Correlated Order Three Gaussian Quadratures in Stochastic Simulation Modelling

---Draft Version of 15.04.2009, not for Citation----

Marco Artavia[*], Harald Grethe[**], Thordis Möller[***], Georg Zimmermann[****]

Paper prepared for the Twelfth Annual Conference on Global Economic Analysis,

Santiago, Chile, June 10-12, 2009

Abstract

A simple approach to generate correlated order three Gaussian quadratures is presented. It is shown how for stochastic simulation modelling purposes the integration over the *n*-cube as suggested in Stroud (1957) is not necessary and thus, a simpler integration formula is given. The theory of inducing a desired covariance matrix is presented and three possibilities to do so are demonstrated. The approach described is implemented in a stochastic version of the European Simulation Model (ESIM) which includes 42 correlated stochastic terms in the yield functions. Model results are presented to validate the proposed approach compared to a Monte Carlo based approach as well as to demonstrate the relevance of stochastic simulation modelling.

---

[*] Contact author: Ph.D. Student, Department of Agricultural Policy, Humboldt University of Berlin Unter den Linden 6, 10099 Berlin, Germany.
Tel.: +(49)-30-20 93 64 28. E-Mail: marco.artavia@agrar.hu-berlin.de

[**] Professor, Agricultural and Food Policy Group, University of Hohenheim

[***] Ph.D. Student, Department of International Agricultural Trade and Development, Humboldt University of Berlin

[****] Privatdozent, Department of Applied Mathematics, University of Hohenheim

1. INTRODUCTION

Current concerns about the impact of climate change on agriculture as well as the food security effects of high volatility in agricultural commodity prices accentuate the importance of taking uncertainty into account when conducting agricultural policy, market and trade analysis. Stochastic simulation modelling can capture the uncertainty attached to variables such as yield, which may vary due to climatic conditions, or model parameters based on weak empirical data, such as supply and demand elasticities. Such simulations generate model results which give more information than point estimates. This additional information may be valuable from a scientific perspective as well as for private and public sector decision makers.

Even though the volatility of agricultural markets has played a major role in recent years, there are only few stochastic applications of large scale simulation models which address the agricultural sector in some detail. Westhoff et al. (2005) introduce stochastic terms to a set of supply-side and demand-side equations in the US component of the FAPRI partial equilibrium model in order to evaluate the probability of failing to cope with World Trade Organization (WTO) commitments. OECD (2003) presents a stochastic application of the AGLINK model assessing yield volatility. Hertel et al. (2005) formulate a stochastic version of the GTAP model to analyze the linkage between supply-side uncertainty and stockholding.

For other large scale partial equilibrium models playing an important role in the analysis of agricultural markets such as AGMEMOD, CAPSIM, CAPRI and IMPACT no stochastic versions have been published (see Verhoog et al. (2008) for references on the models).

Westhoff et al. (2005) as well as OECD (2003) apply the Monte Carlo (MC) procedure in order to approximate the empirical distribution of stochastic variables through correlated random draws and run the models in repeated solves. A major disadvantage of this approach is the high computational requirement which results from the high number of solves required.

Hertel et al. (2005) apply an alternative numerical integration approach: the Gaussian quadratures (GQ). Its main idea is to generate points and associated weights with their weighted sum matching the moments of the original distribution function. This procedure, when compared to Monte Carlo, significantly reduces the number of points needed to approximate a desired distribution. Hertel et al. (2005) characterize the supply side uncertainty based on the Gaussian quadrature method used by DeVuyst and Preckel (1997) and developed by DeVuyst (1993). This procedure is based on the formulation of an objective function which should be minimized. The function is equal to the sum of squared differences between the moments about the origin of the original distribution and the moments of the discrete approximation (DeVuyst, 1993). This system has some difficulties in solving as numerous local minima may occur, so that a sequence of minimization problems must be formulated (see DeVuyst, 1993). Furthermore, the method requires some programming skills and some time investment in order to avoid problems with local minima. Arndt (1996) presents another way to obtain Gaussian quadratures, specifically of order three, based on formulae given by Stroud (1957). These formulae seem to be a practical approach which requires no programming skills and little time investment. In this paper we also base our proposition on Stroud (1957). However, we do not use his point generating formulae, but his theoretical scheme on the conditions that must be fulfilled. Furthermore, we argue that for stochastic modelling purposes more simple formulas can be used than those given by Stroud (1957) and Arndt (1996).

When considering yield stochasticity in market and policy analysis the correlation of the stochastic terms over products as well as countries is an important issue. Nonetheless, the topic of Gaussian quadratures with correlated terms has been rarely addressed in literature. Preckel and DeVuyst (1992), DeVuyst (1993) and Arndt (1996) present practicable procedures to generate correlated GQ, but do not distinguish clearly in their description between the application of the Cholesky decomposition and the eigen-system of the variance-covariance matrix. In presenting different ways to generate correlated GQ we clarify this issue.

In order to validate the accuracy of the numerical integration formulae and the correlation techniques presented in this paper, and to document the enormous reduction of solves required, a stochastic version of the European Simulation Model (ESIM)[1] of the agricultural sector is developed. Results obtained with the GQ approach proposed and with the MC procedure are compared. Finally, we discuss the distribution of model results for endogenous variables and document the relevance of the analysis, for example for the projection of the net trade status of the EU for cereals.

1 See Banse et al. (2005) for documentation of a deterministic ESIM version.

2. ORDER THREE MULTIVARIATE GAUSSIAN QUADRATURES WITH CORRELATED TERMS

2.1 Background

Gaussian quadrature is a method of numerical integration. It approximates an integral by a finite sum, using a set of quadrature points $x_1, \dots, x_N$ together with associated weights $p_1, \dots, p_N$. The points and their weights are chosen in such a way as to maximize the degree of polynomials for which the quadrature formula yields the correct value. From the probability point of view, this can be seen as approximating a continuous distribution by a finite discrete measure, preserving as many moments as possible. This number is the degree of exactness of the quadrature formula.

For example, considering a (univariate) density function f, we estimate the expected value of a function g by

$$E[g(x)] = \int_{\Re} g(x) f(x)\, dx \approx \sum_{k=1}^{N} p_k \; g(x_k).$$

Degree of exactness d, *i.e.*, preserving the first d moments, leads to the following equations obtained by letting $g(x) = x^j$, $j = 1, \dots, d$, in the above:

$$E\left[x^0\right] = \int_{\Re} 1\, f(x)\, dx = 1 = p_1 + p_2 + \dots + p_N$$

$$E\left[x^1\right] = \int_{\Re} x\, f(x)\, dx = p_1 x_1 + p_2 x_2 + \dots + p_N x_N$$

$$E\left[x^2\right] = \int_{\Re} x\ f(x)\,dx \qquad = p_1 x_1^2 + p_2 x_2^2 + \ldots + p_N x_N^2$$

$$\vdots$$

$$E\left[x^d\right] = \int_{\Re} x^d\ f(x)\,dx \quad = p_1 x_1^d + p_2 x_2^d + \ldots + p_N x_N^d\,.$$

In the multivariate case, we consider a multiple integral over $\Re^n$. Consequently, the quadrature points will be vectors with n components (corresponding to n stochastic variables), each again with an associated weight representing a probability.

Naturally, the number N of required quadrature points will depend on the dimension n and on the desired degree d of exactness.

2.2 Description of the Problem

We need to approximate integrals with multivariate normal distributions $N\,(\boldsymbol{o}, \boldsymbol{\Sigma})$ as weights, with covariance matrix $\boldsymbol{\Sigma} \neq \boldsymbol{I_n}$. We shall use a theorem from Stroud (1957) as starting point for the solution of this problem.

More explicitly, we need to find approximate values for integrals of the form

$$I(f) = \int_{\Re} \cdots \int_{\Re} f(\boldsymbol{x}) \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}\ dx_1 \ldots dx_n\,.$$

For this purpose, we want to employ a quadrature formula whose degree of exactness is 3, *i.e.*, points $\boldsymbol{x_1}, \ldots, \boldsymbol{x_N} \in \Re^n$ with weights $w_1, \ldots, w_N$ such that the numerical integration formula

$$I_d(f) = \sum_{k=1}^{N} w_k f(\boldsymbol{x_k})$$

satisfies $I_d(f) = I(f)$ for all polynomials of total degree at most 3.

Polynomials of degree 0 are simply constants, and letting $f \equiv 1$ yields the condition

$$\sum_{k=1}^{N} w_k = \int_{\Re} \cdots \int_{\Re} \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})} dx_1 \dots dx_n = 1$$

Furthermore, we would like the numerical integration formula to be equally weighted; therefore, $w_k = 1/N, \quad k = 1, \dots, N.$

From the probability point of view, this approach amounts to approximating the (continuous) normal distribution $N\ (\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with density

$$\frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$$

by a discrete equidistribution on N points exhibiting the same moments up to order 3.

Finally, for symmetry reasons, we want the vertices to be in pairs of the form $\{\boldsymbol{x}, -\boldsymbol{x}\}$. The following theorem from Stroud (1957) shows that we need $N = 2\,n$ quadrature points to obtain formulas of degree 3 for symmetrical regions in $\Re^n$.

"A necessary and sufficient condition that 2n points $\boldsymbol{x_1}, \dots, \boldsymbol{x_n}, -\boldsymbol{x_1}, \dots, -\boldsymbol{x_n}$ *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: 259).

Here, Stroud uses the following notation:

n is the dimension of the problem (in our case, the number of stochastic variables)

Q_n is the n-dimensional generalized octahedron ("*n*-octahedron")

I_0 is the integral of a constant (in our case, the 0th moment of the distribution)

I_2 is the integral of the square of any variable (in our case, the variance(s) of the distribution)

The fact that I_2 is independent of the variable chosen in the integral stems from the symmetry of the region. In our case, we have to allow for different 2nd moments and therefore need to apply appropriate dilations; see below.

A standard example for a symmetric region of integration is the *n*-cube with vertices $(\pm a, \pm a, \ldots, \pm a)^T$. The standard *n*-octahedron has the vertices $(\pm b, 0, \ldots, 0)^T$, $(0, \pm b, 0, \ldots, 0)^T, \ldots,$ $(0, \ldots, 0, \pm b)^T$. The condition $r = \sqrt{nI_2/I_0}$ from Stroud's theorem means $b = \sqrt{\frac{n}{3}}\, a$ which implies that for $n \geq 4$, the vertices of the standard *n*-octahedron will be outside the *n*-cube. For this reason, the *n*-octahedron has to be rotated to obtain a quadrature formula for the n-cube with vertices inside the region of integration (*cf.* once more Stroud (1957)).

In our problem, we do not integrate over the n-cube, but over the n-dimensional Euclidean space $\Re^n$ (with a weight function). Therefore, there is no need to apply the rotation suggested in Stroud (1957), and we may simply use the vertices of the standard n-octahedron to begin with (*cf.* the following section). We need to transform these vertices, however, in order to introduce the desired covariance terms, as described further below.

2.3 Quadratures for the Multivariate Standard Normal Distribution with Independent Terms ($\boldsymbol{\Sigma} = \boldsymbol{I}_n$)

The density of the multivariate standard normal distribution $N\,(\boldsymbol{0}, \boldsymbol{I}_n)$ shows spherical symmetry. Therefore, all rotations of the n-octahedron are equally well suited. So we might as well use the standard n-octahedron of the appropriated size, which turns out to be given by $r = \sqrt{n}$, as the following calculations show.

The vertices are

$$\boldsymbol{\xi}_{1,2} = \left(\pm\sqrt{n}, 0, \ldots, 0\right)^T, \boldsymbol{\xi}_{3,4} = \left(0, \pm\sqrt{n}, 0, \ldots, 0\right)^T, \ldots, \boldsymbol{\xi}_{2n-1,\,2n} = \left(0, 0, \ldots, \pm\sqrt{n}\ \right)^T.$$

Using these as quadrature points with equal weights $\frac{1}{2\,n}$, we obtain the following moments up to order 3:

0) $$\frac{1}{2n}\sum_{k=1}^{2n} 1 = 1$$

1) $$\frac{1}{2n}\sum_{k=1}^{2n} \xi_{k,\ell} = 0 \quad \text{for all } \ell \in \{1,\dots,n\}$$

$$\Leftrightarrow \quad \frac{1}{2n}\sum_{k=1}^{2n} \boldsymbol{\xi_k} = \boldsymbol{o}$$

2) $$\frac{1}{2n}\sum_{k=1}^{2n} \xi_{k,\ell_1}\, \xi_{k,\ell_2} = \delta_{\ell_1,\ell_2} \quad \text{for all } \ell_1, \ell_2 \in \{1,\dots,n\}$$

$$\Leftrightarrow \quad \frac{1}{2n}\sum_{k=1}^{2n} \boldsymbol{\xi_k}\, \boldsymbol{\xi_k^T} = \boldsymbol{I_n}$$

3) $$\frac{1}{2n}\sum_{k=1}^{2n} \xi_{k,\ell_1}\, \xi_{k,\ell_2}\, \xi_{k,\ell_3} = 0 \quad \text{for all } \ell_1, \ell_2, \ell_3 \in \{1,\dots,n\}$$

Since these are the moments of the standard normal distribution, we indeed have the desired degree of exactness.

Applying any rotation of the $\Re^n$ about the origin will not change these moments, since the symmetry is preserved. Therefore, the use of the vertices suggested in Stroud (1957) for this type of problem is not false, but complicates matters unnecessarily.

2.4 Methods to Induce a Desired Covariance Matrix ($\boldsymbol{\Sigma} \neq \boldsymbol{I}_\mathbf{n}$)

Naturally, our next step is a modification of the quadrature formula obtained above in order to introduce arbitrary variances and correlations. To this end, we firstly study how a linear transformation applied to the vertices of the standard *n*-octahedron will influence the covariance matrix. This will lead us to different methods of generating quadrature points with the desired properties. Finally, we describe the necessary modification in case $E[\boldsymbol{x}] = \boldsymbol{\mu} \neq \boldsymbol{0}$.

To begin with, we consider an equidistribution on N arbitrary points $\boldsymbol{x_1},\ldots,\boldsymbol{x}_N \in \Re^n$ with weight $1/N$ and mean $E[\boldsymbol{x}] = (1/N)(\boldsymbol{x_1} + \cdots + \boldsymbol{x_n}) = \boldsymbol{o}$. In this case, the covariance matrix can be determined by simply gathering these points in a $n \times N$ -matrix

$$\boldsymbol{X} = (\boldsymbol{x_1}|\ldots|\boldsymbol{x}_N)$$

and computing

$$\boldsymbol{COV}[\boldsymbol{x}] = \frac{1}{N}\boldsymbol{X}\ \boldsymbol{X^T} .$$

For the vertices $\boldsymbol{\xi_1},\ldots,\boldsymbol{\xi_{2n}}$ of the standard *n*-octahedron described above, this yields

$$\begin{aligned}\boldsymbol{\Xi} &= (\xi_1|\ldots|\xi_{2n}) \\ &= \begin{pmatrix} \sqrt{n} & -\sqrt{n} & 0 & 0 & \ldots & \ldots & 0 & 0 \\ 0 & 0 & \sqrt{n} & -\sqrt{n} & \ldots & \ldots & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & \ldots & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & \ldots & \sqrt{n} & -\sqrt{n} \end{pmatrix}\end{aligned}$$

and

$$\boldsymbol{COV}[\xi] = \frac{1}{2n}\boldsymbol{\Xi}\ \boldsymbol{\Xi^T} = \boldsymbol{I_n} ,$$

as claimed before.

Now let $\boldsymbol{A}$ be any regular $n \times n$-matrix and consider the points $\boldsymbol{x_x} = \boldsymbol{A}\,\boldsymbol{\xi_k}$, $k = 1,\ldots,2n$. This yields

$$\boldsymbol{X} = \left(\boldsymbol{A\xi_1}\middle|\ldots\middle|\boldsymbol{A\xi_2}\right) = \boldsymbol{A}\,\boldsymbol{\Xi}$$

with

$$E[\boldsymbol{x}] = \frac{1}{2n}\sum_{k=1}^{2n}\boldsymbol{x_k} = \frac{1}{2n}\boldsymbol{A}\sum_{k=1}^{2n}\boldsymbol{\xi_k} = \boldsymbol{A}\,E[\boldsymbol{\xi}] = \boldsymbol{A}\,\boldsymbol{o} = \boldsymbol{o}$$

and

$$\boldsymbol{COV}[\boldsymbol{x}] = \frac{1}{2n}\boldsymbol{X}\,\boldsymbol{X^T} = \frac{1}{2n}\boldsymbol{A}\,\boldsymbol{\Xi}\,\boldsymbol{\Xi^T}\boldsymbol{A^T} = \boldsymbol{A}\,\boldsymbol{A^T}.$$

Thus, our problem is reduced to expressing the desired covariance matrix $\boldsymbol{\Sigma}$ in the form $\boldsymbol{A}\,\boldsymbol{A^T}$ for a regular square matrix $\boldsymbol{A}$. There are countless possibilities of doing this; we describe three standard methods.

- **Diagonalization (Principal Axes Transformation):**

  Since $\boldsymbol{\Sigma}$ is positive semidefinite, it can be written in the form $\boldsymbol{\Sigma} = \boldsymbol{U}\,\boldsymbol{D}\,\boldsymbol{U^T}$, where $\boldsymbol{D}$ is the (non-negative) diagonal matrix of eigen-values of $\boldsymbol{\Sigma}$ and $\boldsymbol{U}$ is orthogonal (consisting of the eigenvectors of $\boldsymbol{\Sigma}$). Letting $\boldsymbol{A} = \boldsymbol{U}\sqrt{\boldsymbol{D}}$ yields $\boldsymbol{A}\,\boldsymbol{A^T} = \boldsymbol{\Sigma}$ as desired.

- **Cholesky Decomposition:**

  The positive definite matrix $\boldsymbol{\Sigma}$ has a Cholesky decomposition $\boldsymbol{\Sigma} = \boldsymbol{L}\,\boldsymbol{L^T}$ where $\boldsymbol{L}$ is lower triangular; choose $\boldsymbol{A} = \boldsymbol{L}$.

- **Reverse Cholesky Decomposition:**

  Instead, we can also use the reverse Cholesky decomposition $\boldsymbol{\Sigma} = \boldsymbol{R}\,\boldsymbol{R}^{\boldsymbol{T}}$ where $\boldsymbol{R}$ is upper triangular and choose $\boldsymbol{A} = \boldsymbol{R}$.

As shown above, for each choice of $\boldsymbol{A}$ with $\boldsymbol{A}\,\boldsymbol{A}^{\boldsymbol{T}} = \boldsymbol{\Sigma}$, the columns of the matrix $\boldsymbol{X} = \boldsymbol{A}\,\boldsymbol{\Xi}$ yield quadrature points with $E[\boldsymbol{x}] = \boldsymbol{o}$ and $\boldsymbol{COV}[\boldsymbol{x}] = \boldsymbol{\Sigma}$.

It is worth noting that different factorizations $\boldsymbol{\Sigma} = \boldsymbol{A}\,\boldsymbol{A}^{\boldsymbol{T}} = \boldsymbol{B}\,\boldsymbol{B}^{\boldsymbol{T}}$ simply differ by an orthogonal matrix factor, *i.e.*, $\boldsymbol{B} = \boldsymbol{A}\,\boldsymbol{O}$ for an orthogonal matrix $\boldsymbol{O}$, and that conversely, each such matrix $\boldsymbol{O}$ will yield a different factorization. Therefore, choosing a different $\boldsymbol{A}$ geometrically simply means rotating the standard *n*-octahedron before applying the transformation to induce the desired covariances.

It may be desirable to induce standard deviations and correlation values separately, *e.g.*, in simulations of future effects of climate change on supply, where the correlations have been observed in the past, but the future volatilities have to be estimated. In these cases, we start by factoring the correlation matrix in the form $\boldsymbol{P} = \boldsymbol{A}\,\boldsymbol{A}^{\boldsymbol{T}}$ and let

$$\boldsymbol{y}_{\boldsymbol{k}} = \boldsymbol{A}\,\boldsymbol{\xi}_{\boldsymbol{k}} \qquad \text{with} \qquad E[\boldsymbol{y}] = \boldsymbol{o} \qquad \text{and} \qquad \boldsymbol{COV}[\boldsymbol{y}] = \boldsymbol{P}.$$

In a second step, we define

$$\boldsymbol{x}_{\boldsymbol{k}} = \boldsymbol{D}_{\sigma}\,\boldsymbol{y}_{\boldsymbol{k}} \qquad \text{where} \qquad \boldsymbol{D}_{\sigma} = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_n \end{pmatrix}$$

and the σ_ℓ are the estimated future volatilities in the form of standard deviations for the n stochastic variables. Altogether, this yields

$$\boldsymbol{x_k} = \boldsymbol{D_\sigma A}\ \boldsymbol{\xi_k} \qquad \text{with} \qquad E[\boldsymbol{x}] = \boldsymbol{o} \qquad \text{and} \qquad \boldsymbol{COV}[\boldsymbol{x}] = \boldsymbol{D_\sigma}\ \boldsymbol{A}\ \boldsymbol{A^T}\boldsymbol{D_\sigma} = \boldsymbol{D_\sigma}\ \boldsymbol{P}\ \boldsymbol{D_\sigma},$$

as desired.

To additionally introduce nonzero means $\boldsymbol{\mu}_\ell \neq 0$, it suffices to add the vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^T$ to each of the transformed points, *i.e.*, to use

$$\boldsymbol{x_k} = \boldsymbol{A}\ \boldsymbol{\xi_k} + \boldsymbol{\mu}, \qquad k = 1, \ldots, 2n.$$

These transformations generate discrete equidistributions with the desired first and second moments. That all third centered moments are still zero – as it is the case with any normal distribution – is a consequence of the symmetry about the vector $\boldsymbol{\mu}$. Therefore, the constructed points yield a quadrature formula $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with degree of exactness 3.

Preckel and DeVuyst (1992), DeVuyst (1993) and Arndt (1996) present the diagonalization method to induce correlation. However, they somewhat misleadingly indicate to compute the Cholesky factors while referring to the matrix of eigen-vectors and to the diagonal matrix of eigen-values.

Richardson, Klose and Gray (2000) use the reverse Cholesky decomposition for factorizing a desired correlation matrix but for a Monte Carlo procedure.

@Risk and SIMETAR, support software for analyzes under uncertainty, include correlation functions using methods such as those presented above; however, in combination with the Monte Carlo framework. @Risk uses a method which applies a product of the factorization of a desired rank correlation matrix (see Iman and Conover 1982 for details on the method). SIMETAR uses the reverse Cholesky decomposition (Richardson, Klose and Gray, 2000).

2.5 Solving a Small Example

Imagine we are working with world market cereal and oilseed models and we would like to study the effect of yield volatility on the European net trade situation. For this purpose we select 6 stochastic variables *n*: barley, rapeseed and wheat yield in the EU and in the rest of the world (ROW), which we would like to analyze in detail. We evaluate time series data for the years 1993 to 2005 and estimate a linear trend. Then, the yearly deviates from the trend, in percent values, would serve as our basis from which the desired variance-covariace matrix $\boldsymbol{\Sigma}$ is taken. Furthermore, the lower triangular matrix $\boldsymbol{L}$ from $\boldsymbol{\Sigma}$ via Cholesky decomposition is computed, as well as the quadratures of the standard $\boldsymbol{Q_n}$, namely $\boldsymbol{\Xi_0}$. Moreover, we define $\boldsymbol{A} = \boldsymbol{L}$ and continue with the matrix multiplication $\boldsymbol{A}\,\boldsymbol{\Xi_0}$ in order to get $\boldsymbol{X}$. Finally, the variance-covariance matrix of $\boldsymbol{X}$: $\boldsymbol{\Sigma(X)}$ is calculated and works as our control parameter to asses whether $\boldsymbol{\Sigma}(\boldsymbol{X}) = \boldsymbol{\Sigma}$ is achieved.

In order to analyze the effect of yield volatility on the European net trade situation, the model should be repeatedly solved over the 12 generated quadratures $\boldsymbol{X} = (\boldsymbol{x_{n1}} | \ldots | \boldsymbol{x_{n12}})$.

Box 1 summarizes the punctual way to obtain the final Gaussian quadratures $\boldsymbol{X}$. As shown, $\boldsymbol{\Sigma}(\boldsymbol{X}) = \boldsymbol{\Sigma}$ is achieved.

Box 1. Solution of the small example

$$\boldsymbol{\Sigma} = \begin{pmatrix} 0,0065 & 0,0018 & 0,0038 & -0,0011 & 0,0038 & 0,0021 \\ 0,0018 & 0,0219 & 0,0009 & 0,0000 & 0,0072 & 0,0000 \\ 0,0038 & 0,0009 & 0,0037 & 0,0004 & 0,0021 & 0,0017 \\ -0,0011 & 0,0000 & 0,0004 & 0,0046 & -0,0021 & 0,0007 \\ 0,0038 & 0,0072 & 0,0021 & -0,0021 & 0,0093 & 0,0014 \\ 0,0021 & 0,0000 & 0,0017 & 0,0007 & 0,0014 & 0,0016 \end{pmatrix}$$

$$\boldsymbol{L} = \begin{pmatrix} 0,0809 & 0 & 0 & 0 & 0 & 0 \\ 0,0225 & 0,1462 & 0 & 0 & 0 & 0 \\ 0,0467 & -0,0009 & 0,0391 & 0 & 0 & 0 \\ -0,0130 & 0,0023 & 0,0259 & 0,0615 & 0 & 0 \\ 0,0469 & 0,0420 & -0,0005 & -0,0257 & 0,0683 & 0 \\ 0,0262 & -0,0038 & 0,0112 & 0,0117 & 0,0097 & 0,0227 \end{pmatrix}$$

$$\boldsymbol{\Xi}_0 = \begin{pmatrix} 2,449 & -2,449 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2,449 & -2,449 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2,449 & -2,449 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2,449 & -2,449 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2,449 & -2,449 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2,449 & -2,449 \end{pmatrix}$$

$$\boldsymbol{X} = \begin{pmatrix} 0,198 & -0,198 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0,055 & -0,055 & 0,358 & -0,358 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0,114 & -0,114 & -0,002 & 0,002 & 0,096 & -0,096 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0,032 & 0,032 & 0,006 & -0,006 & 0,063 & -0,063 & 0,151 & -0,151 & 0 & 0 & 0 & 0 \\ 0,115 & -0,115 & 0,103 & -0,103 & -0,001 & 0,001 & -0,063 & 0,063 & 0,167 & -0,167 & 0 & 0 \\ 0,064 & -0,064 & -0,009 & 0,009 & 0,027 & -0,027 & 0,029 & -0,029 & 0,024 & -0,024 & 0,056 & -0,056 \end{pmatrix}$$

$$\Sigma(X) = \begin{pmatrix} 0{,}0065 & 0{,}0018 & 0{,}0038 & -0{,}0011 & 0{,}0038 & 0{,}0021 \\ 0{,}0018 & 0{,}0219 & 0{,}0009 & 0{,}0000 & 0{,}0072 & 0{,}0000 \\ 0{,}0038 & 0{,}0009 & 0{,}0037 & 0{,}0004 & 0{,}0021 & 0{,}0017 \\ -0{,}0011 & 0{,}0000 & 0{,}0004 & 0{,}0046 & -0{,}0021 & 0{,}0007 \\ 0{,}0038 & 0{,}0072 & 0{,}0021 & -0{,}0021 & 0{,}0093 & 0{,}0014 \\ 0{,}0021 & 0{,}0000 & 0{,}0017 & 0{,}0007 & 0{,}0014 & 0{,}0016 \end{pmatrix}$$

3. INTRODUCING CORRELATED GAUSSIAN QUADRATURES IN THE EUROPEAN SIMULATION MODEL

3.1 Introduction of Stochastic Yield Terms in ESIM and Scenario

ESIM is a comparative static partial equilibrium net-trade multi-country model of agricultural production, consumption of agricultural products, and some first-stage processing activities (Banse et al., 2005). ESIM is a partial model as only a part of the economy, the agricultural sector, is modeled, i.e. macroeconomic variables (such as income or real exchange rates) are exogenous. As a world model it includes all countries, though in greatly varying degrees of disaggregation. All EU Member States as well as accession candidate Turkey plus the US are modeled as individual countries; all others are combined in one aggregate "Rest of the World" (ROW). ESIM has rich cross-commodity relations as well as a detailed representation of EU policies; it depicts price and trade policy instruments as well as direct payments. As ESIM is mainly designed to simulate the development of agricultural markets in the EU and accession candidates, policies are only modeled for these countries (i.e. for the USA and the ROW, production and consumption take place at world market prices). Area allocation, yield and demand functions are isoelastic.

In the stochastic version of ESIM, yield and area allocation variables in the EU and accession candidates are fixed at their deterministic value for 2015[2], and normally distributed and correlated stochastic terms derived according to the approach described below are added for barley, wheat and rapeseed and the model is solved repeatedly in order to finally calculate expected value and standard deviation of endogenous variables.

For this paper, we run the stochastic simulation for the year 2015 in different versions: i) with 42 stochastic yield terms and repeated model solves based on the Monte-Carlo approach (50, 100, 200, 500, 1000 and 2000 solves), and ii) with 42 stochastic yield terms based on quadratures derived based on the Cholesky decomposition as described above with 84 (2n) solves.

As a scenario we apply a standard baseline calibrated to the model base year 2005 and solving for the year 2015 including several standard assumptions with respect to variables which are exogenous to this analysis such as demographic developments, macroeconomic growth and technical progress. Furthermore, many assumptions are made for the development of the Common Agricultural Policy of the EU which include

2 This is because it is assumed that producers have no option to respond to the price signals which result from yield stochasticity in the respective year, but rather base their supply decisions on an expected price which is, in the comparative static version of ESIM, the current period deterministic price. Consistently, also the crop supply variables in non EU countries should be fixed at their deterministic values and the stochastic terms be added on top. This approach, however, due to low elasticities of demand, leads to extreme price volatility by far exceeding real world observed volatility which prevents the model from converging. Therefore, we add the ROW stochastic yield terms to the row supply functions, without fixing supply at its deterministic value before. Alternatively and more realistically, world market price volatility would be reduced by the introduction of stockholding activities. For further development, it is intended to calibrate a stockholding activity in order to match observed price volatility in international agricultural markets.

reforms along the lines of the recently agreed Health Check as well as a conclusion of the Doha Round along the lines of the EU offer made in 2005 (European Commission, 2005). In addition, the world market price development is calibrated to meet projections published by FAPRI for 2015 (FAPRI, 2006) before the recent peak of commodity prices.

3.2 Estimation of Stochastic Yield Terms

FAOSTAT time series data for the period 1962 to 2006 is used for estimating the distribution of the stochastic terms of the yield equations for all countries and regions depicted in the model. Also the correlation between error terms in yields of the considered crops and countries is estimated.

The de-trended stochastic variables θ are derived by dividing the observed yield y by the estimated yield $\hat{y}$ in the linear trend. In order to obtain the relative deviation (above or below) of the trend line we calculate:

$$\left(y / \hat{y}\right) - 1 .$$

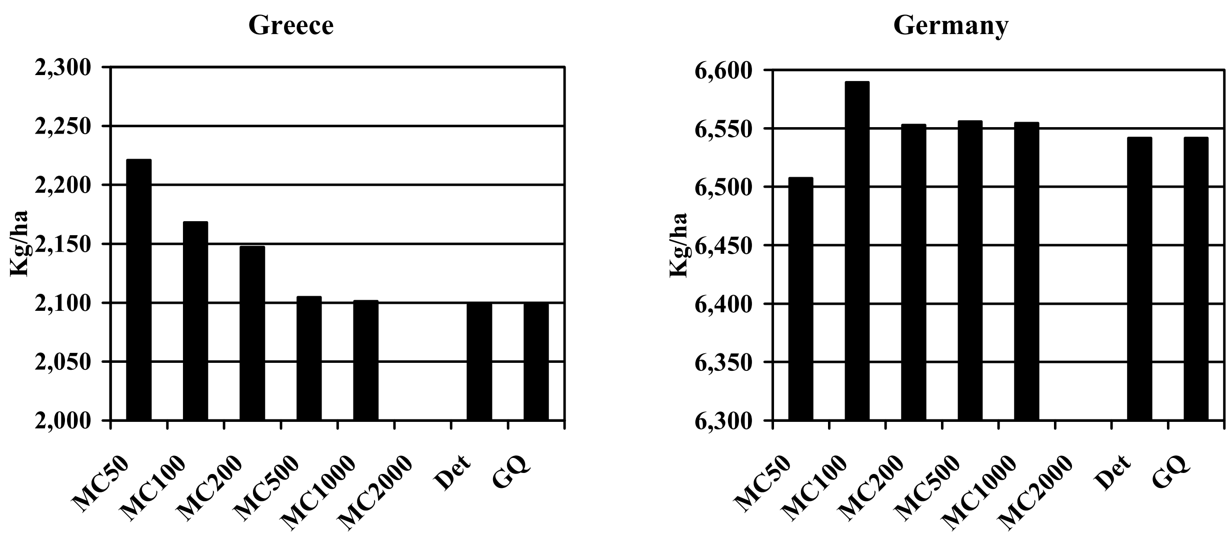
Some countries are grouped based on their correlations in order to reduce, if possible, the number of stochastic variables. Countries which produce rather small quantities are either grouped with large producers they have the best correlation with, or are ignored. The yield functions of those countries run deterministic without a stochastic term. Thus, in total we include 42 stochastic variables into the ESIM model. For more details on the grouping see the Appendix.

3.3 Results

Two types of results are presented. First, results based on GQ are compared to those based on MC simulations with different numbers of draws as well as to deterministic values in order to validate the GQ approach. Second, in order to illustrate the relevance of including yield stochasticity for model results, the EU net trade position for barley is discussed and the difference between a deterministic and a stochastic analysis of the effectiveness of an intervention price for barley is demonstrated.

Graphs 1 and 2 compare the expected values of barley yield in Greece and Germany under different stochastic approaches to the deterministic value.

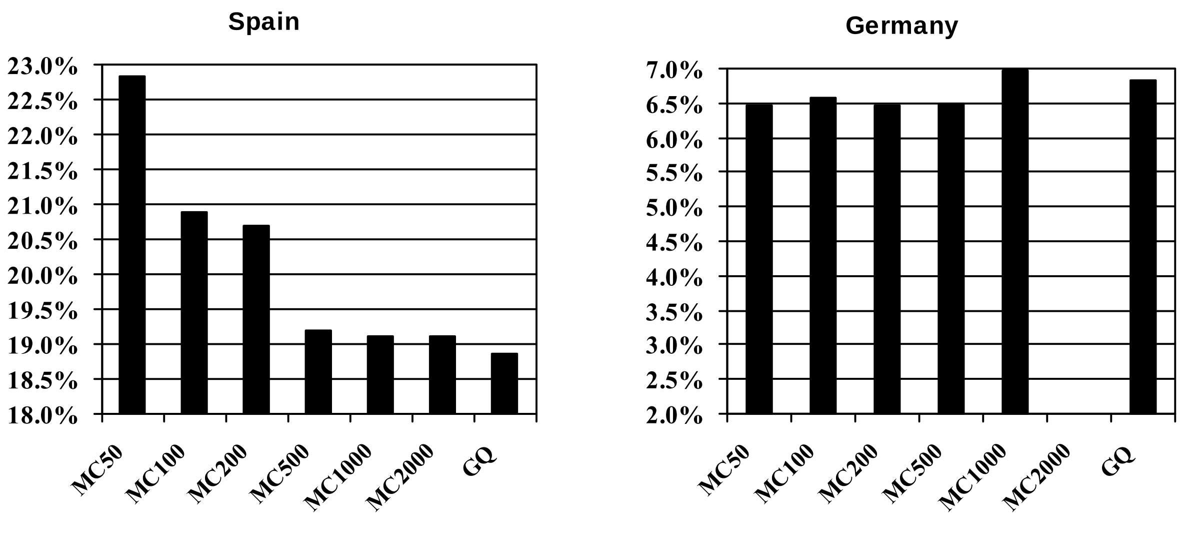
Graphs 1 and 2: Expected Value and Deterministic Solution for Barley Yield in Greece and Germany under Various Stochastic Specifications (kg/ha)



Naturally, a high number of MC draws should result in an approximation of the deterministic value, as the stochastic terms are normally distributed. Graphs 1 and 2 clearly show that a higher number of MC draws results in a better approximation of the deterministic value. For Greece, the expected value with 100 MC draws is still about 3% from the deterministic yield level. The expected values of the Gaussian Quadratures perfectly match the deterministic solution.

Graphs 3 and 4 compare the standard deviation of yield for Spain and Germany under different stochastic approaches.

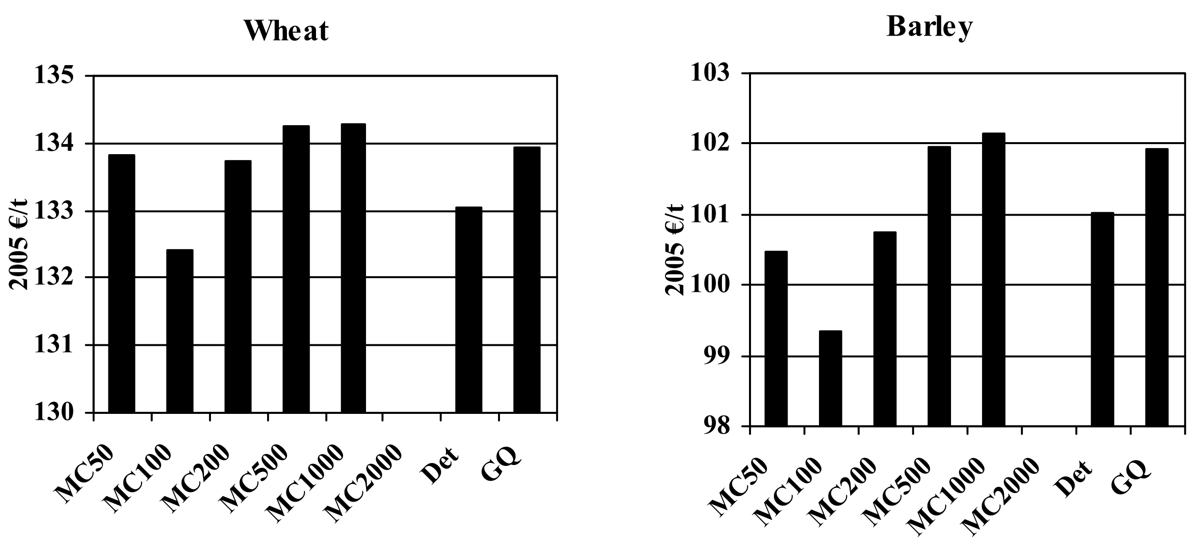
Graphs 3 and 4: Standard Deviation of Yield in % of Deterministic Solution, Spain and Germany



As a first observation, yield volatility is much higher in Spain than in Germany with the standard deviation being about 20% of the deterministic solution compared to about 7% in Germany. Naturally, a higher number of MC draws results in a better representation of yield volatility derived from historical data. Graphs 3 and 4 show that the standard deviation in the quadrature based simulation results are approached by the MC simulations with a high number of draws. But with a low number of MC draws, the standard deviation may substantially differ, e.g. by about 3 percentage points for Spain in case of 100 draws.

More interesting is the comparison of the moments of endogenous model variables under different stochastic specifications which are affected in a complex way by yield volatility. The distribution of these endogenous variables may be non-symmetric caused by the influence of, *i.e.*, tariff rate quotas, production quotas and intervention prices. Especially interesting in this respect are agricultural prices, which are strongly affected by yield volatility. Graphs 5 and 6 compare the expected values of EU prices for wheat and barley under different stochastic approaches to the deterministic value.

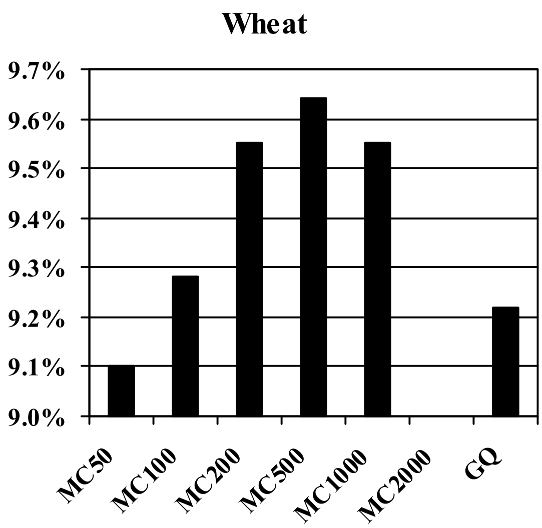
Graphs 5 and 6: Expected Value and Deterministic Solution for Wheat and Barley Price in the EU under Various Stochastic Specifications (2005 €/t)

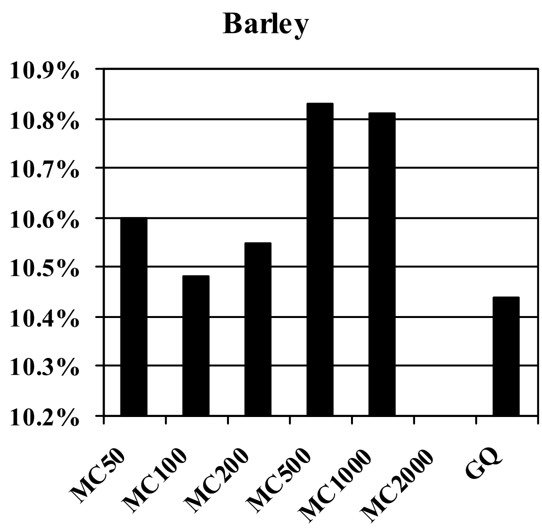


Graphs 5 and 6 show that a high number of MC draws as well as the quadrature based approach result in slightly higher expected values than the deterministic solution. In addition, expected values of the quadrature based approach are close to those based on 2000 MC draws, but save about 95% of computational capacity. Results based on a lower number of MC draws display a deviation in expected value of up to three per cent compared to the solution with 1000 draws.

Graphs 7 and 8 compare the standard deviation of EU prices for wheat and barley as a percentage of the deterministic solution under different stochastic approaches.

Graphs 7 and 8: Standard Deviation of the EU Price for Wheat and Barley in % of Deterministic Solution



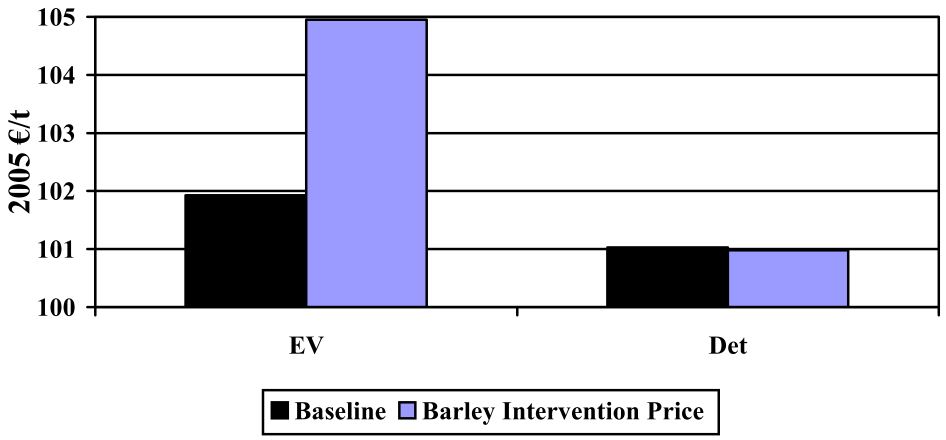


Although differences are small (0.6 percentage points at maximum), it is interesting to note that the standard deviation of price does not seem to move against a limit unambiguously with increasing number draws of up to 2000 as reported here. In addition, the standard deviation based on Gaussian Quadratures is not closer to those of MC simulations with many draws than to those with fewer draws. This may hint at the relationship between the distribution of yields and the distribution of prices being too complex to let the distribution of prices be captured by 2000 MC draws and/or by third order GQ approximating the distribution of yields.

Finally, the relevance of including yield volatility in a simulation model mainly aiming at analyzing EU policies may be documented, for example by looking at the net trade situation of the EU for barley. The deterministic solution suggests that the EU is a net exporter of barley in 2015 (net exports = 2.3 mill. t) and that the EU price level is equal to the world market price level at 101 €/t which is close to the former intervention price level of the EU for cereals (which is abolished in the baseline for barley). The stochastic solution shows that the EU could also be a net importer of barley in some years, *i.e.*, the standard deviation of net exports is at 3.5 mill. t[3] and thus exceeds the deterministic annual net trade value. In the deterministic version, the perpetuation of an intervention price would thus be without impact. However, in a stochastic version it should: especially in situations of low world market prices with net exports the intervention price would become binding. Thus, the expected value of the EU price for barley should be significantly above the deterministic value. To document this effect, ESIM is run with a counterfactual scenario which keeps the EU intervention price for barley at a level of 98 €/t. This scenario is compared to the baseline, where the intervention price is abolished. Results are presented in Graph 9.

[3] The standard deviation of supply relative to the deterministic solution differs between 6% in Ireland and 21% in Romania and Bulgaria. For the EU as a whole it is 6%. Due to inelastic demand, a high share of supply volatility translates into volatile net exports.

Graph 9: Deterministic Solution and Expected Value of the EU Barley Price in the Baseline and a Counterfactual with Intervention Price at 98 €/t



As expected, the effect of the intervention price on the deterministic solution is negligible, as the EU price is above intervention price level in that solution. But it strongly affects the expected value of the stochastic solution, as it keeps the EU price up in case of low world market prices.

4. CONCLUSIONS

We explore the options to use Gaussian quadratures in order to implement correlated stochastic terms in large scale simulation models and reduce the number of required solves compared to a Monte Carlo approach.

First, we use a theorem of Stroud (1957) as starting point for the generation of order three Gaussian quadratures to approximate integrals with multivariate independent normal distributions. The numerical method presented is straightforward and can be rapidly solved. Furthermore, we show that for stochastic modelling purposes there is no need to integrate over the *n*-cube as suggested in Stroud (1957) and followed by other authors, but integration can be over the *n*-dimensional Euclidean space $\Re^n$. Thus, even for dimensions $n \geq 4$ no rotation has to be applied. This makes the problem less complex and easier to imagine geometrically.

Second, we address the issue of inducing a desired covariance or correlation matrix to the generated Gaussian quadratures. We present the theory behind and show that there are many possible procedures to achieve this goal. Three standard procedures are presented. Especially, we show that the diagonalisation and the Cholesky decomposition have sometimes been mixed up in literature on correlated GQ.

Third, we apply the approach developed to a global partial equilibrium model of the agricultural sector, the European Simulation Model which is, to the best of our knowledge, the first application of correlated GQ in a large scale simulation model. In comparing GQ based to MC based results we demonstrate that the GQ approach saves substantially upon computational capacity and approximates the distribution of endogenous variables reasonably compared to a high number of MC draws. Our results hint at second moments being approximated less exactly than first moments of endogenous variables. Further systematic analysis of results as well as the exploration of higher order GQ are envisaged.

Finally, we demonstrate the relevance of stochastic modelling for a small example: the effect of an intervention price for barley in the EU. But stochastic simulation model may be relevant in many other applications, too, such as analyzing effects of increasing yield volatility due to climate change or effects of market integration on price volatility.

5. References

ARNDT, C. (1996), An Introduction to Systematic Sensitivity Analysis via Gaussian Quadrature. Purdue University, Center for Global Trade Analysis, GTAP Technical Paper, No. 2.

BANSE, M., GRETHE, H. and NOLTE, S. (2005): European Simulation Model (ESIM) in GAMS: Model Documentation. Model Documentation prepared for DG AGRI, European Commission, Göttingen and Berlin.

DEVUYST, E. A. (1993): Moment Preserving Approximations of Multivariate Probability Distributions for Decision and Policy Analysis: Theory and Applications. Ph.D. Thesis, Purdue University.

DEVUYST, E. A. and PRECKEL, P. V. (1997): Sensitivity Analysis Revisited: A Quadrature-Based Approach. *Journal of Policy Modeling,* 19(2): 175-185.

EUROPEAN COMMISSION (2005): Making Hong Kong a Success: Europe’s Contribution. Brussels, 28 October.

FAPRI (2006), U.S. and World Agricultural Outlook. Food and Agricultural Policy Research Institute. Iowa State University, University of Missouri-Columbia, Ames, Iowa.

HERTEL, T. W., REIMER, J. J., and VALENZUELA, E. (2005): Incorporating commodity stockholding into a general equilibrium model of the global economy. *Economic Modelling,* 22: 646-664.

IMAN, R. and CONOVER, W. (1982): A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics - Simulation and Computation*, 11(3): 311-334.

OECD (2003): OECD Agricultural Outlook 2003-2008. Paris.

PRECKEL, P. V. and DEVUYST, E. A. (1992): Efficient Handling of Probability Information for Decision Analysis under Risk. *American Journal of Agricultural Economics*, 74(3): 655-662.

RICHARDSON, J., KLOSE, S. and GRAY, A. (2000): An Applied Procedure for Estimating and Simulating Multivariate Empirical (MVE) Probability Distributions in Farm-Level Risk Assessment and Policy Analysis. *Journal of Agricultural and Applied Economics,* 32 (2): 299-315.

STROUD, A. H. (1957), Remarks on the Disposition of Points in Numerical Integration Formulas. *Mathematical Tables and Other Aids to Computation,* 60(11): 257-261.

VERHOOG et al. (2008): Potentials of a harmonised Database for Agricultural Market Modelling. JRC Scientific and Technical Reports, IPTS, Seville.

WESTHOFF, P., BROWN, S., and HART, C. (2005): When Point Estimates Miss the Point: Stochastic Modeling of WTO Restrictions. FAPRI Policy Working Paper #01-05.

APPENDIX

Groupings of countries with identical stochastic terms

	Wheat	Barley	Rapeseed	Deterministic		
				Wheat	**Barley**	**Rapeseed**
1	Germany	Germany	Germany	Cyprus	Cyprus	Cyprus
2	France	France	France	Malta	Malta	Malta
3	UK+Ireland	UK+Ireland	UK+Ireland	Slovenia	Slovenia	Slovenia
4	Spain+Portugal	Spain+Portugal	Poland	Baltic States	Baltic States	Baltic States
5	Italy	Denmark+Sweden	Czech Republic+Slovakia			Belgium
6	Poland	Poland	Denmark+Sweden+ Finland			Luxembourg
7	Czech Republic+Slovakia	Czech Republic+Slovakia	Hungary			Netherlands
8	Romania+Bulgaria	Romania+Bulgaria	Austria			Spain
9	Hungary	Finland	ROW			Portugal
10	Denmark+Sweden+ Finland	Austria				Italy
11	Greece	Hungary				Romania
12	Netherlands+Belgium+ Luxembourg	Italy				Bulgaria
13	Austria	Netherlands+ Belgium+Luxembourg				Greece
14	Turkey	Greece				Turkey
15	US	Turkey				US
16	ROW	US				
17		ROW				