Economic Modelling 45 (2015) 155–168

Contents lists available at ScienceDirect

Economic Modelling

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

Stochastic market modeling with Gaussian Quadratures: Do rotations of Stroud's octahedron matter?

Marco Artavia [a,b,*], Harald Grethe [c,1], Georg Zimmermann [d,2]

[a] Department of Agricultural Economics, Humboldt-Universität zu Berlin, Germany
[b] European Commission, Joint Research Centre, Institute for Prospective Technological Studies (IPTS), Spain
[c] Institute of Agricultural Policy and Markets (420), Universität Hohenheim, 70593 Stuttgart, Germany
[d] Institute of Applied Mathematics and Statistics (110), Universität Hohenheim, 70599 Stuttgart, Germany

ARTICLE INFO

Article history:
Accepted 14 October 2014
Available online 13 December 2014

Keywords:
Sampling methods
Gaussian Quadratures
Monte Carlo
Stochastic modeling
Commodity markets

ABSTRACT

Recently, stochastic applications of large-scale applied simulation models of agricultural markets have become more frequent. However, stochastic modeling with large market models comes with high computational and management costs for data storage, analysis and manipulation. Gaussian Quadratures (GQ), are efficient sampling methods requiring few points to approximate the central moments of the joint probability distribution of stochastic variables and therefore reduce computational costs. For symmetric regions of integration, the vertices of Stroud's *n*-octahedron (Stroud, 1957) are GQ formulas of degree 3 with a minimal number of points which can make the stochastic modeling with large economic models manageable. However, we have the conjecture that rotations of Stroud's *n*-octahedron have an effect on the accuracy of approximation of model results; thus, we test eight different rotations using the European Simulation Model (ESIM). It was found that the 45° rotation yields higher accuracy than the 0° rotation. With the 45° rotation and in models with large regions or variables which strongly determine the outcome of model results such as soft wheat in ESIM, the arrangement of the stochastic variables (A1 or A2) in the covariance matrix or the selected method to introduce correlation (via the Cholesky decomposition –C– or via the diagonalization method –D–) may have a significant effect on the accuracy of the quadratures. With the 45° rotation and with markets where the effect of the different regions or variables on model outcomes are more homogenous as in the case of rapeseed in ESIM, the selection of the arrangements A1 or A2 and the method of introducing correlation C or D may not have a significant effect on the accuracy of the quadratures.

Notation

It is considered useful to first introduce the notation used throughout the article. It can be divided in 4 groups: a) notation for spatial regions, b) matrix and vector notations, c) notation used to describe formulas in the European Simulation Model (ESIM), and d) other notation.

a). Notation for spatial regions

C^n	the *n*-cube with centroid at the origin and vertices $(\pm a, \pm a, \cdots, \pm a)$
E^n	Euclidean *n*-dimensional space

b). Matrix and vector notation

In general, matrices are denoted with upper case, bold letters; column vectors with lower case, bold letters; row vectors with lower case, bold, italic letters and a subscript indicating the row or variable; and vectors of row vectors with lower case, bold, italic letters but without a subscript.

The following are some specificities:

$\mathbf{I}_n$	the identity matrix of size $n \times n$
i	index to determine the elements (variables or coordinates) of a column vector: $i = 1, 2, \cdots, n$
k	index to determine the quadrature points in a matrix of quadratures — these are column vectors: $k = 1, 2, \cdots, N$
$\mathbf{L}$	lower triangular matrix of the Cholesky factorization, $\boldsymbol{\Sigma}[\mathbf{z}] = \mathbf{L}\mathbf{L}^{\mathrm{T}}$
N	number of quadrature points or required evaluations of an integrand

* Corresponding author at: Sonnenallee 105, 12045 Berlin, Germany. Tel.: +49 15773876226; fax: +49 30 20 60 972 99.
E-mail addresses: maartavi@gmail.com (M. Artavia), grethe@uni-hohenheim.de (H. Grethe), Georg.Zimmermann@uni-hohenheim.de (G. Zimmermann).
[1] Tel.: +49 711 459 22631.
[2] Tel.: +49 711 459 22147.

http://dx.doi.org/10.1016/j.econmod.2014.10.017

n number of variables in a column vector — the number of dimensions of a multivariate integration problem
R upper triangular matrix of the reverse Cholesky factorization, $\boldsymbol{\Sigma}[\mathbf{z}] = \mathbf{R}\mathbf{R}^{\mathrm{T}}$
$[\,]^{\mathrm{T}}$ transpose
U the orthogonal matrix consisting of the eigenvectors of $\boldsymbol{\Sigma}[\mathbf{z}]$ — these are situated in the columns of **U**
X matrix of quadrature points for the approximation of a multivariate normal probability density function: $\begin{bmatrix} x_{1,1} & \cdots & x_{1,N} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,N} \end{bmatrix}$
x an independent variable
$\mathbf{x}$ a vector of variables whose components are denoted by subscripts: $\mathbf{x} = (x_1, x_2, \cdots, x_n)$
Z matrix of deviates with n variables and m observations: $\begin{bmatrix} z_{1,1} & \cdots & z_{1,m} \\ \vdots & \ddots & \vdots \\ z_{n,1} & \cdots & z_{n,m} \end{bmatrix}$
Γ matrix of quadrature points for the multivariate standard normal probability density function or the C^n with vertices $(\pm 1, \pm 1, \cdots, \pm 1)$: $\begin{bmatrix} \gamma_{1,1} & \cdots & \gamma_{1,N} \\ \vdots & \ddots & \vdots \\ \gamma_{n,1} & \cdots & \gamma_{n,N} \end{bmatrix}$
$\boldsymbol{\mu}$ the vector of means of the stochastic variables in a multivariate probability density function: $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_n)$
$\boldsymbol{\Sigma}$ the covariance matrix: $\begin{bmatrix} \sigma_{1,1} & \cdots & \sigma_{1,n} \\ \vdots & \ddots & \vdots \\ \sigma_{n,1} & \cdots & \sigma_{n,n} \end{bmatrix}$ where $\sigma_{i,i} = \sigma_i^2 = \mathrm{Var}(x_i)$
0 the zero vector with n-elements: $\mathbf{0} = (0, 0, \cdots, 0)^{\mathrm{T}}$

c). notation used to describe formulas in ESIM

$elastsp_{crops,crops}$ own and cross price elasticity of supply
elastyd own price elasticity of yield
elastyi yield elasticity with respect to the index of intermediate input costs
elastyl yield elasticity with respect to the index of labor costs
indi cost index of intermediate inputs
indl cost index of labor inputs
intsp constant in the supply function
tp technical progress of yields or supply
intyd constant in the yield function

d). Other notation

d degree of precision of a quadrature formula
E[ ] expected value of the data in the content of the brackets
Q^n the n-octahedron from Stroud's theorem from 1957

1. Introduction

Recently, stochastic applications of large-scale applied simulation models of agricultural markets have become more frequent. These analyses have different purposes. First, some intend to identify the main sources of market uncertainty and to give an indication of alternative possible outcomes around baseline commodity market projections (FAPRI-MU, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011b, 2012, 2013; European Commission, 2011, 2012; OECD/FAO, 2007, 2008, 2011, 2013; OECD, 2011). Second, some aim at calibration of model parameters and model validation (Beckman et al., 2011; Hertel et al., 2005; Valenzuela et al., 2007). Third, some aim at policy analysis (Arndt and Hertel, 1997; Hertel et al., 2010; Verma, Hertel and Valenzuela, 2011; Westhoff, Brown and Hart, 2006). With the diversity of possible applications as well as the current interest in understanding and assessing uncertainty, the increase in stochastic modeling applications is likely to continue.

One important aspect of stochastic modeling with large economic models is how to handle dimensionality. Stochastic simulations come with high computational and management costs for data storage, analysis and manipulation. The Food and Agricultural Policy Research Institute (FAPRI) model, AGLINK-COSIMO,[3] the European Simulation Model (ESIM) and the Global Trade Analysis Project (GTAP) model are frequently applied large market models that have been modified to incorporate stochastic features, although they handle dimensionality in different ways (see FAPRI-MU, 2011a; Burrell and Nii-Naate, 2013; Artavia, 2014; Pearson and Arndt, 2000 respectively, for a description of the stochastic features in those models). The FAPRI model uses Latin Hypercube Sampling (LHS) in combination with a reduced model version with less detail than the deterministic version; AGLINK-COSIMO runs over 500 LHS points with high computational and simulation time requirements; and ESIM and the GTAP model make use of Gaussian Quadratures (GQ), which are efficient sampling methods requiring few points to approximate the central moments of the joint probability distribution of the stochastic variables.

Stroud (1957) proposed a theorem for the generation of GQ formulas which have the advantage of being very simple to compute and requiring the minimal number of points (Mysovskih, 1966 as cited by Haber, 1970). In his theorem, Stroud has shown that for symmetric regions of integration one can use the vertices of an n-dimensional octahedron, Q^n, with certain characteristics as a quadrature formula. In stochastic modeling, a symmetric region of integration is for example the multivariate normal distribution of the stochastic variables. With Stroud's theorem we are able to generate a Q^n for the multivariate standard normal distribution with independent stochastic variables and we call it the 'reference Q^n'. However, since we work with uncertainties around commodity markets we want the quadratures to consider the observed variance and covariance of the stochastic variables. In order to induce a desired covariance matrix a linear transformation must be applied to the matrix of quadratures obtained from the reference Q^n. We call the n-octahedron considering specific covariance matrices the 'transformed Q^n'.

The transformed Q^n can be rotated in space without changing the degree of precision of the quadrature formula. Nonetheless, in stochastic modeling, the interest is not only in the degree of precision of the quadrature, but also in the location of the sample points. With the rotation of the transformed Q^n, the location of points sampling the marginal distributions of the stochastic variables changes significantly and this may have an effect on the accuracy of the quadratures. For example, functions in large market models are often highly dimensional, depending on complex cross relationships or market policies triggered at threshold points (i.e., tariffs, tariff rate quotas — TRQs, export subsidies with quantity limits, production quotas, etc.). This can make the evaluation point of the function an issue that matters.

Stroud-based sampling approaches are convenient as they may result in a considerable reduction of the number of model solves needed in stochastic analyses. However, the level of accuracy of approximation of model results of different rotations of both, the reference and the transformed Q^n, has not been sufficiently studied in market models. Furthermore, the accuracy of different rotations has not been tested in large partial equilibrium (PE) models of the agricultural sector as for example

[3] The model is used by The Organisation for Economic Co-operation and Development (OECD) and the Food and Agriculture Organization (FAO) of the United Nations for their annual agricultural outlook. The model is also used by the European Commission (EC).

AGLINK-COSIMO, the FAPRI model or ESIM. Two articles deal with the topic of accuracy of GQ in computable general equilibrium (CGE) models. Preckel et al. (2011b) assess the performance of two linear transformations of the reference Q^n, namely, the ones computed via the Cholesky decomposition and via the Eigensystem of the covariance matrix (the procedure making use of the Eigensystem of the covariance matrix is called the diagonalization method in this article (see Section 3.3). The reference Q^n evaluated is the one generated by using Stroud's degree 3 formula (see Stroud, 1957) which presents 45° of rotation (see Section 3.1). To avoid biases created by different models or different correlation levels, Preckel et al. (2011b) use three models and different covariance matrices. Their results show that for the two smaller models, the choice of the transformation method does not have a large impact on the accuracy of the quadrature. For the third model, a highly aggregated version of a global general equilibrium model (GTAP), results are inconclusive. Preckel et al. (2011a) explore the sensitivity of model results to the breadth of the GQ sampling using a reduced version of the model applied by Hertel, Martin and Leister (2010) which is a GTAP model version that has been specifically tailored to agricultural applications. They make two copies of the GQ derived by either Stroud or Liu which are available in the Systematic Sensitivity Analysis (SSA) in GEMPACK. They stretch one of the copies and shrink the other so as to make it possible to maintain the variances (Preckel et al., 2011b). Thus, their method results in a duplication of the sample size. They find that the sampling breadth is an important aspect for highly non-linear models.

In this paper we have the hypothesis that different rotations of the transformed Q^n may have an effect on the accuracy of approximation of the central moments of the output variables of commodity market simulation models. Thus, the article has the purpose of assessing the performance of alternative rotations of the transformed Q^n. For that purpose different quadrature formulas are tested using ESIM. Like Preckel et al. (2011b), we evaluate the accuracy of quadratures obtained with the linear transformations of the reference Q^n with 45° of rotation computed via the Cholesky decomposition and via the diagonalization method. Moreover, we also tests two alternative procedures in the generation of the transformed Q^n which have not been studied before: (i) to use the reference Q^n with the vertices lying on the coordinate axes (0° rotation), and (ii) to use different arrangements of the coordinate system (different arrangements of the stochastic variables in the covariance matrix) when computing the linear transformation of the 45° and 0° reference Q^ns. Note that the different procedures all result in different rotations of the transformed Q^n.

The article is organized as follows: Section 2 gives an overview of ESIM and of the stochastic version used; Section 3 provides the theoretical background of Stroud's theorem and its application to stochastic modeling; Section 4 presents the rotations evaluated; Section 5 introduces the benchmark (the approximation of the true value of model results) for the assessment of the accuracy of the rotations tested; Section 6 presents the results and identifies several factors which contribute to the explanation of the observed differences in accuracy; the final section offers conclusions and briefly describes the future research agenda.

2. Stochastic version of ESIM

ESIM is a world, comparative static, partial equilibrium, net trade, multimarket model which covers the main agricultural products. It has been designed with a focus on the simulation of medium-term (five to ten years) developments of agricultural markets in the European Union (EU) and accession candidate countries (Croatia, Turkey, and Western Balkan). Thus, it depicts the EU at the Member State (MS) level and includes the market policies of the Common Agricultural Policy (CAP) in detail. The 'Rest of the World' (ROW) is modeled as one aggregate with the exception of the United States of America (USA) which is modeled separately. ESIM has rich cross and input/output relationships between commodities and covers first-stage processing products (dairy, plant oils and biofuels). The model is run with a set of coherent macroeconomic and policy assumptions and solves for the equilibrium where world net exports equal world net imports to determine equilibrium world market prices. For this article, an ESIM version documented in Banse, Grethe and Nolte (2005) is used and policy specifications as well as stochastic extensions are as documented in Artavia (2014).

In the stochastic version the yields of wheat, barley and rapeseed in all countries/regions are considered as stochastic, accounting for their uncertainty around linear trends. The stochastic terms are incorporated into the respective supply and yield equations. The crops are selected due to their importance in the EU. All 3 crops are significant in terms of their share of production in the EU and wheat and rapeseed especially significant in terms of their political importance, wheat as a sensitive food product and rapeseed as the main feedstock for the production of biodiesel.

Due to its focus on the EU, in ESIM the equations for crop supply for the EU Member States are divided in area and yield equations, supply being then obtained by:

$$\mathrm{SUPPLY} = \mathrm{AREA} \times \mathrm{YIELD}. \tag{1}$$

Crop supply for the rest of countries and regions – USA and ROW – is only one isoelastic function:

$$\mathrm{SUPPLY} = \mathrm{intsp} \prod\nolimits_{\mathrm{crops}} PP^{\mathrm{elastsp}_{\mathrm{crops,crops}}}\,\mathrm{tp}. \tag{2}$$

For the EU Member States crop yields are modeled with isoelastic functions dependent on producer prices, on intermediate cost indices and on technical progress as follows:

$$\mathrm{YIELD} = \mathrm{intyd}\, PP^{\mathrm{elastyd}}\, \mathrm{indi}^{\mathrm{elastyi}}\, \mathrm{indl}^{\mathrm{elastyl}}\, \mathrm{tp}. \tag{3}$$

Note that the yield equations are only subject to own price elasticities. The cross price elasticities are considered in the area equations which are not presented here.

For the stochastic version, the stochastic term *stoch* has been included for the corresponding crops in all regions as a factor into Eqs. (4) and (5) as follows.

For the USA and the ROW:

$$\mathrm{SUPPLY} = \mathrm{intsp} \prod\nolimits_{\mathrm{crops}} PP^{\mathrm{elastsp}_{\mathrm{crops,crops}}}\,\mathrm{tp}(1 + \mathrm{stoch}), \tag{4}$$

for the European countries:

$$\mathrm{YIELD} = \mathrm{intyd}\, PP^{\mathrm{elastyd}}\, \mathrm{indi}^{\mathrm{elastyi}}\, \mathrm{indl}^{\mathrm{elastyl}}\, \mathrm{tp}(1 + \mathrm{stoch}). \tag{5}$$

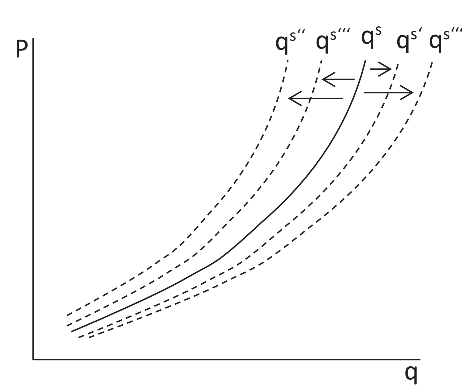


Fig. 1. Effect of the stochastic variable on supply curves.

stoch contains the generated quadratures for $\mathbf{x} \sim N(\mathbf{0}, \mathbf{\Sigma})$ which is the assumed stochastic space of the random vector.

By running the model several times over the quadrature points contained in *stoch*, the consequences of yield uncertainty on the stability of markets can be studied. Fig. 1 shows the effect of the stochastic factor on the crop supply curves. It can be observed how the factor shifts the curve to the left and to the right.

For the determination of the stochastic terms, the EU MS are grouped both to reduce the dimensionality of stochastic space and to focus on the major sources of price uncertainty. Groups are formed according to the level of correlation of yield deviates between MS as well as to their share of production in the EU. MS with high correlation are grouped together, e.g., Romania and Bulgaria; MS with small shares of production are either put together with a larger neighbor, e.g., Ireland is grouped with the UK, or are left in their deterministic version, e.g., the Baltic States, Malta, and Cyprus. The grouping results in a total of 42 stochastic variables: 16 for wheat, 17 for barley, and 9 for rapeseed (see Table 1).

After the grouping, the stochastic part of the yield time series is determined as the deviates from estimated linear trends, and the deviates are captured as shares based on FAOSTAT data (1961–2006). For example, let $y_{i,j}$ be any of the observed values for the variable i, where $i = (1, 2, \cdots, n)$, in the year j, where $j = (1, 2, \cdots, m)$; and let $\hat{y}_{i,j}$ be the estimated trend value for the same variable in the same year. Then the observed deviate $z_{i,j}$ is captured by

$$z_{i,j} = \frac{y_{i,j}}{\hat{y}_{i,j}} - 1. \tag{6}$$

Proceeding in this way for all variables and all years, the matrix of deviates $\mathbf{Z}_{n \times m}$ is computed. The final matrix of deviates is arranged as follows:

$$\mathbf{Z}_{n\times m} = \begin{bmatrix} z_{FR.w,1} & z_{FR.w,2} & \cdots & z_{FR.w,m} \\ z_{GE.w,1} & z_{GE.w,2} & \cdots & z_{GE.w,m} \\ \vdots & \vdots & & \vdots \\ z_{ROW.w,1} & z_{ROW.w,2} & \cdots & z_{ROW.w,m} \\ z_{FR.b,1} & z_{FR.b,2} & \cdots & z_{FR.b,m} \\ z_{GE.b,1} & z_{GE.b,2} & \cdots & z_{GE.b,m} \\ \vdots & \vdots & & \vdots \\ z_{ROW.b,1} & z_{ROW.b,2} & \cdots & z_{ROW.b,m} \\ z_{FR.r,1} & z_{FR.r,2} & \cdots & z_{FR.r,m} \\ z_{GE.r,1} & z_{GE.r,2} & \cdots & z_{GE.r,m} \\ \vdots & \vdots & & \vdots \\ z_{ROW.r,1} & z_{ROW.r,2} & \cdots & z_{ROW.r,m} \end{bmatrix}, \tag{7}$$

for $n = 42$ variables and $m = 46$ years.

In the elements of $\mathbf{Z}$, the first index indicates the country and the crop at the same time and the second index determines the observation year. For example, *FR. w*, 1 is the value for wheat in France in year one (1961). Thus, the variables of wheat in Eq. (7) are located in the upper block of the matrix, those of barley in the middle block (indices with: *country*. *b*) and rapeseed in the bottom block (indices with: *country*. *r*).

Since the elements of $\mathbf{Z}$ are deviations from linear trends, the expected value of each stochastic variable is zero. The standard deviations of the stochastic variables differ to an important degree from each other and are presented in Table 1. Statistical tests showed that most variables are stationary and normally distributed. Thus, for simplicity and since the focus of this article is on the evaluation of different sampling techniques, we assume stationarity and normal distribution for all stochastic variables.

In its deterministic version, ESIM has been developed to simulate medium-term adaptation of agricultural markets to external shocks. In the stochastic version, the model is run to simulate the effects of weather induced short-term yield and supply shocks. For this reason, the parameterization of behavioral yield and supply functions in ESIM is adjusted to simulate the short-run adaptation of markets. Own, input and cross price elasticities of crops and livestock commodities are reduced such that ESIM reproduces observed world market price volatility. Deviates from linear price trends calculated from data by Anderson and Valenzuela (2008) are taken as the basis for the estimation of world market price uncertainty. For the adaptation of elasticities, a new parameter per commodity, *ea* (elasticity adjustment parameter), is created and both input and output price elasticities are adjusted by it. The final values of the elasticity adjustment parameter are presented in Table 2.

In the deterministic version, own price and input price elasticities are relatively similar across countries while cross price elasticities diverge to a higher degree. This is a result of the calibration of the system of elasticities to fulfill the conditions derived from economic theory, which are homogeneity of degree zero in input and output prices, symmetry of compensated substitution effects and non-negativity of the own price effect (Banse et al., 2007). In the calibration process, own and input price elasticities are held constant and only cross price elasticities are allowed to vary (Banse et al., 2007).

Table 1
Stochastic variables integrated to the yield and supply equations of soft wheat, barley and rapeseed in different countries and country-groups and their standard deviation (in %: 0.10 = 10%).

	Wheat		Barley		Rapeseed	
EU-15	AT	0.10	AT	0.10	AT	0.12
	BE–NL	0.09	BE–NL	0.09		
	DK–SW–FI	0.09	DK–SW	0.09	DK–SW–FI	0.10
			FI	0.14		
	ES–PT	0.16	ES–PT	0.19		
	FR	0.08	FR	0.08	FR	0.12
	GE	0.07	GE	0.07	GE	0.10
	GR	0.17	GR	0.15		
	IT	0.08	IT	0.13		
	UK–IR	0.09	UK–IR	0.06	UK–IR	0.14
EU-12	CZ–SK	0.15	CZ–SK	0.17	CZ–SK	0.16
	HU	0.23	HU	0.20	HU	0.17
	PL	0.11	PL	0.13	PL	0.17
	RO–BG	0.21	RO–BG	0.21		
Other regions	TU	0.11	TU	0.10		
	US	0.07	US	0.08		
	ROW	0.06	ROW	0.09	ROW	0.07

Note: The countries and regions are abbreviated in the following way: Austria (AT), Belgium–Luxembourg (BE), BG: Bulgaria (BG), Cyprus (CY), Czech Republic (CZ), Denmark (DK), Finland (FI), France (FR), Germany (GE), Greece (GR), Hungary (HU), Ireland (IE), Italy (IT), Netherlands (NL), Poland (PL), Portugal (PT), Romania (RO), 'Rest of the World' (ROW), Slovak Republic (SK), Spain (ES), Sweden (SW), Turkey (TU), United Kingdom (UK), United States of America (USA).

3. Theoretical background

The following 3 subsections present the theoretical background of Stroud's theorem of numerical integration, its correspondence with discrete approximations of probability distributions, and the transformation of quadratures for the joint standard normal space to quadratures for specific stochastic spaces.

Table 2
Final values of the elasticity adjustment parameter (*ea*) used as a factor to reduce the input and output price elasticities of yield and supply equations.

$ea_{for\ all\ commodities}$	0.300
ea_{wheat}	0.150
ea_{barley}	0.200
$ea_{rapeseed}$	0.155

3.1. Stroud's theorem

Stroud's theorem from 1957 states:

> *A necessary and sufficient condition that 2n points $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n, -\boldsymbol{x}_1, \cdots, -\boldsymbol{x}_n$ form an equally weighted numerical integration formula of degree 3 for a symmetrical region is that these points form the vertices of a Q^n whose centroid coincides with the centroid of the region and lie on an n-sphere of radius $r = \sqrt{nI_2/I_0}$.*
>
> [Stroud (1957, p. 259).]

Here, I_0 is the n-volume of the region of integration and I_2 is the integral of the square of any variable over that region.

Stroud (1957), based on his theorem and with the purpose of generating quadrature points which are interior to the integration region, proposed the following formula for the n-cube (C^n) with vertices $(\pm 1, \pm 1, \cdots, \pm 1)$.

Let γ_k denote the quadrature point $(\gamma_1, \gamma_2, \cdots, \gamma_n)^{\mathrm{T}}$, where:

$$\gamma_{k,\,2r-1} = \sqrt{\frac{2}{3}}\cos\frac{(2r-1)k\pi}{n} \quad \gamma_{k,\,2r} = \sqrt{\frac{2}{3}}\sin\frac{(2r-1)k\pi}{n} \quad \text{for } r = 1, 2, \cdots, \left[\frac{1}{2}n\right] \tag{8}$$

and if n is odd:

$$\gamma_{k,n} = (-1)^k/\sqrt{3} \tag{9}$$

then, the points $\gamma_1, \cdots, \gamma_N$ for $N = 2n$, satisfy the conditions of the theorem and all are interior to C^n (Stroud, 1957, p. 260).

Note that Stroud was confronted with the problem that for the C^n with vertices $(\pm a, \pm a, \cdots, \pm a)$ and for $n > 3$, the radius of the n-sphere on which the vertices that the Q^n must lie on is greater than a. Put another way, for $n > 3$ we have $r > a$. As a result, if the quadrature points lie on the coordinate axes, these would be outside C^n as the calculation below shows.

For $C^n = [-a, a]^n$ we obtain as its n-volume:

$$I_0 = \int_{C^n} x^0 dx = (2a)^n, \tag{10}$$

and the integral of the square of any variable over C^n is:

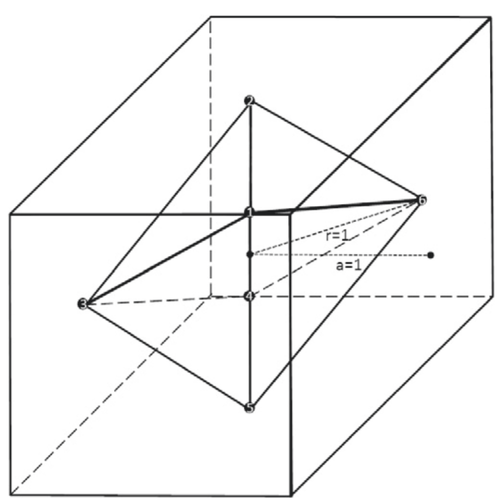
$$\begin{aligned} I_2 = \int_{C^n} x^2 dx &= \int_{C^{n-1}} dx \int_{-a}^{a} x^2 dx = (2a)^{n-1} \int_{-a}^{a} x^2 dx \\ &= (2a)^{n-1} \left[\frac{1}{3}x^3\right]_{-a}^{a} \\ &= (2a)^{n-1}\left(\frac{2}{3}a^3\right) = \frac{2^n}{3}a^{n+2}. \end{aligned} \tag{11}$$

Consequently, for the radius of the Q^n for the C^n we obtain:

$$r = \sqrt{n\frac{I_2}{I_0}} = \sqrt{n\left(\frac{2^n}{3}a^{n+2}/2^n a^n\right)} = \sqrt{n\frac{a^2}{3}} = a\sqrt{\frac{n}{3}}. \tag{12}$$

Thus, for $n > 3$, we have $r > a$.

Quadrature points outside the region of integration have the problem that the function being integrated may not be defined at those points. To avoid points lying outside C^n, for $n > 2$, Stroud rotates the Q^n so that the vertices are as far as possible from the coordinate axes. Fig. 2 shows the quadrature points obtained with Stroud's formula for the C^n with vertices $(\pm 1, \pm 1, \cdots, \pm 1)$ for $n = 3$. Note that the quadrature points are enumerated – where vertices 1–4, 2–5, and 3–6 are opposite to each other – and that the point in the middle indicates the center of the region. It can be observed that the points are the vertices of a Q^n whose centroid coincides with the centroid of the region C^n and lie on an n-sphere of radius $r = a\sqrt{n/3} = 1$. It can also be seen how the Q^n is rotated. Compared to a 3-octahedron with vertices on the coordinate axes, the vertices of the octahedron in Fig. 2 are rotated



Coordinates of the quadrature points:

	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6
γ_1	0.41	-0.41	-0.82	-0.41	0.41	0.82
γ_2	0.71	0.71	0.00	-0.71	-0.71	0.00
γ_3	-0.58	0.58	-0.58	0.58	-0.58	0.58

Fig. 2. The octahedron for the 3 dimensional cube obtained with Stroud's degree 3 formula from 1957.

by 45° with respect to two axes (such as x and y) and then again by 45° with respect to two other axes (such as x and z). It is difficult to imagine how with growing dimensions, the growing Q^n can still be fitted into the C^n. However, note that in Eq. (8), the coordinates for all i and all k take the minimum and maximum values of $\pm\sqrt{2/3}$, which occur when the cosine or the sine part takes a value of ± 1. This shows that these points are inside C^n for all n.

3.2. The reference Q^n

Numerical integration formulas for joint distributions have the characteristic that the regions of integration have an associated probability distribution. Thus, their degree of precision is defined by the capacity of matching the central moments of the joint probability distribution. For example, a formula of degree 3 reproduces all central moments up to the 3rd one.

The main difference between Stroud's Q^n for the C^n with vertices $(\pm 1, \pm 1, \cdots, \pm 1)$ and the reference Q^n (for the joint standard normal distribution) is its size. This is shown by the derivations below.

For the joint standard normal distribution:

$$I_0 = \int_{E^n} x^0 \frac{1}{(2\pi)^{n/2}} e^{-\frac{\|\mathbf{x}\|^2}{2}} d\mathbf{x} = (1)^n = 1 \tag{13}$$

and:

$$\begin{aligned} I_2 &= \int_{E^n} x_i^2 \frac{1}{(2\pi)^{n/2}} e^{-\frac{\|\mathbf{x}\|^2}{2}} d\mathbf{x} \\ &= \int_{E^{n-1}} \frac{1}{(2\pi)^{(n-1)/2}} e^{-\frac{\|\mathbf{x}\|^2 - x_i^2}{2}} d\mathbf{x} \int_{-\infty}^{\infty} x_i^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}} dx \\ &= (1)^{n-1} 1 = 1 \end{aligned} \tag{14}$$

where $\|\mathbf{x}\|$ is the norm of the vector $\mathbf{x}$.

Thus:

$$r = \sqrt{n \frac{I_2}{I_0}} = \sqrt{n}. \tag{15}$$

Note that the joint standard normal distribution is supported on the entire Euclidian space, E^n (see Eqs. (14) and (15)). In this way the vertices of the reference Q^n will never lie outside the region of integration. However, the marginal distributions could be limited to consider only the intervals of high probability, for example $[-3\sigma, 3\sigma]$ which would cover around 99.7% of their probabilities. As a result, the region of integration would be an n-sphere and the vertices of the reference Q^n could be outside of it.

Note also that the weights of Stroud's theorem, $w_k = \frac{I_0}{2n}$, applied to multivariate probability distributions result in: $p_k = \frac{1}{2n} I_0 = \frac{1}{2n}$.

3.3. The transformed Q^n

Specific stochastic spaces (the stochastic space expanded by the stochastic variables used in an uncertainty analysis with specific variance and covariance) divert from the definition of symmetric regions used by Stroud (1957). These are still symmetric regions but the Q^n can no longer lie on an n-sphere. Different variances result in the expansion or contraction of marginal probability distributions. Moreover, positive and negative correlation between stochastic variables results in the partial connection of the marginal distributions. If we limit the probability distribution of the stochastic variables to the intervals of high probability, as mentioned in Section 3.2, the region of integration would be an n-ellipsoid. In this subsection we show how to transform the reference Q^n to obtain quadratures which consider specific covariance matrices. In other words, if $\mathbf{z}$ is the vector of stochastic variables, it is shown how to induce a desired covariance matrix, $\mathbf{\Sigma}[\mathbf{z}]$, to the reference Q^n, in order to get a new matrix of quadratures, $\mathbf{X}$, with $\mathbf{\Sigma}[\boldsymbol{x}] = \mathbf{\Sigma}[\mathbf{z}]$. Further details of this procedure and a small example are given in Artavia et al. (2009).

To begin with, we consider an equidistribution of N arbitrary points $\mathbf{x}_1, \cdots, \mathbf{x}_N \in E^n$ with weight $1/N$ and mean $\mathrm{E}[\boldsymbol{x}] = (1/N)(\mathbf{x}_1 + \cdots + \mathbf{x}_N) = \mathbf{0}$. In this case the covariance matrix can be determined by simply gathering these points in an $n \times N$-matrix:

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,N} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,N} \end{bmatrix} \tag{16}$$

and then computing:

$$\mathbf{\Sigma}[\boldsymbol{x}] = \frac{1}{N}\mathbf{X}\mathbf{X}^{\mathrm{T}}. \tag{17}$$

For the vertices $\boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_N$ of the reference Q^n gathered in the matrix Γ:

$$\mathbf{\Sigma}[\boldsymbol{\gamma}] = \frac{1}{N}\mathbf{\Gamma}\mathbf{\Gamma}^{\mathrm{T}} = \mathbf{I}_{\mathrm{n}}, \tag{18}$$

where $\mathbf{I}_{\mathrm{n}}$ is the identity matrix of size $n \times n$.

Now, let $\mathbf{A}$ be any regular $n \times n$-matrix and consider the points $\mathbf{x}_k = \mathbf{A}\boldsymbol{\gamma}_k$ for $k = 1, 2, \cdots, N$. This yields:

$$\mathbf{X} = [\mathbf{A}\boldsymbol{\gamma}_1|\cdots|\mathbf{A}\boldsymbol{\gamma}_N] = \mathbf{A}\mathbf{\Gamma} \tag{19}$$

with:

$$\mathrm{E}[\boldsymbol{x}] = \frac{1}{2n}\sum_{k=1}^{2n}\mathbf{x}_k = \frac{1}{2n}\mathbf{A}\sum_{k=1}^{2n}\boldsymbol{\gamma}_k = \mathbf{A}\,\mathrm{E}[\boldsymbol{\gamma}] = \mathbf{A}\mathbf{0} = \mathbf{0} \tag{20}$$

and:

$$\mathbf{\Sigma}[\boldsymbol{x}] = \frac{1}{2n}\mathbf{X}\mathbf{X}^{\mathrm{T}} = \frac{1}{2n}\mathbf{A}\mathbf{\Gamma}(\mathbf{A}\mathbf{\Gamma})^{\mathrm{T}} = \mathbf{A}\frac{1}{2n}\mathbf{\Gamma}\mathbf{\Gamma}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}} = \mathbf{A}\mathbf{I}_{\mathrm{n}}\mathbf{A}^{\mathrm{T}} = \mathbf{A}\mathbf{A}^{\mathrm{T}}. \tag{21}$$

Thus, our problem is reduced to expressing the desired covariance matrix, $\mathbf{\Sigma}[\mathbf{z}]$, in the form $\mathbf{A}\mathbf{A}^{\mathrm{T}}$ for a regular square matrix $\mathbf{A}$. There are countless possibilities of doing this. Three standard methods are described below.

a) Diagonalization method (principal axes transformation):
Since $\mathbf{\Sigma}[\mathbf{z}]$ is positive semidefinite, it can be written in the form $\mathbf{\Sigma}[\mathbf{z}] = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{T}}$, where $\mathbf{D}$ is the (non-negative) diagonal matrix of eigenvalues of $\mathbf{\Sigma}[\mathbf{z}]$ and $\mathbf{U}$ is orthogonal (consisting of the eigenvectors of $\mathbf{\Sigma}[\mathbf{z}]$). Notice that the vectors (eigenvectors) will be the columns of $\mathbf{U}$. Then, letting $\mathbf{A} = \mathbf{U}\sqrt{\mathbf{D}}$ yields $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{\Sigma}[\mathbf{z}]$ as desired.

b) Cholesky decomposition:
The positive semidefinite matrix $\mathbf{\Sigma}[\mathbf{z}]$ has a Cholesky decomposition $\mathbf{\Sigma}[\mathbf{z}] = \mathbf{L}\mathbf{L}^{\mathrm{T}}$ where $\mathbf{L}$ is a lower triangular matrix as follows:

$$\mathbf{L} = \begin{bmatrix} L_{1,1} & 0 & \cdots & 0 \\ L_{2,1} & L_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ L_{n,1} & \cdots & L_{n,n-1} & L_{n,n} \end{bmatrix}. \tag{22}$$

Then, choose $\mathbf{A} = \mathbf{L}$.
The Cholesky decomposition can also be of the form $\mathbf{\Sigma}[\mathbf{z}] = \mathbf{L}\mathbf{D}\mathbf{L}^{\mathrm{T}}$, where $\mathbf{L}$ is a lower triangular matrix and $\mathbf{D}$ is a diagonal matrix as

$$
\mathbf{Z}_{A2} = \begin{bmatrix}
z_{ROW.w,1} & z_{ROW.w,2} & \cdots & z_{ROW.w,m} \\
z_{ROW.b,1} & z_{ROW.b,2} & \cdots & z_{ROW.b,m} \\
z_{ROW.r,1} & z_{ROW.r,2} & \cdots & z_{ROW.r,m} \\
z_{FR.w,1} & z_{FR.w,2} & \cdots & z_{FR.w,m} \\
z_{GE.w,1} & z_{GE.w,2} & \cdots & z_{GE.w,m} \\
\vdots & \vdots & & \vdots \\
z_{US.w,1} & z_{US.w,2} & \cdots & z_{US.w,m} \\
z_{FR.b,1} & z_{FR.b,2} & \cdots & z_{FR.b,m} \\
z_{GE.b,1} & z_{GE.b,2} & \cdots & z_{GE.b,m} \\
\vdots & \vdots & & \vdots \\
z_{US.b,1} & z_{US.b,2} & \cdots & z_{US.b,m} \\
z_{FR.r,1} & z_{FR.r,2} & \cdots & z_{FR.r,m} \\
z_{GE.r,1} & z_{GE.r,2} & \cdots & z_{GE.r,m} \\
\vdots & \vdots & & \vdots \\
z_{US.r,1} & z_{US.r,2} & \cdots & z_{US.r,m}
\end{bmatrix}
\begin{matrix}
\left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} \text{Wheat, Barley and Rapeseed in ROW} \\
\left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} \text{Wheat} \\
\left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} \text{Barley} \\
\left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} \text{Rapeseed}
\end{matrix}
$$

Note: The elements (z) of the matrix of stochastic variables data are identified with the following subscripts for the years $1, 2, \cdots, m$: (FR.w, FR.b, FR.r) denote wheat, barley, and rapeseed in France; (GE.w, GE.b, GE.r) denote wheat, barley, and rapeseed in Germany; (ROW.w, ROW.b, ROW.r) denote wheat, barley, and rapeseed in the 'Rest of the World'; and (US.w, US.b, US.r) wheat, barley, and rapeseed in the USA.

Fig. 3. The arrangement A2 of the stochastic variables in the matrix of data Z.

follows:

$$
\mathbf{LDL}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ L_{2,1} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ L_{n,1} & \cdots & L_{n,n-1} & 1 \end{bmatrix} \begin{bmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_n \end{bmatrix} \times \begin{bmatrix} 1 & L_{1,2} & \cdots & L_{1,n} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & L_{n-1,n} \\ 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{23}
$$

Then, choose $\mathbf{A} = \mathbf{L}\sqrt{\mathbf{D}}$.

c) Reverse Cholesky decomposition:
The positive semidefinite matrix $\boldsymbol{\Sigma}[\mathbf{z}]$ can also be decomposed in $\boldsymbol{\Sigma}[\mathbf{z}] = \mathbf{RR}^{\mathrm{T}}$ where $\mathbf{R}$ is an upper triangular matrix. In this case, choose $\mathbf{A} = \mathbf{R}$ to obtain the desired conditions.
It is worth noting that different factorizations $\boldsymbol{\Sigma}[\mathbf{z}] = \mathbf{AA}^{\mathrm{T}} = \mathbf{BB}^{\mathrm{T}}$ simply differ by an orthonormal matrix factor $\mathbf{O}$ (the length of the columns of $\mathbf{O}$ must be 1), e.g., $\mathbf{B} = \mathbf{AO}$. Consequently, each such matrix $\mathbf{O}$ will yield a different factorization. Therefore, choosing a different factorization, e.g., $\mathbf{B}$ instead of $\mathbf{A}$, will result in $\mathbf{B\Gamma} = \mathbf{AO\Gamma}$. Geometrically this means rotating the quadratures for the multivariate standard normal distribution (the reference Q^n) before applying the transformation to induce the desired covariance matrix.

4. Rotations evaluated

The analysis consists of eight rotations of the transformed Q^n. For the generation of the quadratures we follow the procedure described in Section 3: first, to generate the reference Q^n and second, to transform the reference Q^n for the approximation of the specific stochastic space.

The eight rotations tested are obtained by combining two rotations (0° and 45°) of the reference Q^n, two methods to introduce correlation (via the Cholesky decomposition of the covariance matrix and the diagonalization method; see Section 3.3), and two arrangements (A1 and A2) of $\mathbf{z}$ (the vector of stochastic variables in $\mathbf{Z}$). With A1, the stochastic variables are arranged in $\mathbf{Z}$ as in Eq. (7). With A2, the variables $\mathbf{z}_{ROW.\,w}$, $\mathbf{z}_{ROW.\,b}$, and $\mathbf{z}_{ROW.\,r}$, which are important for the price determination in ESIM, are located at the beginning of $\mathbf{Z}$. Fig. 3 illustrates the arrangements with A2. To distinguish between the matrices of stochastic data with different arrangements these are called $\mathbf{Z}_{A1}$ and $\mathbf{Z}_{A2}$.

The combination of 0° and 45° rotations of the reference Q^n, the methods to induce correlation, and the different arrangements of the coordinate system, results in different rotations of the transformed Q^n. Table 3 provides an overview of the tested rotations.

5. True value

The accuracy with which the numerical integration formulas approximate the integral of $f(x)$ depends not only on the degree of accuracy of the formula, but also on the accuracy with which $f(x)$ itself can be approximated by polynomials (Haber, 1970). This depends on the smoothness of $f(x)$, i.e., the number of times that the integrand is (continuously) differentiable (Haber, 1970). If we name the value

Table 3
The rotations tested.

0°A1C	45°A1C
0°A1D	45°A1D
0°A2C	45°A2C
0°A2D	45°A2D

Note: The names given to the rotations are composed of the following elements: (0°) denotes the reference n-octahedron with vertices lying on the coordinate axes; (45°) denotes the reference n-octahedron with rotation from Stroud's degree 3 formula from 1957; (A1) and (A2) denote the arrangements 1 and 2 of the stochastic variables in the covariance matrix; (C) and (D) denote the methods to induce correlation to the quadratures, via Cholesky decomposition (C) and via the diagonalization method (D).

obtained with a quadrature formula, $Q(f(x))$, then we can express its approximation error by (Haber, 1970):

$$\int_a^b f(x)dx - Q(f(x)). \tag{24}$$

For large-scale, complex stochastic simulation models, the true value of $\int_a^b f(x)dx$ cannot be computed since equilibrium conditions and multidimensionality make the analytical determination of the true value practically impossible. Therefore, the true value can only be approximated.

To compare the accuracy obtained by applying different rotations, we choose global prices for the stochastic products as important solution variables and compare them to the approximated true value. The true value is approximated using the Monte Carlo approach of numerical integration; more specifically, the LHS method (see Vose,

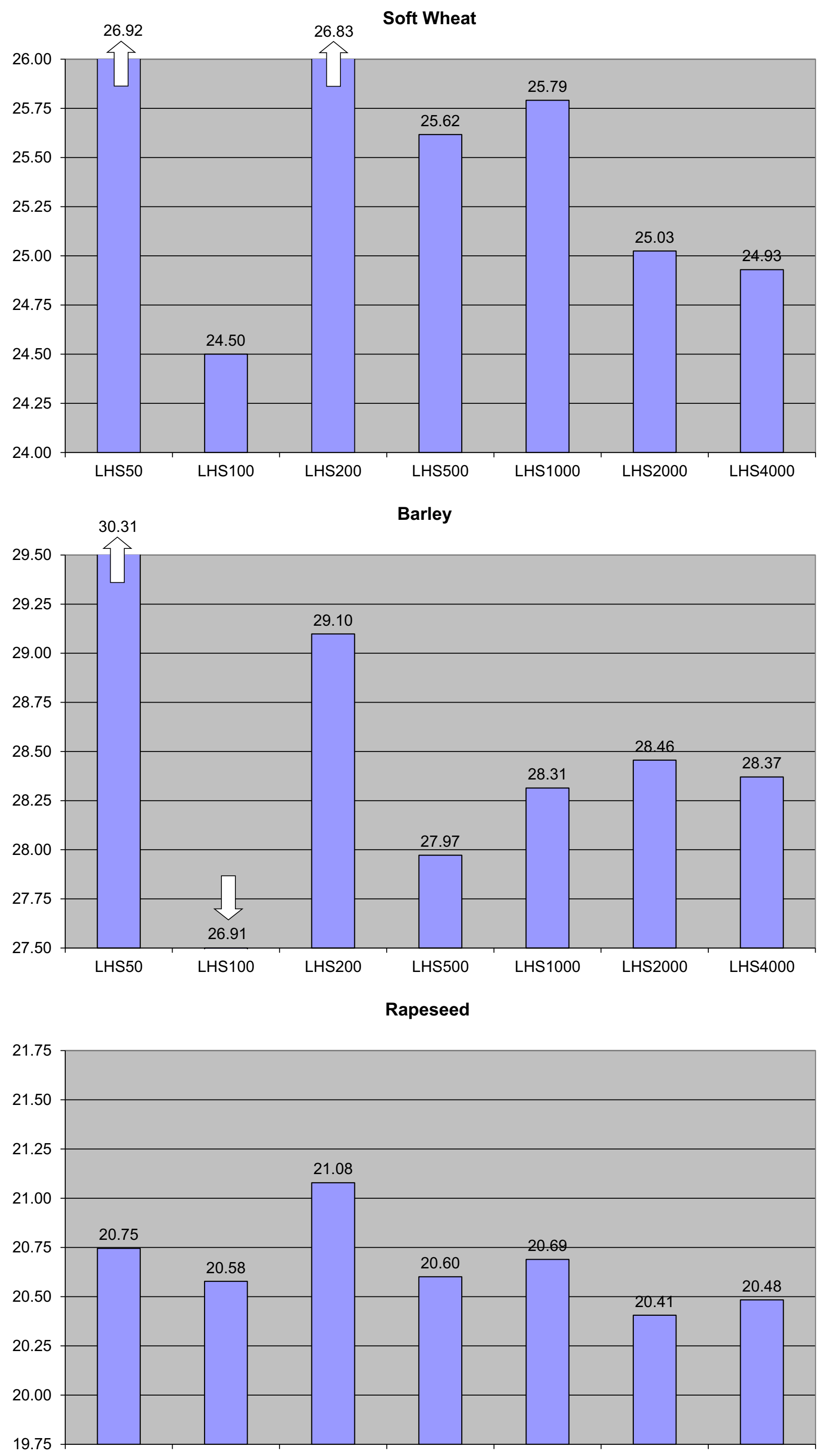


Fig. 4. Coefficient of variation (in %) of soft wheat, barley, and rapeseed world market prices with LHS samples of increasing size ($N = 50, 100, \ldots, 4000$).

2000 for an explanation of the LHS). This approach is not based on interpolation procedures and it has the advantage that due to its nature, the smoothness of $f(x)$ is not essential. Its basic idea is to treat the evaluation of the multiple integral as a probabilistic problem and to investigate it by statistical experiments (Metropolis and Ulam, 1949). However, the disadvantage is that due to its reliance on random sampling, to obtain the same degree of precision as quadrature formulas based on interpolation, a much greater number of points are required.

In this article, the true values are estimated by searching the convergence point of model results with progressively increasing LHS sample sizes of the stochastic space described in **Z**. ESIM is run using LHS sample sizes of 50, 100, 200, 500, 1000, 2000, and 4000 points. Fig. 4 shows the results of the analysis. The coefficient of variation (CV), which is an indicator of the second moment of the probability distribution of prices, is the reference for the evaluation of convergence. The first moments are easier to approximate and are thus not presented here.

Fig. 4 shows that convergence differs among crops. Convergence of the CV of soft wheat prices is the most difficult to achieve, while convergence of the CV of rapeseed prices is the easiest. For our analysis, LHS4000 is used as the reference point for all three commodities.

Lesser convergence for soft wheat may result from the complex price policies for this product, namely, an intervention price, tariff rate quotas and a threshold price, which are triggered depending on the net trade position of the EU. To test the effect of those price policies, ESIM is run again with LHS50–LHS4000, yet under the assumption of a liberalized EU wheat market. The results are presented in Fig. 5.

By comparing the soft wheat graph in Fig. 4 with Fig. 5 it can be seen that convergence is achieved faster in the liberalized market. This indicates that the simulated price policies add complexity to the numerical integration problem. Apparently, the price polices generate function surfaces or manifolds of complicated shape, which make the integration problem more complex.

In Figs. 4 and 5, it can also be seen that it is more difficult to approximate the CV of prices of soft wheat (even liberalized, see Fig. 5) and barley than rapeseed. Among other reasons, this behavior may occur because, in ESIM, soft wheat and barley have stronger cross relationships with each other than with rapeseed. Thus, stochastic dimensions with a significant effect on the soft wheat and barley markets are more than that for rapeseed, adding further complexity to the numerical integration of these variables.

6. Results and discussion

6.1. Results

The accuracy of the eight rotations of the transformed Q^n (see Table 3) is evaluated in comparison to the estimated LHS4000 value. Fig. 6 displays results and shows that the accuracy of the rotations differs substantially.

The core conclusions are: 1) the quadratures based on the 45° rotation are more accurate than those based on the 0° rotation; 2) for quadratures based on the 45° rotation and for soft wheat, the difference between the arrangements of the coordinate system or the selected method to introduce correlation are important; 3) for quadratures based on the 45° rotation and for barley and rapeseed, the differences between linear transformations of the reference Q^n are ambiguous; and 4) for quadratures based on the 45° rotation and for barley, systematic lower coefficients of variations than the approximated ‘true value’ are obtained.

Several factors are identified which contribute to the explanation of the observed differences in accuracy:

a) the rotations tested result in very different samples of the stochastic space (in another words, the rotations result in different discrete approximations of the marginal probability distributions of the stochastic variables included as factors to the supply and yield equations — see Eqs. (4) and (5));
b) in ESIM, supply changes in large producing countries have stronger effects on prices than changes in small producing countries (asymmetrical consequences on prices of supply shocks in different countries); and
c) in ESIM, the simulated supply shocks to the left and to the right in one country (see Fig. 1) result in asymmetrical positive and negative price changes due to the functional forms of supply and demand.

6.2. Rotations result in different samples

Fig. 7 illustrates how the rotations result in different samples of the stochastic space. It presents the discrete approximations (samples) of the marginal probability distributions of the stochastic variables for soft wheat in the regions ‘ROW’ and the USA which are the two largest producing regions in ESIM. In addition to the histograms of the samples obtained with the 8 rotations, the histogram of the sample obtained

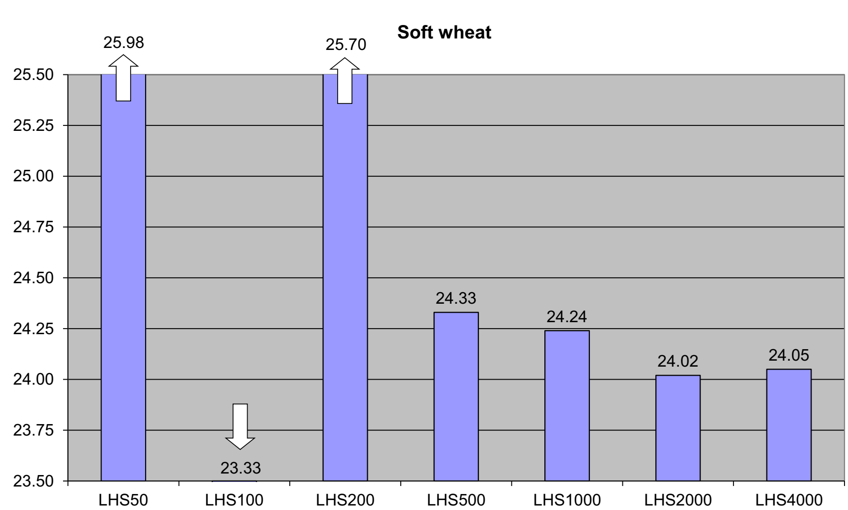
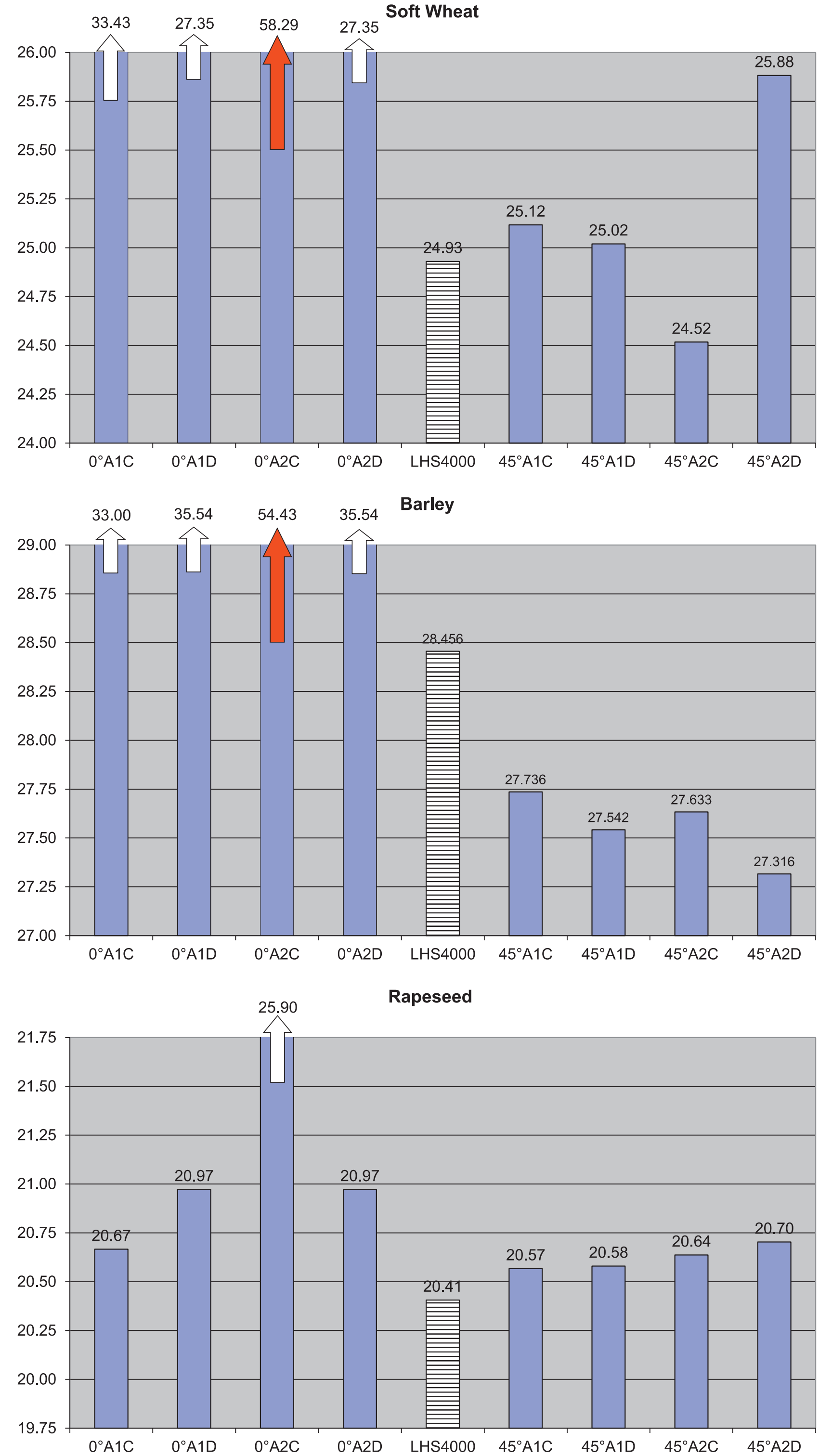
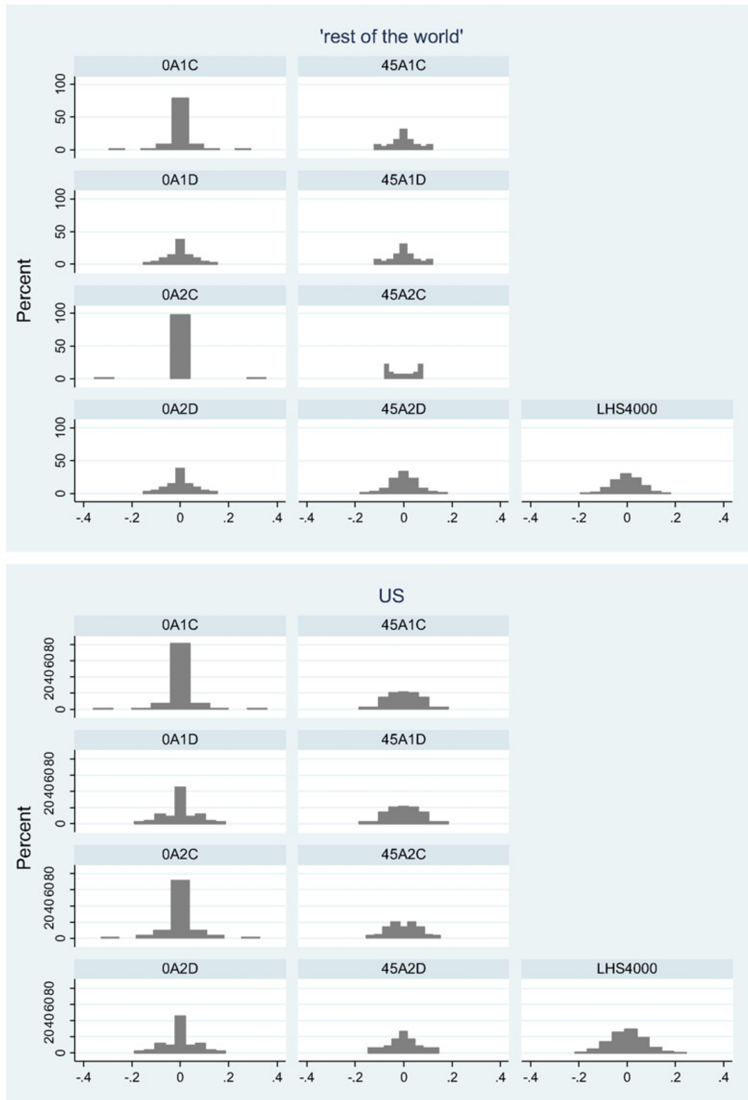


Fig. 5. Coefficient of variation (in %) of soft wheat world market prices with LHS samples of increasing size ($N = 50, 100, \ldots, 4000$) under the scenario of a liberalized EU wheat market.



Note: The names given to the rotations are composed of the following elements: (0°) denotes the reference n-octahedron with vertices lying on the coordinate axes; (45°) denotes the reference n-octahedron with rotation from Stroud's degree 3 formula from 1957; (A1) and (A2) denote the arrangements 1 and 2 of the stochastic variables in the covariance matrix; (C) and (D) denote the methods to induce correlation to the quadratures, via Cholesky decomposition (C) and via the diagonalization method (D). LHS4000 indicates the Latin Hypercube Sample of size N=4000 and is the approximated 'true value'.

Fig. 6. Coefficient of variation (in %) of prices of soft wheat, barley, and rapeseed in the 'Rest of the World' obtained with the rotations tested and the LHS4000.



Note: The names given to the rotations are composed of the following elements: (0°) denotes the reference n-octahedron with vertices lying on the coordinate axes; (45°) denotes the reference n-octahedron with rotation from Stroud's degree 3 formula from 1957; (A1) and (A2) denote the arrangements 1 and 2 of the stochastic variables in the covariance matrix; (C) and (D) denote the methods to induce correlation to the quadratures, via Cholesky decomposition (C) and via the diagonalization method (D). LHS4000 indicates the Latin Hypercube Sample of size N=4000 and is the approximated 'true value'.

Fig. 7. Histograms of the discrete probability distribution approximations of the stochastic variables regarding the two largest soft wheat producers in ESIM obtained with the rotations tested and the LHS4000.

with the LHS method with the size of 4000 points (the one used for the generation of the 'true value') is also given and serves as the reference. In Fig. 7 it can be seen how the rotation of the reference Q^n, 0° or 45°, is crucial for the determination of the allocation of the sample points. With the 0° rotations (see left column in the figure) the samples are often characterized by a concentration of points around zero and a few extreme values at the ends of the tails of the marginal distributions. In the case of the quadratures generated with the 45° rotations, the samples are characterized by avoiding points on the tails of the distributions; a systematic concentration of points in one or some parts of the domain of integration cannot be identified.

By analyzing the samples of all the stochastic variables we find that for the 0° rotations, 48% of the variables are approximated with extreme values which are at least 20% lower or 20% higher than the reference *min* and *max* values of LHS4000, respectively. For the 45° rotations, 98% of the discrete approximations present *min* and *max* points which are at least 20% higher or 20% lower than the reference *min* and *max* values, respectively.

The combination of the different samples of the stochastic space with the 0° and 45° rotations and the asymmetries in ESIM (points b and c above), explain the differences in the accuracy of the rotations. The remainder of this section discusses the implication of the asymmetries in ESIM.

6.3. Price changes and large and small producing countries

Fig. 8 shows a simplification of the price mechanism in ESIM, with prices being determined at the equilibrium point where net exports equal net imports. In addition, the effect on world market prices of a supply shock in one of the regions is shown. If the region is a large producer, the effect of a supply shock will result in a significant price change. If a country/region is a small producer, the same supply shock (in %) will have only a small effect on the world market price.

Fig. 9 shows the supply and demand quantities of soft wheat, barley and rapeseed in the largest regions in ESIM in 2015, the year in which the stochastic runs are applied. Note that ESIM depicts the markets in the EU at the MS level and not as an aggregate; nonetheless, the EU is included in the figure for reference purposes only (bars with line fillings).

In Fig. 9, it can be seen how soft wheat is particularly asymmetrical with a large region (ROW), a medium size country (USA), and some other rather small countries in the EU. Thus, the approximation of the probability distribution of the stochastic variables of the large regions is of special importance since changes in supply in these regions result in significant changes in prices. This characteristic of the soft wheat market explains the difference in accuracy obtained with the 45° rotation and arrangements A1 and A2 of the covariance matrix **Z**. The samples of the probability distributions of the stochastic variables of these large regions appear to be very significant in the determination of prices and thus, the small changes in the allocation of the points of those samples in A1 and A2 result in different coefficients of variations of soft wheat prices (see graph for soft wheat in Fig. 6). Note also that significant differences are observed for rotations 45°A2C and 45°A2D. This indicates that again, if some regions strongly influence the developments of one market in the model, then small changes in the stochastic sample may significantly affect the approximation of the moments of model results (i.e. the coefficient of variation of prices).

6.4. Behavioral functions and price changes

The second important asymmetry in ESIM is that with isoelastic supply and demand functions, positive and negative shocks of equal magnitude result in price changes of different sizes. These differences are particularly strong with inelastic curves, exemplified in Fig. 10.

A situation as depicted in Fig. 10 occurs in the market for barley in ESIM. The reaction of demand in the ROW, a region that covers around 70% of world demand, is very low. This explains why for all 45° rotations the approximated CVs of barley prices are below the true value. Those rotations avoid points on the tails of the distributions which would result in large price changes. Without those points, the generated CVs are smaller.

The accuracy of the approximation of the CV of rapeseed prices is easier to achieve since this market has fewer asymmetries. Unlike for barley and wheat, neither world supply nor world demand are heavily concentrated in one of the regions of the model. Furthermore, the demand functions are more elastic which results in more symmetrical price changes.

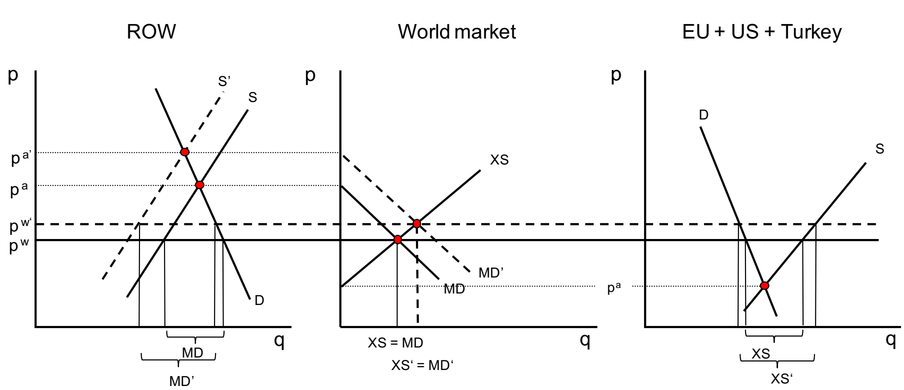
7. Conclusions and outlook

In this article, we explore the accuracy of i) an alternative (0°) rotation to Stroud's degree 3 formula 45° rotation of the reference Q^n, ii) different methods to introduce covariance matrices to standard normal quadratures (via the Cholesky decomposition or the diagonalization method of the covariance matrix), and iii) different arrangements of the coordinate system (A1 and A2). The quadratures are tested in ESIM which is a complex PE simulation model of global agricultural markets. The evaluation of the accuracy of the different quadratures is achieved through comparison with an approximated true value of model results.

It was found that the choice between the 0° and 45° rotation of the reference Q^n is crucial for the determination of the accuracy of the quadratures. With high n, the quadratures based on the 0° rotation result in samples with a concentration of points near the center and some few extreme points at the latter ends of the tails of the marginal probability distributions of the stochastic variables. The location of points of those samples combined with asymmetries in ESIM result in an inaccurate approximation of model results. The quadratures based on the 45° rotation result in sample points which avoid the tails of the marginal distributions of the stochastic variables; a systematic concentration of points in one or some parts of the domain of integration cannot be identified. These rotations yield greater accuracy.

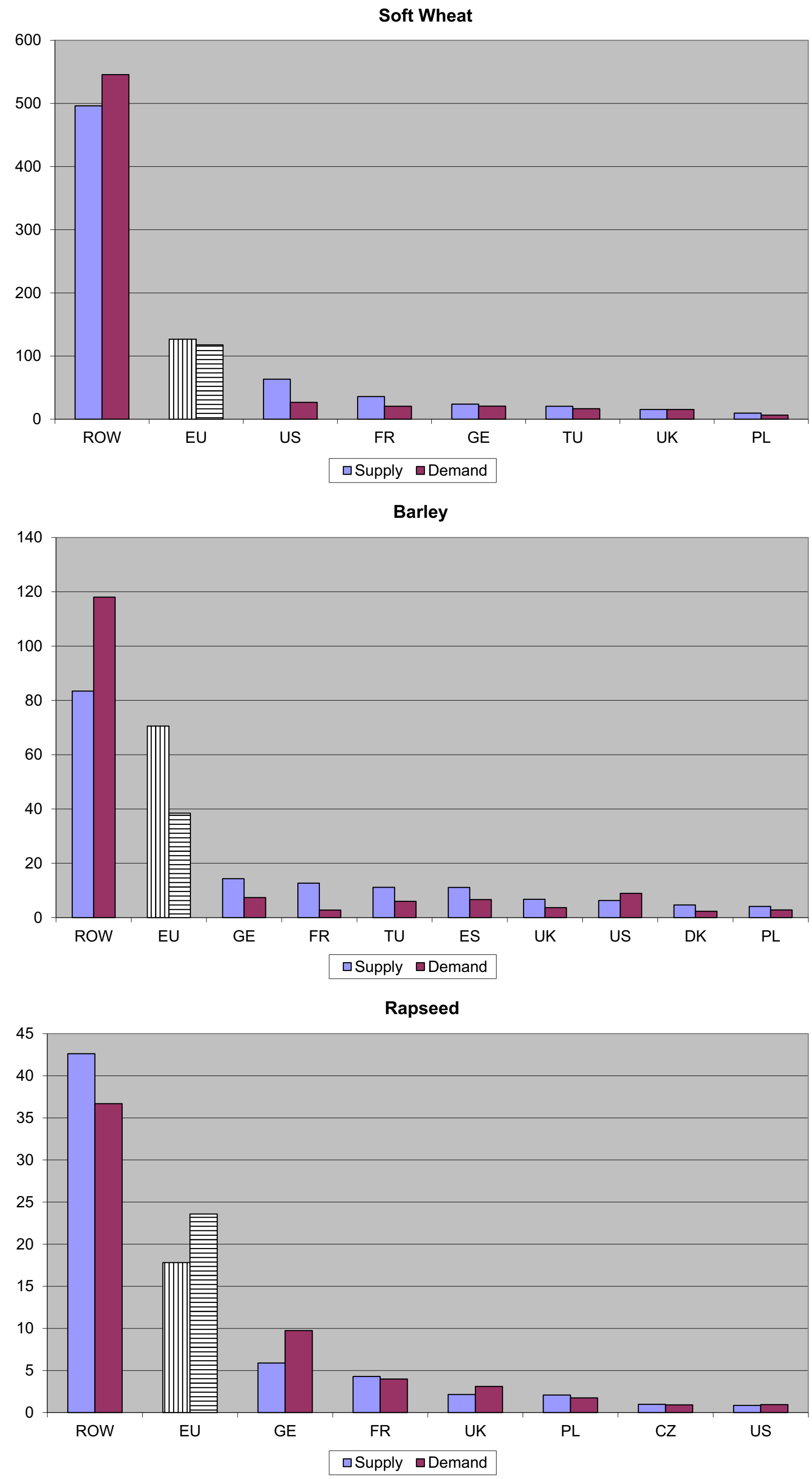
In models with large regions or variables which strongly determine the outcome of model results as is the case of soft wheat in ESIM, the arrangement of the coordinate system or the selected method to introduce correlation may also have a significant effect on the accuracy of the quadratures. In these cases, an analysis of the stability of results is recommended. This can be done by repetitive model solves with different arrangements of the stochastic variables in the covariance matrix or by using different linear transformation methods.

With the 45° rotation and with markets where the effect of the different regions or variables on model outcomes are more homogenous as in the case of rapeseed in ESIM, the selection of different arrangements of the stochastic variables in the covariance matrix or of different



Note: (p^a) denotes the price by autarky; (p^w) denotes the world market price which under free trade equals the domestic price; (S) denotes the supply curve; (D) denotes the demand curve; (XS) denotes the total export supply curve; (MD) denotes the total import demand curve; and the apostrophe (') denotes the situation after a shock to the system.

Fig. 8. A simplification of the price determination mechanism in ESIM.

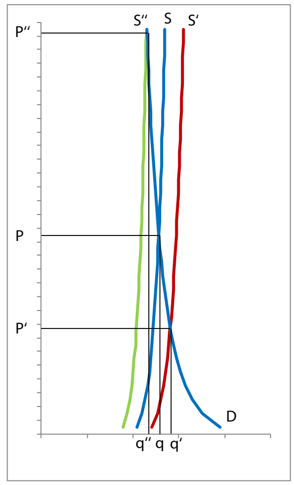


Note: The countries and regions are abbreviated in the following way: Czech Republic (CZ), Denmark (DK), European Union of 27 MS without Croatia (EU), France (FR), Germany (GE), Poland (PL), 'Rest of the World' (ROW), Spain (ES), Turkey (TU), United Kingdom (UK), United States of America (US).

Fig. 9. Supply and demand (in million tonnes) quantities of soft wheat, barley and rapeseed in the largest regions in ESIM in 2015.

methods to introduce correlation may not have a significant effect on the accuracy of the quadratures.

The quadratures based on the 45° rotation give an accurate estimate of the uncertainty of model results in ESIM and simplify stochastic analyses by strongly reducing the number of solves required when compared to the LHS sampling method. However, if models are highly asymmetric or with many and strong threshold points, the quadratures may lose accuracy. In these cases, the alternative of using Monte Carlo-



Note: (P) denotes the price at the equilibrium point; (S) denotes the supply curve; (D) denotes the demand curve; and the single and double apostrophe (') and (") denote the situations after rightward and leftward shocks of equal magnitude.

Fig. 10. Price effect of supply shocks with inelastic supply and demand functions.

based approaches should be evaluated. Factors such as higher computational and management costs vs. accuracy gains must be considered.

As a future research agenda for the refinement of the application of Gaussian Quadratures in large-scale simulation models, one may test rotations from the reference Q^n slightly different from 45° (for example 30° or 40°). This may result in the inclusion of values from the tails of the marginal distributions of the stochastic variables and in a higher diversity of points. On the other hand, the dependency on the arrangement of the coordinates may increase.

Another area for future research is the stability of the quadratures based on the 45° rotation with higher dimensions. Will the discrete approximations (samples) of the marginal distributions of the stochastic variables always avoid the tail ends of the distributions? One may also test whether and to what extent higher order quadratures approximate the LHS4000 values more accurately.

Finally, the GQ can be evaluated in other PE and CGE models. Through comparison of the performance of the GQ and of model characteristics, factors affecting the accuracy of the quadratures may be identified.

References

Anderson, K., Valenzuela, E., 2008. Estimates of Distortions to Agricultural Incentives: 1955 to 2007. World Bank, Washington D.C. (spreadsheet at www.worldbank.org/agdistortions).
Arndt, C., Hertel, T.W., 1997. Revisiting the fallacy of free trade. Rev. Int. Econ. 5 (2), 221–229.
Artavia, M., 2014. Stochastic multi-market modeling with efficient quadratures: does the rotation of Stroud's octahedron matter?PhD dissertation Humboldt University of Berlin.
Artavia, M., Möller, T., Grethe, H., Zimmermann, G., 2009. Correlated order three Gaussian quadratures in stochastic simulation modelling. Paper presented at the 12th Annual Conference on Global Economic Analysis (10–12 June, Santiago de Chile).
Banse, M., Grethe, H., Nolte, S., 2005. European Simulation Model (ESIM) in GAMS: Model Documentation, Model documentation prepared for DG. AGRI, European Commission.
Banse, M., Grethe, H., Nolte, S., 2007. European Simulation Model (ESIM) in GAMS: model documentation. Prepared for DG AGRI. European Commission.
Beckman, J., Hertel, T.W., Tyner, W., 2011. Validating energy-oriented CGE models. Energy Econ. 33, 799–806.
Burrell, A., Nii-Naate, Z., 2013. Partial stochastic analysis with the European Commission's version of the AGLINK-COSIMO model. Eur. Comm. JRC Sci. Policy Rep.
European Commission, 2011. Prospects for Agricultural Markets and Income in the EU 2011–2021. European Commission, Brussels.
European Commission, 2012. Prospects for Agricultural Markets and Income in the EU 2012–2022. European Commission, Brussels.
FAPRI-MU, 2002. U.S. Baseline Briefing Book.
FAPRI-MU, 2003. U.S. Baseline Briefing Book.
FAPRI-MU, 2004. U.S. Baseline Briefing Book.
FAPRI-MU, 2005. U.S. Baseline Briefing Book.
FAPRI-MU, 2006. U.S. Baseline Briefing Book.
FAPRI-MU, 2007. U.S. Baseline Briefing Book.
FAPRI-MU, 2008. U.S. Baseline Briefing Book.
FAPRI-MU, 2009. U.S. Baseline Briefing Book.
FAPRI-MU, 2010. U.S. Baseline Briefing Book.
FAPRI-MU, 2011a. FAPRI-MU Stochastic U.S. Crop Model Documentation. FAPRI-MU Report No. 09-11.
FAPRI-MU, 2011b. U.S. Baseline Briefing Book.
FAPRI-MU, 2012. U.S. Baseline Briefing Book.
FAPRI-MU, 2013. U.S. Baseline Briefing Book.
Haber, S., 1970. Numerical evaluation of multiple integrals. SIAM Rev. 12 (4), 481–526.
Hertel, T.W., Reimer, J.J., Valenzuela, E., 2005. Incorporating commodity stockholding into a general equilibrium model of the global economy. Econ. Model. 22, 646–664.
Hertel, T.W., Martin, W., Leister, A., 2010. Potential Implications of a special safeguard mechanism in the World Trade Organization: the case of wheat. World Bank Econ. Rev. 24 (2), 330–359.
Metropolis, N., Ulam, S., 1949. The Monte Carlo method. J. Am. Stat. Assoc. 44 (247), 335–341.
OECD, 2011. Aggregate Model Analysis of Exogenous Risk and Price Variability, (Paris).
OECD/FAO., 2007. OECD-FAO Agricultural Outlook 2007–2016. OECD Publishing and FAO.
OECD/FAO., 2008. OECD-FAO Agricultural Outlook 2008–2017. OECD Publishing and FAO.
OECD/FAO., 2011. OECD-FAO Agricultural Outlook 2011–2020. OECD Publishing and FAO.
OECD/FAO., 2013. OECD-FAO Agricultural Outlook 2013–2022. OECD Publishing and FAO.
Pearson, K.R., Arndt, C., 2000. How to carry out systematic sensitivity analysis via Gaussian quadrature and GEMPACK. GTAP Technical Paper No. 3.
Preckel, P.V., Verma, M., Hertel, T.W., Martin, W., 2011a. Implications of broader Gaussian quadrature sampling strategy in the contest of the special safeguard mechanism. Paper presented at the 14th Annual Conference on Global Economic Analysis (16–18 June, Venice, Italy).
Preckel, P.V., Verma, M., Hertel, T.W., Martin, W., 2011b. Systematic Sensitivity Analysis for GTAP — Where We've Been and Where We're Going. Paper Presented at the 14th Annual Conference on Global Economic Analysis (16–18 June, Venice, Italy).
Stroud, A.H., 1957. Remarks on the disposition of points in numerical integration formulas. Math. Tables Other AIDS Comput. 11 (60), 257–261.
Valenzuela, E., Hertel, T.W., Keeney, R., Reimer, J.J., 2007. Assessing global computable general equilibrium model validity using agricultural price volatility. Am. J. Agric. Econ. 89 (2), 383–397.
Verma, M., Hertel, T.W., Valenzuela, E., 2011. Are the poverty effects of trade policies invisible? World Bank Econ. Rev. 25 (2), 190–211.
Vose, D., 2000. Risk Analysis: A Quantitative Guide, 2nd ed. John Wiley & Sons Ltd.
Westhoff, P., Brown, S., Hart, C., 2005. When point estimates miss the point: stochastic modeling of WTO restrictions. J. Int. Agric. Trade Dev. 2 (1), 87–107.