetar this range was simulated as = UNIFORM(50,000, 80,000) to get random costs between the two costs and avoid using an average of the two costs. Likewise, fertilizer cost between 0 and 100 was simulated as = UNIFORM(0,100). Table A1 presents a range of costs used for inputs and labor, the simulated total costs and the proportion of each input and labor cost used in the mode. Production costs were generally high in Morogoro than Dodoma for maize and rice. The production cost for sorghum almost the same for all the two regions (Fig. A1).

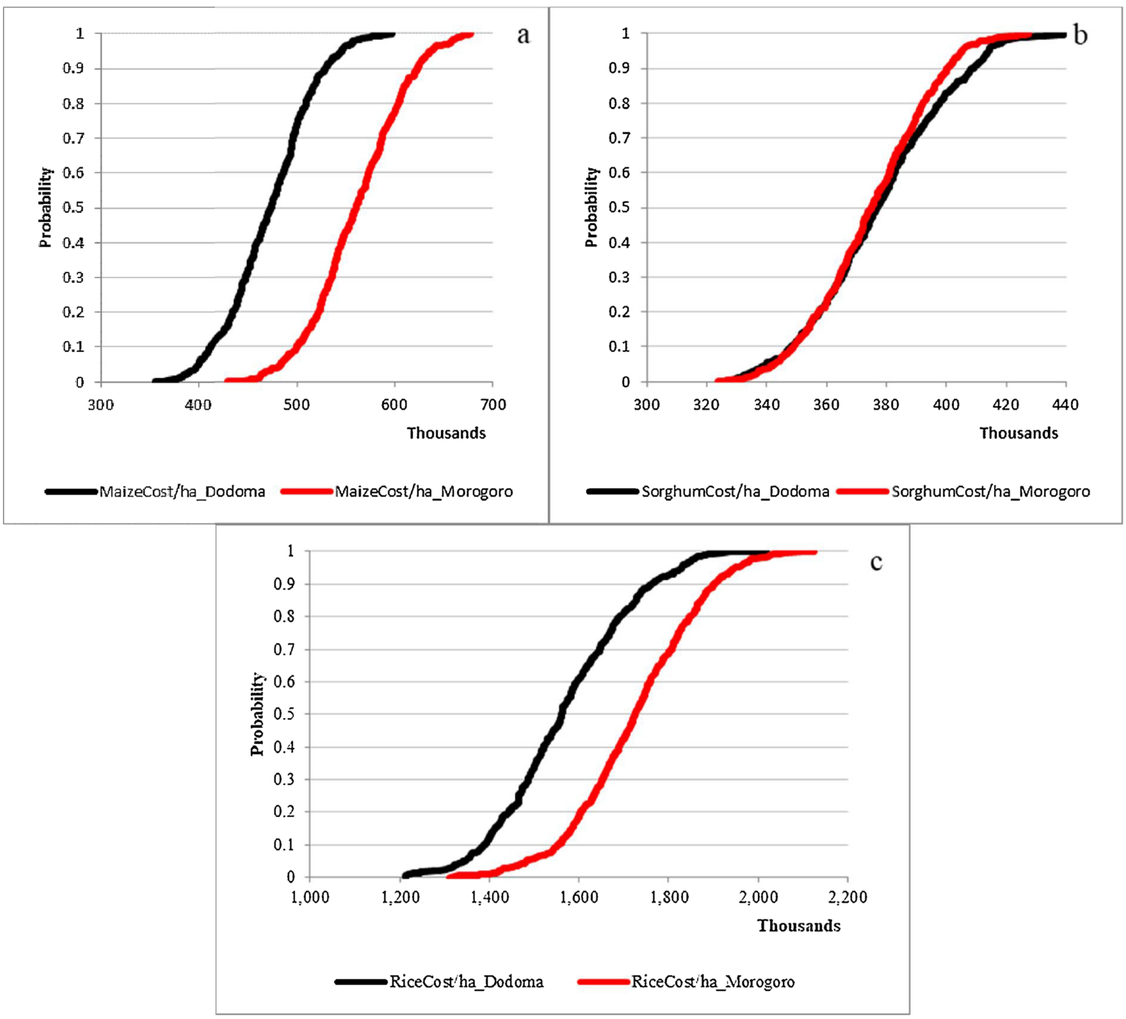

Fig. A1. CDF of per ha production costs: a for maize; b for sorghum; and c for rice.

Input and labor costa	Dodoma			Morogoro		
	Maize	Sorghum	Rice	Maize	Sorghum	Rice
	TZS/ha ($\times$ 1000)			TZS/ha ($\times$ 1000)		
Land preparation	50–80	50–80	80–200	60–100	45–90	80–250
Seeds	45–100	30–60	130–285	80–110	45–60	150–300
Planting	30–60	15–40	150–400	30–60	20–30	200–400
Weeding	80–120	110–140	300–540	60–190	80–120	350–600
Bird scaring	0	0	120–200	0	0	120–200
Fertilizers	0–100	0	0	50–100	0–50	0
Pesticides	0–60	0	0	12–60	0–24	0
Harvesting	30–80	50–90	100–380	20–60	60–90	120–350
Postharvest handling	40–70	30–60	50–200	50–80	50–60	80–250
Simulated total cost using = UNIFORM() Command						
	TZS/ha	TZS/ha	TZS/ha	TZS/ha	TZS/ha	TZS/ha
Mean	472.50	377.50	1567.48	561.00	375.00	1725.01
STDV	44.15	22.53	145.93	47.11	19.65	140.09
CV	9.34	5.97	9.31	8.40	5.24	8.12
Minimum	355.16	327.56	1212.13	429.66	323.97	1310.34
Maximum	596.53	439.49	2017.61	676.49	427.64	2123.92
Percentage of input and labor cost per ha per crop						
	Percent	Percent	Percent	Percent	Percent	Percent
Land preparation	12.5	17.6	7.0	14.3	22.5	5.8
Seeds	16.1	14.5	13.0	16.9	14.1	15.8
Planting	8.0	9.9	21.7	8.0	7.7	19.0
Weeding	21.3	30.4	27.7	22.3	24.5	23.3
Bird scaring	0.0	0.0	8.8	0.0	0.0	10.7
Fertilizers	13.1	0.0	0.0	13.4	0.0	0.0
Pesticides	6.3	0.0	0.0	6.4	0.0	0.0
Harvesting	10.3	16.7	13.3	7.1	16.1	17.3
Postharvest handling	9.1	10.9	8.6	11.6	15.2	8.0
Total	100.0	100.0	100.0	100.0	100.0	100.0

Appendix B. Steps for estimating the parameters for an MVE distribution

The steps expressed in Section 2.3 are presented in Table B1. The residuals/deviates of prices and yields from cubic regression were used in the model to develop the stochastic component of the MVE.

Obs.	Yields for Dodoma (in t/ha)			Price for Dodoma ($\times$ 100,000)			Yields for Morogoro (in t/ha)			Prices Morogoro ($\times$ 100,000)		
	MzY_1	SoY_1	RcY_1	MzP_1	SoP_1	RcP_1	MzY_2	SoY_2	RcY_2	MzP_2	SoP_2	RcP_2
Unsorted deviations from cubic regression step 1 (from Eqs. (4) and (8))												
1	0.067	0.086	−0.046	−0.139	−0.280	0.598	0.126	0.123	0.003	−0.327	−0.309	−0.083
2	−0.216	−0.262	−0.121	0.340	0.388	−0.230	−0.394	−0.248	−0.249	0.474	0.163	0.595
3	0.090	0.184	0.290	−0.231	−0.038	−1.676	0.156	0.046	0.467	−0.098	0.612	−1.091
4	0.293	0.106	0.184	−0.450	−0.105	−0.743	0.470	0.065	0.034	−0.132	−0.303	−0.804
5	−0.191	−0.117	−0.276	0.526	−0.030	3.586	−0.270	0.026	−0.194	0.461	−0.223	3.239
6	−0.186	−0.001	−0.137	1.180	1.389	0.736	−0.209	0.076	−0.444	0.735	0.903	−0.501
7	0.120	0.033	−0.019	−1.144	−1.285	−3.014	0.034	0.003	0.345	−1.556	−0.797	−2.842
8	0.048	−0.092	−0.009	−1.026	−1.356	0.403	0.062	−0.136	0.092	−0.814	−1.702	0.763
9	−0.037	0.034	0.099	0.147	0.235	−0.367	−0.009	−0.043	0.020	0.686	1.443	0.913
10	0.032	0.070	0.188	1.621	2.320	1.086	0.098	0.121	−0.080	1.567	1.071	0.149
11	−0.020	−0.040	−0.153	−0.826	−1.238	−0.379	−0.066	−0.033	0.005	−0.996	−0.859	−0.339
Unsorted deviations from cubic regression as a percent of predicted step 1 (Eqs. (5) and (9))												
1	0.272	−0.195	−0.007	−0.086	−0.109	−0.121	0.112	0.211	0.116	−0.044	−0.058	−0.040
2	−0.438	−0.331	−0.040	−0.003	0.015	−0.054	−0.303	−0.156	−0.209	0.049	−0.003	0.050
3	0.066	0.299	0.427	−0.119	−0.094	−0.087	0.183	0.079	0.282	−0.116	0.067	−0.091
4	0.370	0.262	0.238	−0.115	−0.060	0.020	0.449	0.046	−0.041	−0.090	−0.070	−0.055
5	−0.304	0.022	−0.345	0.139	0.016	0.324	−0.261	−0.043	−0.166	0.092	−0.030	0.265
6	−0.231	0.094	−0.258	0.301	0.334	0.109	−0.246	−0.036	−0.276	0.201	0.136	−0.026
7	0.188	0.066	−0.177	−0.095	−0.121	−0.154	−0.046	−0.127	0.220	−0.197	−0.054	−0.189
8	0.116	−0.114	−0.165	−0.069	−0.118	0.030	−0.026	−0.249	0.092	−0.038	−0.140	0.052
9	0.016	−0.036	0.000	0.084	0.099	−0.041	−0.069	−0.102	0.052	0.195	0.167	0.054
10	0.042	−0.009	0.216	0.221	0.311	0.022	0.118	0.174	−0.024	0.250	0.094	0.003
11	−0.091	−0.063	0.118	−0.281	−0.312	−0.062	0.091	0.204	−0.041	−0.312	−0.125	−0.028
Sorted deviations from cubic regression as a percent of predicted step 2 (Eqs. (6) and (10)) and Step 3 ($Pmin_{\hat{e}}$, $Pmax_{\hat{e}}$),												
F(x)	MzY_1	SoY_1	RcY_1	MzP_1	SoP_1	RcP_1	MzY_2	SoY_2	RcY_2	MzP_2	SoP_2	RcP_2
0	−0.438	−0.331	−0.345	−0.281	−0.312	−0.154	−0.303	−0.249	−0.276	−0.312	−0.140	−0.189
0.045	−0.438	−0.331	−0.345	−0.281	−0.312	−0.154	−0.303	−0.249	−0.276	−0.312	−0.140	−0.189
0.136	−0.304	−0.195	−0.258	−0.119	−0.121	−0.121	−0.261	−0.156	−0.209	−0.197	−0.125	−0.091
0.227	−0.231	−0.114	−0.177	−0.115	−0.118	−0.087	−0.246	−0.127	−0.166	−0.116	−0.070	−0.055
0.318	−0.091	−0.063	−0.165	−0.095	−0.109	−0.062	−0.069	−0.102	−0.041	−0.090	−0.058	−0.040

0.409	0.016	−0.036	−0.040	−0.086	−0.094	−0.054	−0.046	−0.043	−0.041	−0.044	−0.054	−0.028
0.500	0.042	−0.009	−0.007	−0.069	−0.060	−0.041	−0.026	−0.036	−0.024	−0.038	−0.030	−0.026
0.591	0.066	0.022	0.000	−0.003	0.015	0.020	0.091	0.046	0.052	0.049	−0.003	0.003
0.682	0.116	0.066	0.118	0.084	0.016	0.022	0.112	0.079	0.092	0.092	0.067	0.050
0.773	0.188	0.094	0.216	0.139	0.099	0.030	0.118	0.174	0.116	0.195	0.094	0.052
0.864	0.272	0.262	0.238	0.221	0.311	0.109	0.183	0.204	0.220	0.201	0.136	0.054
0.955	0.370	0.299	0.427	0.301	0.334	0.324	0.449	0.211	0.282	0.250	0.167	0.265
1	0.370	0.299	0.427	0.301	0.334	0.324	0.449	0.211	0.282	0.250	0.167	0.265
Correlation matrix estimated using the residuals from cubic regression (Step 4)												
	MzY_1	SoY_1	RcY_1	MzP_1	SoP_1	RcP_1	MzY_2	SoY_2	RcY_2	MzP_2	SoP_2	RcP_2
	1	0.71	0.69	−0.50	−0.30	−0.54	0.98	0.55	0.74	−0.89	−0.51	−0.81
MzY_1		1	0.73	−0.04	0.11	−0.37	0.73	0.98	0.76	−0.55	0.12	−0.89
SoY_1			1	−0.01	0.23	−0.50	0.10	0.10	0.09	−0.05	−0.03	−0.07
MzP_1				1	0.96	0.57	−0.80	−0.30	−0.81	0.99	0.74	0.52
SoP_1					1	0.37	−0.53	−0.06	−0.74	0.90	0.83	0.19
RcP_1						1	−0.78	−0.78	−0.68	0.64	−0.08	0.98
MzY_2							1	0.51	0.57	−0.22	−0.12	−0.45
SoY_2								1	0.09	0.19	0.30	−0.20
RcY_2									1	−0.57	−0.26	−0.52
MzP_2										1	0.81	0.49
SoP_2											1	0.09
RcP_2												1
Deterministic forecasts $\widehat{Y}_{t,\mathrm{ij}} = \hat{a} + \hat{b}\,T_{t,\mathrm{ij}}$ and $\widehat{P}_{t,\mathrm{ij}} = \hat{a} + \hat{b}\,T_{t,\mathrm{ij}}$ forecasts without risk (error term)												
	MzY_1	SoY_1	RcY_1	MzP_1 (x 100,000)	SoP_1	RcP_1	MzY_2	SoY_2	RcY_2	MzP_2 (x 100,000)	SoP_2	RcP_2
12	1.1505	1.1676	1.1100	7.205	7.909	19.694	0.932	0.995	2.246	6.596	12.521	18.520
13	1.2122	1.1947	1.1377	7.513	8.342	20.428	0.910	0.988	2.345	6.858	13.280	19.331
14	1.2739	1.2217	1.1653	7.822	8.775	21.162	0.889	0.981	2.443	7.120	14.039	20.141
15	1.3356	1.2488	1.1930	8.130	9.208	21.896	0.867	0.974	2.541	7.382	14.798	20.952
16	1.3973	1.2758	1.2207	8.438	9.641	22.630	0.845	0.966	2.639	7.644	15.557	21.763
17	1.4590	1.3029	1.2484	8.747	10.074	23.365	0.824	0.959	2.738	7.906	16.316	22.573
18	1.5207	1.3300	1.2761	9.055	10.507	24.099	0.802	0.952	2.836	8.168	17.074	23.384

Appendix C. Model simulation and evaluation

This appendix provides the evaluation tests used to check for completeness, accuracy, and forecasting ability of the model we developed in Appendix B. The evaluation of the model indicates that the MVE procedures appropriately correlated the random variables as none of the Student t statistics in Table C1-a are greater than the critical value of 2.94. The *t*-Tests of the means for the random variables in the year 2019 indicated that the simulated means are statistically equal to their deterministic means at the 95% level. The test statistics (test values) are less than the critical value of 2.25 for the random variables. The p-values are also higher than 0.1 at the alpha equal to 5% ($p > .05$), then we failed to reject the null hypothesis that the means are equal (Table C1-b). The remaining seven years have similar results, hence based on the evaluation tests, the simulated yields and prices can be used for future decision-making, particularly national and household food production trend over time. The null and alternative hypotheses for the Student's-t test are as follows:

Ho: $\widehat{Y}_{\mathrm{ij}} = Y_{\mathrm{ij}}$

HA: $\widehat{Y}_{\mathrm{ij}} \neq Y_{\mathrm{ij}}$

Ho: $\widehat{P}_{\mathrm{ij}} = P_{\mathrm{ij}}$

HA: $\widehat{P}_{\mathrm{ij}} \neq P_{\mathrm{ij}}$

Ho: $\rho_{\widehat{Y}\widehat{P},\mathrm{ij}} = \rho_{\mathrm{YP,ij}}$

HA: $\rho_{\widehat{Y}\widehat{P},\mathrm{ij}} \neq \rho_{\mathrm{YP,ij}}$

where: $\widehat{Y}_{\mathrm{ij}}$and $\widehat{P}_{\mathrm{ij}}$is the simulated mean yield and price for crop i, for region j, respectively;$Y_{ij}P_{ij}$is the mean from historical yield and price for crop i; for region j, respectively;$\widehat{\rho}_{\widehat{Y}\widehat{P},\mathrm{ij}}$is the individual correlation coefficient between the simulated variables for i and j and $\rho_{YP,\ ij}$is the historical correlation coefficient between variables i and j used to simulate the multivariate distribution.

a) Test correlation coefficient

Confidence level			99.6586%								
Critical value			2.94								
			Dodoma region				Morogoro region				
	SoY1	RcY1	MzP1	SoP1	RcP1		SoY2	RcY2	MzP2	SoP2	RcP2
MzY1	1.07	0.81	0.53	0.12	1.00	MzY2	1.78	0.63	0.38	2.34	2.06
SoY1		1.94	0.37	1.00	1.08	SoY2		0.90	0.07	1.44	1.29
RcY1			0.13	0.67	1.73	RcY2			0.59	1.30	1.08
MzP1				1.44	0.29	MzP2				1.26	0.08
SoP1					0.08	SoP2					0.63

b) Test parameters (test for simulated vs deterministic means)

Confidence level			95.0000%	Critical value		2.25

Simulated vs deterministic maize mean yield for 2019 in Dodoma				
	Given value	Test value	P-value	
t-Test	1.150503	0.43	0.66	*Fail to reject the Ho that the mean is equal to 1.150503*
Simulated vs deterministic sorghum mean yield for 2019 in Dodoma				
	Given value	Test value	P-value	
t-Test	1.1676286	−0.04	0.97	*Fail to reject the Ho that the mean is equal to 1.167629*
Simulated vs deterministic rice mean yield for 2019 in Dodoma				
	Given value	Test value	P-value	
t-Test	1.1099606	0.15	0.88	*Fail to reject the Ho that the mean is equal to 1.109960*
Simulated vs deterministic maize mean price for 2019 in Dodoma				
	Given value	Test value	P-value	
t-Test	720,520	−0.36	0.72	*Fail to reject the Ho that the mean is equal to 720,520*
Simulated vs deterministic sorghum mean price 2019 in Dodoma				
	Given value	Test value	P-value	
t-Test	790,919	−0.62	0.53	*Fail to reject the Ho that the mean is equal to 790,919*
Simulated vs deterministic rice mean price for 2019 in Dodoma				
	Given value	Test value	P-value	
t-Test	1,969,380	−0.23	0.82	*Fail to reject the Ho that the mean is equal to 1,969,380*
Simulated vs deterministic maize mean yield for 2019 in Morogoro				
	Given value	Test value	P-value	
t-Test	0.931736	−0.04	0.97	*Fail to reject the Ho that the mean is equal to* 0.931736
Simulated vs deterministic sorghum mean yield for 2019 in Morogoro				
	Given value	Test value	P-value	
t-Test	0.994937	−0.07	0.94	*Fail to reject the Ho that the mean is equal to 0.994937*
Simulated vs deterministic rice mean yield for 2019 in Morogoro				
	Given value	Test value	P-value	
t-Test	2.246236	0.59	0.55	*Fail to reject the Ho that the mean is equal to 2.246236*
Simulated vs deterministic maize mean price for 2019 in Morogoro				
	Given value	Test value	P-value	
t-Test	659,637	−0.15	0.88	*Fail to reject the Ho that the mean is equal to 659,637*
Simulated vs deterministic sorghum mean price for 2019 in Morogoro				
	Given value	Test value	P-value	
t-Test	1,252,099	−0.39	0.70	*Fail to reject the Ho that the mean is equal to 1,252,099*
Simulated vs deterministic rice mean price for 2019 in Morogoro				
	Given value	Test value	p-value	
t-Test	1851,977	−0.23	0.94	*Fail to reject the Ho that the mean is equal to 1851,977*

Appendix D. The summary statistics for stochastic prices in the model

	2019	2020	2021	2022	2023	2024	2025
Maize price/t in Dodoma (× 1000)							
Mean	718.61	749.50	780.95	813.13	840.95	872.30	904.05
STDV	117.16	121.33	126.24	133.22	139.18	143.59	143.32
CV	16.30	16.19	16.17	16.38	16.55	16.46	15.85
Min	518.10	540.27	562.43	584.60	606.77	628.93	651.10
Max	937.66	977.78	1017.90	1058.02	1098.14	1138.26	1178.37
Maize price/t in Morogoro (× 1000)							
Mean	659.06	684.98	711.96	739.32	762.95	790.15	816.89
STDV	110.41	117.18	120.71	125.26	130.11	134.01	137.00
CV	16.75	17.11	16.95	16.94	17.05	16.96	16.77
Min	453.67	471.69	489.70	507.71	525.72	543.74	561.75
Max	824.52	857.26	889.99	922.73	955.47	988.21	1020.94
Sorghum price/t in Dodoma (× 1000)							
Mean	786.94	831.22	872.52	916.71	959.72	1001.69	1046.95
STDV	142.97	150.42	155.11	167.62	175.38	182.54	185.15
CV	18.17	18.10	17.78	18.28	18.27	18.22	17.69
Min	543.77	573.53	603.30	633.06	662.82	692.59	722.35
Max	1055.39	1113.15	1170.92	1228.68	1286.45	1344.22	1401.98
Sorghum price/t Morogoro (× 1000)							
Mean	1250.91	1325.32	1401.80	1478.07	1553.24	1629.43	1705.72
STDV	122.90	129.87	136.88	145.33	152.60	159.97	166.03
CV	9.82	9.80	9.76	9.83	9.82	9.82	9.73
Min	1076.22	1141.45	1206.68	1271.91	1337.14	1402.37	1467.60
Max	1461.13	1549.69	1638.24	1726.80	1815.36	1903.92	1992.48
Rice price/t in Dodoma (× 1000)							
Mean	1966.96	2040.32	2113.62	2186.94	2260.14	2333.55	2406.97
STDV	239.69	248.76	257.57	266.52	275.24	284.38	293.36
CV	12.19	12.19	12.19	12.19	12.18	12.19	12.19
Min	1667.00	1729.14	1791.29	1853.43	1915.58	1977.72	2039.86

Max	2606.56	2703.73	2800.90	2898.07	2995.24	3092.41	3189.58
Rice price/t Morogoro (× 1000)							
Mean	1851.30	1932.41	2013.40	2094.45	2175.37	2256.49	2337.61
STDV	192.76	201.30	209.65	218.11	226.37	234.96	243.43
CV	10.41	10.42	10.41	10.41	10.41	10.41	10.41
Min	1501.11	1566.82	1632.54	1698.25	1763.96	1829.68	1895.39
Max	2343.51	2446.10	2548.69	2651.29	2753.88	2856.47	2959.06

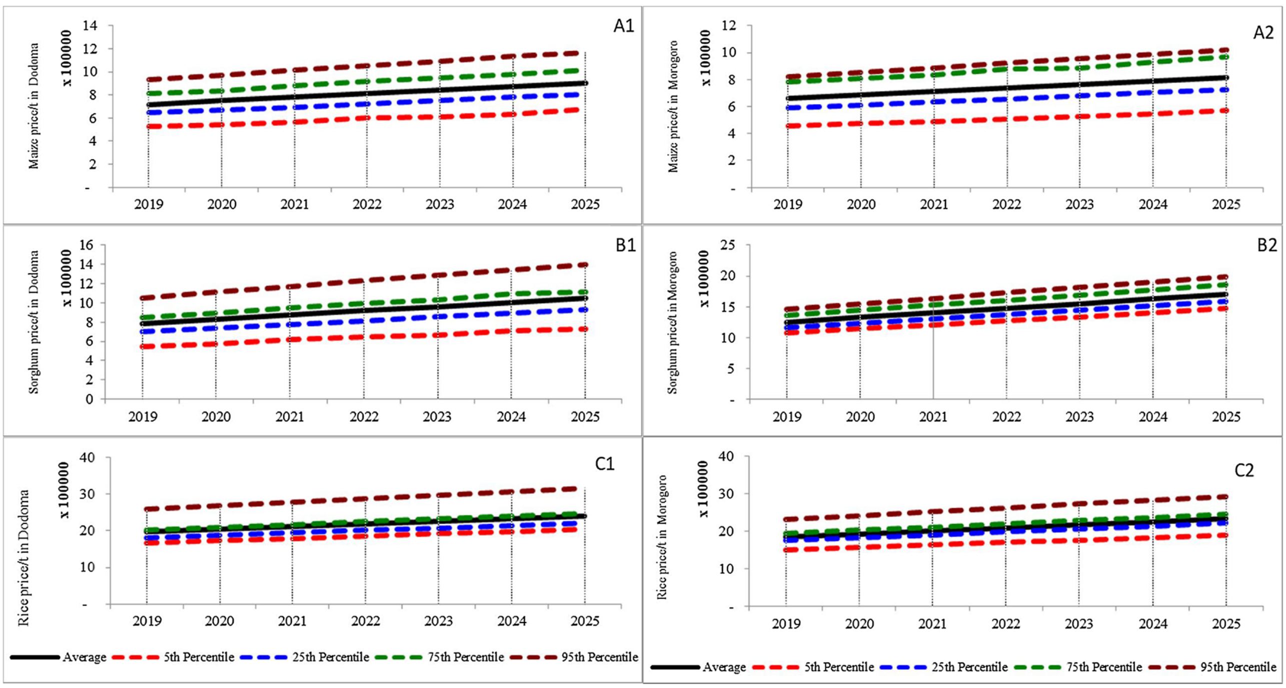

Fig. D1. Fan graph showing how the relative variability of c stochastic prices change over the 7-year period through 2025. Fig. A1, B1 and C1 for Dodoma region and A2, B2 and C2 for Morogoro region.

Appendix E

	2019	2020	2021	2022	2023	2024	2025
Annual net returns for maize in Dodoma (× 1000,000,000)							
Mean	121.33	137.08	153.48	166.43	173.86	184.05	191.94
STDV	67.58	69.13	68.56	72.92	66.89	69.80	66.93
CV	55.70	50.43	44.67	43.82	38.47	37.93	34.87
Min	−51.77	−40.43	−4.64	4.66	22.68	21.05	37.12
Max	344.50	348.40	373.96	412.12	397.00	422.33	440.70
P(< 0)	3.7%	1.5%	0.5%	–	–	–	–
Annual net returns for maize in Morogoro (× 1000,000,000)							
Mean	45.59	44.83	43.25	42.37	38.94	37.54	36.29
STDV	53.98	50.86	48.62	44.34	39.05	37.59	34.61
CV	118.42	113.45	112.42	104.65	100.27	100.13	95.35
Min	−75.38	−78.81	−77.02	−70.70	−59.30	−51.00	−47.17
Max	232.76	230.54	210.01	202.08	163.98	165.90	143.27
P(< 0)	21.4%	19.8%	17.6%	15.74%	18.2%	16.8%	13.9%
Annual net returns for sorghum in Dodoma (× 1000,000,000)							
Mean	81.54	85.53	88.09	90.50	90.62	91.86	92.70
STDV	36.54	38.21	36.92	37.37	33.51	35.56	34.49
CV	44.81	44.67	41.91	41.30	36.98	38.71	37.21
Min	2.70	5.93	8.54	14.30	20.28	18.84	23.71
Max	191.78	193.15	202.89	192.71	192.60	196.18	193.30
P(< 0)	–	–	–	–	–	–	–
Annual net returns for sorghum in Morogoro (× 1000,000,000)							
Mean	7.92	7.79	7.68	7.47	7.24	7.06	6.79
STDV	2.38	2.38	2.40	2.22	1.99	2.08	1.93
CV	30.00	30.57	31.29	29.65	27.53	29.43	28.39
Min	3.03	3.07	3.47	3.23	2.95	2.96	3.33
Max	14.72	14.92	14.19	13.79	12.67	12.34	12.34

P(< 0)	–	–	–	–	–	–	–
Annual net returns for rice in Dodoma (× 1000,000,000)							
Mean	5.45	6.24	6.89	7.46	7.79	8.27	8.46
STDV	4.52	4.58	4.30	4.46	4.02	4.16	4.00
CV	82.80	73.42	62.43	59.84	51.63	50.25	47.31
Min	−5.55	−4.67	−2.83	−2.98	−1.05	−0.41	0.65
Max	21.18	23.10	20.09	23.04	23.87	23.81	22.40
P(< 0)	9.8%	8.1%	4.0%	2.0%	1.0%	0.3%	–
Annual net returns for rice in Morogoro (× 1000,000,000)							
Mean	411.07	437.19	453.72	472.92	483.15	494.92	500.82
STDV	113.98	118.07	110.41	113.20	114.17	110.38	115.99
CV	27.73	27.01	24.33	23.94	23.63	22.30	23.16
Min	181.00	188.21	212.63	214.60	234.01	226.45	193.84
Max	730.01	818.89	800.69	837.34	853.59	832.61	1019.03
P(< 0)	–	–	–	–	–	–	–

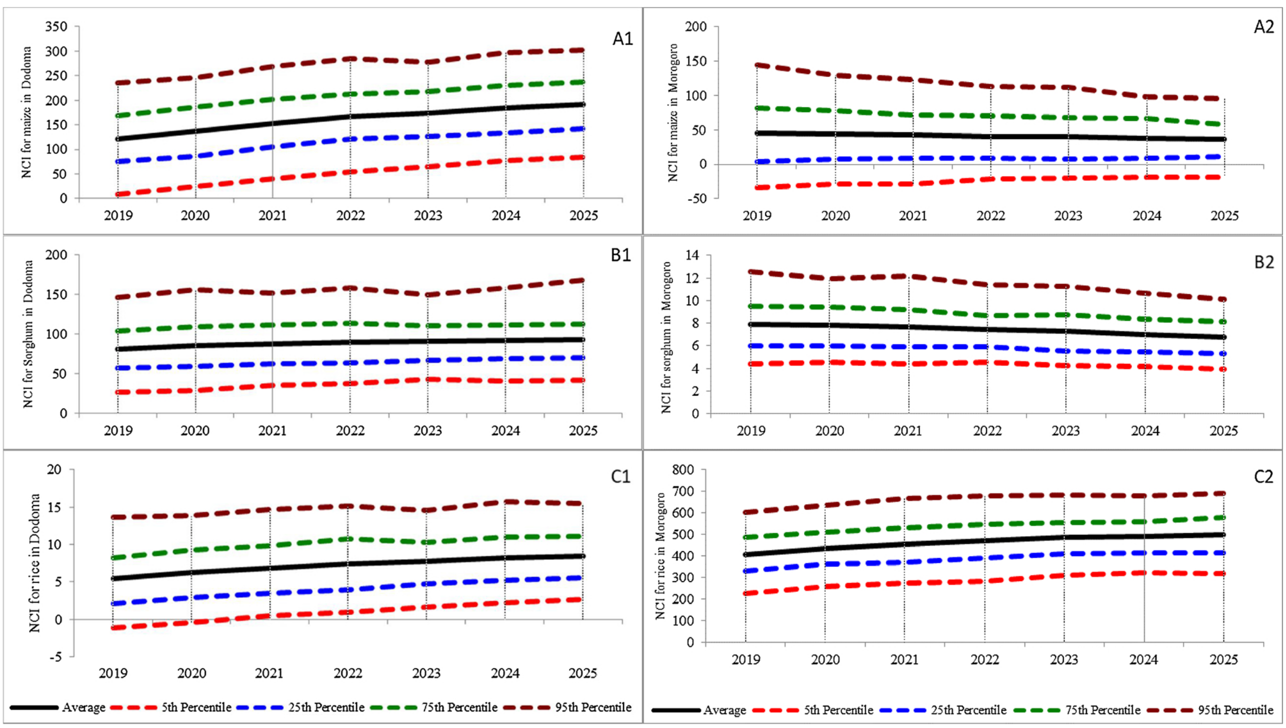

Fig. E1. Fan graph showing how the relative variability of stochastic annual net income (in TZS billion) change over the 7-year period through 2025: A1 = maize, B1 = sorghum and C1 = rice for Dodoma region and A2 = maize, B2 = sorghum and C2 = rice for Morogoro region.