AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

SCHRIFTEN DER GESELLSCHAFT FÜR WIRTSCHAFTS- UND SOZIALWISSENSCHAFTEN DES LANDBAUES E.V.

Liu, X., Xu, W., Odening, M.: Lassen sich Ertragsrisiken in der Landwirtschaft global diversifizieren? In: Weingarten, P., Banse, M., Gömann, H., Isermeyer, F., Nieberg, H., Offermann, F., Wendt, H.: Möglichkeiten und Grenzen der wissenschaftlichen Politikanalyse. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., Band 46, Münster-Hiltrup: Landwirtschaftsverlag (2011), S. 245-258.

Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., Bd. 46, 2011: 245 – 258

LASSEN SICH ERTRAGSRISIKEN IN DER LANDWIRTSCHAFT GLOBAL DIVERSIFIZIEREN?

Xiaoliang Liu, Wei Xu und Martin Odening[1]

Zusammenfassung

In dieser Arbeit wird die Stochastizität landwirtschaftlicher Ernteerträge in wichtigen Erzeugerregionen der Welt am Beispiel von Winterweizen untersucht. Besondere Aufmerksamkeit wird der stochastischen Abhängigkeit der Erträge in den verschiedenen Regionen gewidmet. Damit verbindet sich die Frage, ob und in welchem Maße Ertragsschwankungen durch globalen Handel ausgeglichen werden können. Während statistische Zusammenhangsanalysen üblicherweise auf linearen Korrelationen basieren, werden in diesem Beitrag Copulas verwendet. Im Vergleich zu linearen Korrelationen ist die Anwendung von Copulas an weniger restriktive Voraussetzungen gebunden. Insbesondere das gemeinsame Auftreten extremer zufälliger Ereignisse lässt sich mit Copulas genauer modellieren. Unsere Berechnungen zeigen, dass eine Angebotskonstellation, wie sie in 2007 aufgetreten ist, kein Jahrhundertereignis darstellt, sondern sich c.p. etwa alle 15 Jahre wiederholen kann.

Keywords

Ertragsrisiken, Copulas, Diversifikation

1 Einleitung

Es gilt als weithin akzeptiert, dass die weltweite Liberalisierung der Agrarmärkte und der damit verbundene Wegfall preisstabilisierender Instrumente (variable Zölle, Quoten, Exporterstattungen) zu einer Erhöhung der Preisschwankungen auf diesen Märkten führt, zumindest aus Sicht der Länder, in denen solche Maßnahmen bislang zur Anwendung kamen (ARTAVIA ET AL. 2010). Das konkrete Ausmaß der in der jüngeren Vergangenheit zu beobachtenden Preisausschläge, insbesondere der drastische Preisanstieg für pflanzliche Produkte in 2007 und zu Beginn des Jahres 2008, erscheint für Agrarökonomen indes überraschend oder zumindest bemerkenswert. So hat beispielsweise die Gesellschaft für Wirtschaft- und Sozialwissenschaften des Landbaues ihre 49. Jahrestagung 2009 unter das Leitthema „Die Agrar- und Ernährungsmärkte nach dem Boom“ gestellt. Der Preisboom für landwirtschaftliche Produkte in 2007/2008 hat mehrere Aspekte: Zum einen hat er den Optimismus der in der Branche Tätigen beflügelt. Vom Ende der landwirtschaftlichen Tretmühle war die Rede (VON WITZKE 2008), und Landwirte wurden als künftige Scheichs gesehen (z.B. RETTBERG 2005). Auf der anderen Seite hat sich infolge hoher Nahrungsmittelpreise die Hungerproblematik in Niedrigeinkommensländern zugespitzt. Aus wissenschaftlicher Sicht stellen sich in diesem Zusammenhang folgende Fragen: Erstens, welche Faktoren waren für den starken Preisanstieg verantwortlich? Zweitens, wie wahrscheinlich es ist, dass sich eine solche Konstellation preistreibender Faktoren wiederholt und drittens, wie lassen sich derartige Preisausschläge im Bedarfsfall dämpfen?

Als mögliche Gründe für den Preisboom wurden unter anderem Mindererträge in wichtigen Erzeugerregionen der Welt, die erhöhte Nachfrage nach Biokraftstoffen sowie das spekulative Engagement von Investoren auf Märkten für Agrarrohstoffe genannt (cf. SARRIS 2009). In diesem Beitrag soll einer der preisbestimmenden Faktoren auf der Angebotsseite näher betrachtet werden und zwar das Ertragsrisiko. Es ist bekannt, dass landwirtschaftliche Erträge, insbe-

[1] Department für Agrarökonomie, Humboldt-Universität zu Berlin, Philippstr. 13 – Haus 12A, 10115 Berlin, E-Mail: m.odening@agrar.hu-berlin.de

sondere in der Pflanzenproduktion in starkem Maße witterungsabhängig sind und Wetterrisiken sich mehr oder weniger direkt in Ertragsschwankungen niederschlagen (ODENING ET AL. 2008). Wenn es darum geht, den Einfluss von Ertragsschwankungen auf Preisschwankungen auf deregulierten globalen Märkten zu bestimmen, ist die Wahl der Perspektive wichtig. Eine regional begrenzte Analyse greift in diesem Fall zu kurz, denn Mindererträge in einer Region können möglicherweise durch Importe aus anderen Regionen kompensiert werden (LOTZE-CAMPEN 2007). Mit anderen Worten: In integrierten Märkten muss Diversifikation durch Handel berücksichtigt werden, woraus die Notwendigkeit einer globalen Analyse von Ertragsschwankungen folgt. Vor diesem Hintergrund wird in der vorliegenden Arbeit die Stochastizität landwirtschaftlicher Ernteerträge in wichtigen Erzeugerregionen der Welt untersucht. Die Analyse erfolgt beispielhaft für das Produkt Weizen. Besondere Aufmerksamkeit wird der stochastischen Abhängigkeit der Erträge in den verschiedenen Regionen gewidmet. Während statistische Zusammenhangsanalysen üblicherweise auf linearen Korrelationen basieren, werden in diesem Beitrag Copulas verwendet. Im Vergleich zu Korrelationen ist die Anwendung von Copulas an weniger restriktive Voraussetzungen gebunden. Insbesondere das gemeinsame Auftreten extremer zufälliger Ereignisse lässt sich mit Copulas genauer modellieren als mit linearen Korrelationen. Durch die Anwendung dieser in der Agrarökonomie noch nicht sonderlich verbreiteten Technik leistet die vorliegende Arbeit einen methodischen Beitrag zur Risikomessung in diesem Bereich. Im Folgenden werden zunächst die statistischen Grundlagen von Copulas erläutert. Anschließend wird diese Technik herangezogen, um die Wahrscheinlichkeit weltweiter Ertragsausfälle für Weizen zu schätzen.

2 Modellierung stochastischer Abhängigkeiten mittels Copulas

2.1 Theoretischer Hintergrund

Die wahrscheinlichkeitstheoretische Darstellung mehrdimensionaler Risiken erfolgt mit Hilfe von multivariaten Verteilungsfunktionen $F(x_1, x_2, \cdots, x_d)$. Darin bezeichnet x_i die Realisation einer Zufallsvariable X_i, $i = 1, \cdots, d$. Eine direkte empirische Schätzung von F scheitert in der Regel an der ausreichenden Verfügbarkeit von Daten. Stattdessen werden zur Bestimmung von F üblicherweise zum einen die eindimensionalen Randverteilungen $F_i(x_i)$ und zum anderen die (linearen) Korrelationen zwischen den einzelnen Zufallsvariablen geschätzt. Durch die Verknüpfung der Randverteilungen mit den Korrelationen wird dann die gemeinsame Verteilung abgeleitet. Dieses Vorgehen ist allerdings nur dann korrekt, wenn die gemeinsame Verteilung F der Klasse der elliptischen Verteilungen angehört, zu denen die multivariate Normalverteilung zählt (vgl. EMBRECHTS ET AL. 2002). Ist F nicht multivariat normal, enthalten lineare Korrelationskoeffizienten nicht die gesamte Information über die Abhängigkeitsstruktur der Einzelrisiken. Dieses Defizit macht sich besonders in den Rändern der Verteilung bemerkbar, also bei der Quantifizierung der Wahrscheinlichkeit des gleichzeitigen Auftretens extremer Ereignisse, die als "tail dependence" bezeichnet wird. XU ET AL (2010) zeigen am Beispiel von Wetterrisiken, wie stark die Fehler bei der Verwendung linearer Korrelationen sein können, wenn die Anwendungsvoraussetzungen nicht erfüllt sind. Die Nutzung von Copulas stellt nun einen Kompromiss zwischen der praktisch unmöglichen direkten Schätzung der gemeinsamen Verteilung F und der Verwendung linearer Korrelationen dar.

Die theoretische Begründung zur Verwendung von Copulas leitet sich aus SKLAR's Theorem (SKLAR 1959) ab, demzufolge jede d-dimensionale Verteilungsfunktion $F(x_1, x_2, \cdots, x_d)$ mit Hilfe einer Copula-Funktion $C(\cdot)$ dargestellt werden kann:

$$(1) \quad F(x_1, \ldots, x_d) = C\{F_1(x_1), \ldots, F_d(x_d)\}, \quad \forall\, x_1, \ldots, x_d \in \overline{R}.$$

Darin sind $F_i(x_i)$ (eindimensionale) Randverteilungen, und $u_i = F_i(x_i) \to [0,1]$ ist folglich uniform verteilt. Die Copula lässt sich darstellen als (vgl JOE 1997):

$$(2) \quad C(u_1,\ldots,u_d) = F\{F_1^{-1}(u_1),\ldots,F_d^{-1}(u_d)\}, \quad F_i^{-1}(u_i) = x_i.$$

Gleichung (1) bringt zum Ausdruck, dass sich die in der gemeinsamen Verteilungsfunktion enthaltene Information in zwei Komponenten zerlegen lässt: erstens die Randverteilungen und zweitens die Abhängigkeitsstruktur, die durch die Copula abgebildet wird. Die Dichtfunktion einer d-dimensionalen Copula ergibt sich als d-te Ableitung aus Gleichung (2)

$$(3) \quad c(u_1,\ldots,u_d) = \frac{\partial^d C(u_1,\ldots,u_d)}{\partial u_1 \cdots \partial u_d}, \quad \forall\, u_1,\ldots,u_d \in [0,1].$$

Copulas können parametrisch oder nicht-parametrisch (z.B. als Kernel-Copulas) modelliert werden (CHEN und HUANG 2007). Beide Methoden haben Vor- und Nachteile. Bei Vorliegen einer geringen Zahl von Beobachtungen kann die Anwendung von nicht-parametrische Copula vorteilhaft sein, weil parametrische Methoden in diesem Fall zu großen Schätzfehlern führen können (VEDENOV 2008). Andererseits sind parametrische Modelle für bestimmte Anwendung besser geeignet, etwa wenn der Randbereich der Verteilung genau beschrieben werden soll (CHARPENTIER, FERMANIAN und SCAILLET 2007). Im Folgenden wird ein parametrischer Ansatz verfolgt, da die Wahrscheinlichkeit gleichzeitiger extremer Ernteausfälle in verschiedenen Regionen im Mittelpunkt der Betrachtung steht.

2.2 Typen und Eigenschaften von parametrischen Copulas

Man unterscheidet zwei Klassen von Copulas und zwar elliptische und archimedische Copula. Im Folgenden beschränken wir uns auf archimedische Copulas, die sich in geschlossener Form darstellen lassen. Dies ist vorteilhaft wenn Copulas für Simulationen genutzt werden sollen. Die allgemeine Schreibweise für archimedische Copulas lautet:

$$(4) \quad C(u_1,\ldots,u_d) = \varphi^{-1}(\varphi(u_1) + \cdots + \varphi(u_d)),$$

wobei $\varphi(\cdot)$ die sogenannte Erzeuger-Funktion ist und $\varphi^{-1}(\cdot)$ deren Inverse.

Zu der Klasse der archimedischen Copulas gehören drei wichtige Typen: die Gumbel-Copula, die Clayton-Copula und Frank-Copula. Das Gumbel-Copula ist wie folgt definiert

$$(5) \quad C_{Gu}(u_1,\ldots,u_d,\theta) = \varphi^{-1}\{\varphi(u_1) + \ldots + \varphi(u_d)\} = \exp\left[-\left\{(-\ln u_1)^\theta + \ldots + (-\ln u_d)^\theta\right\}^{\frac{1}{\theta}}\right], 1 \le \theta \le \infty,$$

Darin ist θ ein zu schätzender Parameter der Copula. Die Gumbel-Copula ist asymmetrisch und modelliert eine starke Abhängigkeit für Werte in den jeweils oberen Bereichen der Randverteilungen. Man spricht von „upper tail dependence“ (OKHRIN 2007). Die Clayton-Copula ist ebenfalls eine asymmetrische Archimedische Copula:

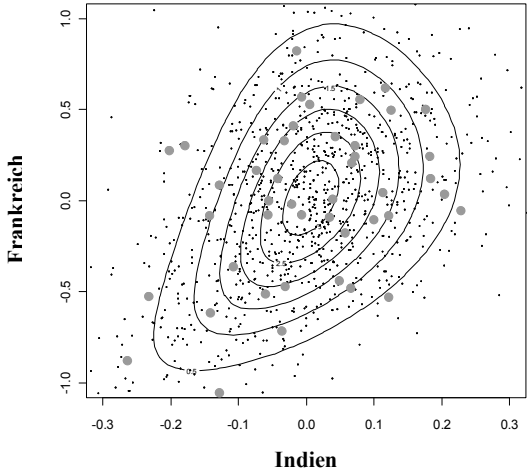
$$(6) \quad C_C(u_1,\ldots,u_d,\theta) = \left\{u_1^{-\theta} + \ldots + u_d^{-\theta} - (d-1)\right\}^{-\frac{1}{\theta}}, -1 \le \theta \le \infty, \theta \ne 0.$$

Im Gegensatz zu der Gumbel Copula eignet sich die Clayton-Copula zur Modellierung von Abhängigkeiten im unteren Bereich der Verteilungen, etwa Verlusten (lower tail dependence). Anders als die Gumbel- und die Clayton-Copula implziert die Frank-Copula Unabhängigkeit in den Rändern der Verteilungen, und sie ist für die Modellierung symmetrischer Abhängigkeiten geeignet. Eine Frank-Copula ist definiert als

$$(7) \quad C_F(u_1,\ldots,u_d,\theta) = -\frac{1}{\theta}\ln\left[1 + \frac{\prod_{i=1}^{d}(e^{-\theta u_i} - 1)}{(e^{-\theta} - 1)^{d-1}}\right]$$

Zur Veranschaulichung sind in Abbildung 1 im Vorgriff auf die nachfolgende empirische Anwendung die Konturlinien bivariater Verteilungen des stochastischen Teils der Weizenerträge in Indien und Frankreich abgebildet. Den Verteilungen liegt eine Clayton-Copula zugrunde.

Abbildung 1: Konturlinien der gemeinsamen Verteilung des Störterms von Weizenerträgen (dt/ha) in Indien und Frankreich. Copulatyp: Clayton; rot = beobachtet, schwarz = simuliert



2.3 Hierarchische Konstruktionen für mehrdimensionale archimedische Copulas

Die im vorherigen Abschnitt beschriebene Vorgehensweise stellt die einfachste Form zur Modellierung mehrdimensionaler Copulas dar. Zum einen wird die gesamte multivariate Abhängigkeitsstruktur mit nur einem Copulaparameter θ ausgedrückt. Zum anderem bleibt die Sub-Struktur der stochastischen Abhängigkeiten unsichtbar. Darüber hinaus impliziert diese Methode, dass die einzelnen Argumente u_i innerhalb der mehrdimensionalen Copula-Funktion austauschbar sind, also z.B. für eine dreidimensionale Copula $C(u_1,u_2,u_3)=C(u_3,u_1,u_2)$ gilt. Diese Voraussetzung ist nicht für alle Anwendungen plausibel (vgl. NELSEN 2005). Eine flexiblere Modellierung mehrdimensionaler Copulas ist mit Hilfe sogenannter „Fully Nested-Archimedian-Copulas“ (FNAC) möglich, die auf einer wiederholten Verknüpfung bivariater Copulas basiert (WHELAN 2004). Die Struktur einer d-dimensionalen archimedischen Copula kann dann ausgedruckt werden als:

$$C(u_1,\ldots,u_d)=\varphi_{d-1}^{-1}\{\ \varphi_{d-1}\circ\varphi_{d-2}^{-1}[\ldots(\varphi_2\circ\varphi_1^{-1}[\varphi_1(u_1)+\varphi_1(u_2)]+\varphi_2(u_3))+\ldots+\varphi_{d-2}(u_{d-1})]+\varphi_{d-1}(u_d)\ \} \tag{8}$$

Darin sind $\varphi_1,\ldots,\varphi_{d-1}$ Erzeuger-Funktion und „$\circ$“ steht für den Kompositionsoperator. Eine FNAC ist charakterisiert durch insgesamt $d(d-1)/2$ mögliche Kombinationen verschiedener bivariater Verteilungen $(F_i(x_i),F_j(x_j)), i\neq j$, aus den insgesamt $d-1$ Copulas mit entsprechenden Parametern geschätzt werden. Dabei ist die Substruktur der Abhängigkeitsbeziehungen hierarchisch aufgebaut (siehe Abb. 2 in Abschnitt 3.2) Angefangen mit der niedrigsten

Ebene werden zunächst zwei Variablen zu einer bivariaten Copula kombiniert. Dabei werden zunächst diejenigen Variablen zusammengefasst, die den stärksten Zusammenhang aufweisen, z.B. gemessen durch Kendall's Tau. Bezogen auf die in Abb. 2 dargestellte Situation ergibt sich:

(9) $$z_1 = C_1(u_5, u_6) = \varphi_1^{-1}[\varphi_1(u_5) + \varphi_1(u_6)]$$

Auf der Zweiten Ebene wird anschließend z_1 mit u_2 zu $C_2(z_1, u_2)$ kombiniert.

(10) $$C_2(z_1, u_2) = \varphi_2 \circ \varphi_1^{-1}[\varphi_1(u_5) + \varphi_1(u_6)] + \varphi_2(u_2) = \varphi_2^{-1}[\varphi_2(z_1) + \varphi_2(u_2)]$$

Dieser Prozess wird bis zum höchsten Ebene fortgesetzt, und im Ergebnis liegt eine *d*-dimensionale Copula vor. In der Regel wird derselbe Copulatyp für alle bivariaten Kombinationen verwendet, so dass sich Unterschiede nur in den Copulaparametern ausdrücken. Man kann zeigen, dass der Wert der geschätzten Copulaparameter von der niedrigsten Ebene zu höchsten Ebene monoton abnehmen muss, d.h. $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_{d-1}$ (JOE 1997, EMBRECHTS ET AL. 2003). Durch diese Bedingung ist sichergestellt, dass die gemeinsame Verteilung, die sich aus der Kombination zweier Copulafunktionen ergibt, wiederum eine Copulafunktion ist.

2.4 Schätzen und Testen von Copulas

Für die Schätzung der Copulaparameter stehen drei Methoden zur Verfügung (vgl. CHERUBINI, ET AL. 2004). Die exakte Maximum Likelihood Methode schätzt die Copula-Parameter θ und die Parameter der marginalen Verteilungen $\alpha_1, \cdots, \alpha_d$ simultan. Alternativ kann die Schätzung in zwei Schritten vollzogen werden, d.h., zunächst werden die Parameter für die marginalen Verteilungen geschätzt und für die nachfolgende Schätzung der Copula-Parameter vorgegeben. Diese als "the Inference for Margin Methods" bezeichnete Methode ist im Vergleich zur simultanen Schätzung weniger effizient, zeichnet sich aber durch einen vergleichsweise geringen Rechenaufwand aus. Eine Variante stellt die sogenannte „Canonical Maximum-Likelihood (CML)" Methode dar (HAERDLE ET AL. 2008). Die Log-Likelihood-Funktion lautet für diesen Fall.

(11) $$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \sum_{j=1}^{k} \ln[c\{\hat{F}_1(x_{1j}), \ldots, \hat{F}_d(x_{dj}); \theta\}]$$

$\hat{F}(x_{ij})$ ist die i-te empirische Randverteilung basierend auf den Beobachtungswerten x_{ij}, $j = 1, \cdots, k$. $\hat{\theta}$ wird auch als Maximum-Pseudolikelihood-Schätzer oder rangbasierter Maximum-Likelihood-Schätzer bezeichnet. Im Unterschied zu den beiden zuvor genannten Methoden werden bei der CML-Methode an Stelle parametrischer marginalen Verteilungen die empirischen Randverteilungen verwendet. Dies bietet sich an, wenn nicht genügend empirischen Daten zur Verfügung stehen, um eine parametrische Verteilung eindeutig zu bestimmen und zuverlässig zu schätzen. Die CML-Methode wird für die Schätzung der Copulaparameter in dieser Untersuchung verwendet.

Gleichzeitig muss der den Daten angemessene Copulatyp bestimmt werden. Analog zur Schätzung univariater Verteilugen können hierbei Goodness-of-Fit Tests (GoF) zur Anwendung kommen. Das Problem gestaltet sich allerdings schwieriger, da es hier um den Vergleich mehrdimensionaler Verteilungen geht. GENEST und REMILLARD (2008) schlagen in diesem Zusammenhang einen Test vor, der zu der Klasse der Dimensionsreduktionsverfahren gehört und auf einer Cramer-von-Mises Statistik basiert. Sei C die wahre *d*-dimensionale Copula für die zu analysierenden Daten und C_θ der zu testende Copulatyp, dann können die zu testenden Hypothesen wie folgt formuliert werden:

(12) $H_0 : C \in \boldsymbol{C} = \{C_\theta ; \theta \in \Theta\}$ vs. $H_1 : C \notin \boldsymbol{C} = \{C_\theta ; \theta \in \Theta\}$

Darin bezeichnet Θ den Parameterraum. In dem Test wird die Distanz zwischen der empirischen Copula und der angenommenen parametrischen Copula gemessen und für den Signifikanz-Test verwendet. Die Verteilung der Teststatistik besitzt keine geschlossene Form und muss mit Hilfe eines Bootstrap-Verfahrens ermittelt werden. Eine ausführliche Beschreibung dieses Tests, auf den sich die Analyse in dieser Arbeit stützt, findet sich bei GENEST und REMILLARD (2008).

2.5 Copulabasierte Simulation simultaner Risiken

Liegen Schätzungen für die Randverteilungen und die Copula vor, dann können mittels Monte-Carlo-Simulation Realisationen von d-dimensionalen Zufallsvektoren generiert werden, die der gemeinsamen Verteilung $F(x_1, \cdots, x_d)$ folgen. Speziell für die Simulation mehrdimensionaler Copulas wurde die sogenannte „conditional inverse method" entwickelt (HAERDLE und OKHRIN 2009). Die Grundidee diese Methode ist, mehrdimensionale Zufallsvariablen rekursiv aus der bedingten Verteilung zu simulieren. Zuerst werden d-dimensionale unabhängige Zufallsvariable $v_1, \ldots, v_d$ aus einer Gleichverteilung $U \in [0,1]$ gezogen. Danach wird muss eine d-dimensionale Gleichverteilung $u_1, \ldots, u_d$ generiert werden. Im Gegensatz zu den $v_1, \ldots, v_d$ sollen die $u_1, \ldots, u_d$ stochastisch abhängig sein, wobei die Abhängigkeitsstruktur durch die Copula $C(u_1, \cdots, u_d)$ beschrieben wird. Um dies zu erreichen, wird für den Anfangswert $u_1 = v_1$ gesetzt. Die weiteren Werte $u_2, \ldots, u_d$ werden bestimmt durch.

(13) $u_i = \Lambda^{-1}(v_i),\ i = 2, \cdots, d$

mit

(14) $$\Lambda(u_i | u_1, \cdots, u_{i-1}) = P((U_i \le u_i | U_1 = u_1, \ldots, U_{i-1} = u_{i-1})) = \frac{c_{i-1}^{i}(u_1, \cdots, u_i)}{c_{i-1}^{i-1}(u_1, \cdots, u_{i-1})}$$

$c_{i-1}^{i}(u_1, \cdots, u_i)$ und $c_{i-1}^{i-1}(u_1, \cdots, u_{i-1})$ in der Gleichung (14) sind die Dichtefunktion der i bzw. $i-1$ dimensionalen Copula, die analog zu (3) definiert sind als:

(15) $$c_{i-1}^{i}(u_1, \cdots, u_i) = \frac{\partial^{i-1} C_i(u_1, \cdots, u_i)}{\partial u_1, \cdots, \partial u_{i-1}}$$

und

(16) $$c_{i-1}^{i-1}(u_1, \cdots, u_i) = \frac{\partial^{i-1} C_{i-1}(u_1, \cdots, u_{i-1})}{\partial u_1, \cdots, \partial u_{i-1}}$$

Im letzten Schritt werden die für jede Zufallsvariable X_i Realisationen x_i gemäß

(17) $x_i = F_i^{-1}(u_i),\ \ i = 1, \cdots, d$

erzeugt, wobei $F_i^{-1}(u_i)$ die Inverse der Randverteilung $F_i(x_i)$ bezeichnet. Die so simulierten Realisationen $x_1, \ldots, x_d$ folgen der gewünschten gemeinsamen Verteilung $F(x_1, \cdots, x_d)$.

3 Empirische Analyse globaler Ertragsausfallrisiken

3.1 Vorgehensweise und Datengrundlage

Die empirische Analyse basiert auf den Weizenproduktionsmengen der weltweit größten Erzeugerländer. Um den Rechenaufwand zu beschränken, erfolgt eine Eingrenzung auf die acht wichtigsten Nationen. Dies sind China, Deutschland, Frankreich, Kanada, Indien, Pakistan,

die Türkei und USA[2]. Das Produktionsvolumen dieser Länder machte in 2007 58 Prozent der Welterzeugung aus. Für diese Länder liegen Ertragszeitreihen der FAO für die Jahre zwischen 1961 und 2007, also 47 Beobachtungen, vor (http://faostat.fao.org).

Bei der Analyse wurde wie folgt vorgegangen: Zunächst wurden die Gesamterträge auf die Erntefläche des jeweiligen Jahres bezogen, also in Hektarerträge umgerechnet. Dadurch lässt sich die Produktion in unterschiedlich großen Ländern besser vergleichen, und die jährlichen Schwankungen der Ernteflächen werden nicht in die Analyse übertragen[3]. Anschließend erfolgt eine lineare Trendbereinigung, um technische Fortschritte zu eliminieren. Die nachfolgende statistische Analyse bezieht sich also auf die Abweichungen der Hektarerträge von ihrem linearen Trend. Ein Teil dieser Daten ist aufsteigend sortiert in Tabelle A1 im Anhang dargestellt. Es wird deutlich, dass das Erntejahr 2007 in 4 der betrachteten acht Länder tatsächlich zu den drei schlechtesten Jahren seit 1961 gehört. Anderseits ist auch zu erkennen, dass dies nicht für alle Länder zutrifft; so war in 2007 beispielsweise in Pakistan eine vergleichsweise gute Ernte zu verzeichnen. Weiterhin vermitteln die Zahlen einen Eindruck über die Höhe der Ertragsschwankungen und deren Unterschiede in den einzelnen Untersuchungsländern.

Basierend auf den in Tabelle A1 enthaltenen Daten wird für jedes Land eine univariate Ertragsverteilung spezifiziert, geschätzt und mittels χ^2-Test auf Signifikanz geprüft. Für die acht Zeitreihen der Ertragsresiduen wird eine hierarchische archimedische Copula (FNAC) geschätzt. Die Auswahl des Copulatyps und die Festlegung der hierarchischen Copulastruktur basieren auf der oben beschriebenen Cramer-von Mises Statistik. Die CML Methode wird verwendet, um die Copulaparameter zu schätzen. Aus dieser Copula werden mit dem in Abschnitt 2.5 beschrieben Verfahren achtdimensionale Zufallsvektoren generiert. Zusammen mit der Information aus den Randverteilungen werden so Ertragsschwankungen simuliert, die der gewünschten gemeinsamen Verteilung folgen. Schließlich werden die deterministischen Werte der Trendfunktionen für das letzte Beobachtungsjahr (2007) wieder addiert und die simulierten hektarbezogenen Weizenerträge der einzelnen Länder nach Maßgabe der relativen Produktionsumfänge aggregiert. Durch Multiplikation mit den Anbauumfängen ergibt sich die Gesamterzeugung. Nach 1000 maliger Wiederholung ergibt sich eine Verteilungsfunktion für die Hektarerträge bzw. die Gesamtproduktion der betrachteten Erzeugerländer, für die sich beliebige Kennzahlen, insbesondere Quantile, bestimmen lassen.

3.2 Ergebnisse

Die Ergebnisse der Anpassungstests für die Randverteilungen der Ertragsschwankungen sind in Tabelle A2 im Anhang enthalten. Die Ertragsschwankungen in den einzelnen Ländern werden am besten durch unterschiedliche Verteilungen modelliert und zwar durch die Weibull-Verteilung, die Normalverteilung und die logistische Verteilung. Dieser Befund steht in Einklang mit vergleichbaren empirischen Untersuchungen, z.B. von MORIONDO ET. AL. (2009) oder UPADHYAY und SMITH (2005). Das bedeutet gleichzeitig, dass die gemeinsame Verteilungsfunktion nicht multivariat normal sein kann und somit die Voraussetzung zur korrekten Anwendung linearer Korrelationskoeffizienten nicht gegeben ist.

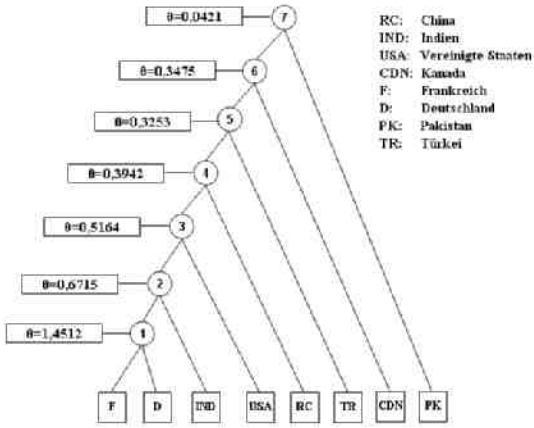
Mit Hilfe des in Abschnitt 2.4 beschriebenen GoF wurde untersucht, welcher archimedische Copulatyp am besten geeignet ist. Als mögliche Kandidaten standen die Clayton, die Gumbel und die Franck Copula zur Auswahl, die jeweils unterschiedliche Abhängigkeitsstrukturen repräsentieren. Der Test wurde für alle 28 Kombinationsmöglichkeiten für Länderpaarungen durchgeführt. In der Mehrzahl der Fälle weist die Clayton Copula die beste Anpassung an die

[2] Russland zählt zwar zu den größten Erzeugerländern, musste aber aus der Analyse ausgeschlossen werden, da keine ausreichend lange Ertragszeitreihe vorlag.

[3] Die Weizenernteflächen der untersuchten Länder für den Betrachtungszeitraum sind im Anhang A1 enthalten.

Daten auf. Dieser Typ wird daher für die gesamte FNAC verwendet. Auch dieses Ergebnis unterstreicht die Unzulässigkeit der Verwendung linearer Korrelationen. Nach der Spezifikation des Copulatyps muss die Struktur der hierarchischen Copula bestimmt werden. Die Zusammenfassung der Regionen erfolgt dabei nach Maßgabe der Test-Statistik des GoF-Tests. Abbildung 2 zeigt das Ergebnis dieser Prozedur. Die ausgewiesenen Schätzwerte der Copulaparameter nehmen – wie theoretisch gefordert - von der niedrigsten Ebene zu höchsten ab, mit Ausnahme der Ebene 6. Diese Abweichung kann auf die Länge der empirischen Datenreihe zurückgeführt werden, die mit 47 Beobachtungen recht kurz ist. Die Ergebnisse zeigen wenig überraschend, dass die Weizenerträge zwischen räumlich nahe gelegenen Ländern wie Frankreich und Deutschland eine relativ hohe stochastische Abhängigkeit aufweisen. Überraschend ist dagegen, dass die Abhängigkeit des Weizenertrags zwischen Indien einerseits sowie Frankreich und Deutschland andererseits höher ist, als zwischen Indien und Pakistan.

Abbildung 2: Hierarchische Struktur der Archimedischen Copula



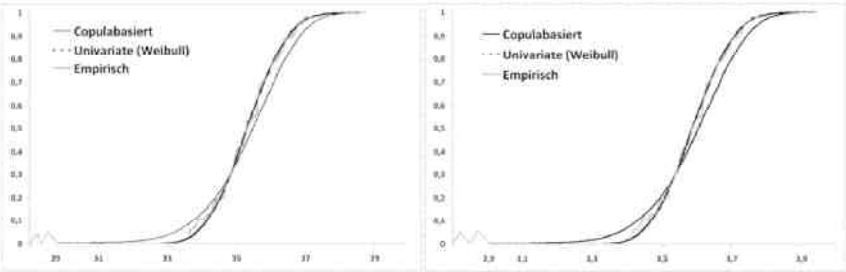
Ebene	Goodness-of-Fit-Test	
1	Test-Statistics	0,0216
	p -Wert	0,5790
2	Test-Statistics	0,0195
	p -Wert	0,7050
3	Test-Statistics	0,0272
	p -Wert	0,3940
4	Test-Statistics	0,0463
	p -Wert	0,0920
5	Test-Statistics	0,0198
	p -Wert	0,7330
6	Test-Statistics	0,0215
	p -Wert	0,6470
7	Test-Statistics	0,0226
	p -Wert	0,6500

Abbildung 3 und Tabelle 1 stellen das zentrale Ergebnis der Analyse dar und zwar die Verteilungsfunktionen für die Hektarerträge (Abb. 3a) und die Gesamtproduktion (Abb. 3b) bezogen auf das Ertragsniveau und die Anbauumfänge von 2007. Da die Verteilung in Abb. 3a aus der Aggregation sehr unterschiedlicher Regionen hervorgegangen ist, ist die enthaltene Information relativ schwer zu bewerten. Als Bezugspunkt mag das Erntejahr 2007 hilfreich sein, also das Jahr, das als besonders schlecht in Erinnerung ist und in dem der eingangs beschriebene weltweite Preisanstieg zu verzeichnen war. In diesem Jahr war ein Durchschnittsertrag von 33,52 dt/ha zu beobachten. Es ist zu erkennen, dass dieses Ertragsniveau c.p. mit einer Wahrscheinlichkeit von 7 Prozent unterschritten wird. Es handelt sich also um ein durchaus seltenes Ereignis, allerdings eines, das sich statistisch unter sonst gleichen Bedingungen etwa alle 15 Jahre wiederholt.

Zu Vergleichszwecken sind neben der copulabasierten Schätzung der Ertragsverteilung noch zwei weitere Verteilungen dargestellt, die alternativ geschätzt wurden. Zum einen wurde eine parametrische univariate Verteilung der Hektarerträge für die Gesamtregion gefittet und geschätzt. Dabei handelt es sich um eine Weibull-Verteilung. Darüber hinaus wurde noch die entsprechende empirische Verteilung bestimmt. Eine solche direkte Schätzung der Ertragsverteilung ist natürlich mit viel weniger Aufwand verbunden, als die Herleitung aus einer multivariaten Verteilung der Einzelländer. Das Vorgehen hat aber auch Nachteile. Die univariate Analyse stützt sich allein auf 47 Beobachtungswerte, und die Schätzung der Verteilung wird

entsprechend unpräzise, insbesondere in den Rändern, die hier von besonderem Interesse sind. Darüber hinaus bleibt die innere Abhängigkeitsstruktur der einzelnen Länder verborgen. Schließlich ist zu befürchten, dass Aggregationsfehler auftreten, wenn Marginalverteilungen unterschiedlichen Typs durch nur eine parametrische Verteilung repräsentiert werden sollen. Insofern verwundert es nicht, dass sich die in Abb. 3 und Tabelle 1 dargestellten Verteilungen unterscheiden. Es ist zu erkennen, dass trotz ähnlicher Erwartungswerte das Risiko extremer Mindererträge in der Gesamtregion durch die direkten Schätzungen der Verteilungen im Vergleich zu der copulabasierten Schätzung unterschätzt wird. So lässt sich das 1 Prozent Quantil aus der empirischen Verteilung aufgrund der Nichtstetigkeit überhaupt nicht bestimmen, und die univariate Weibull-Verteilung weist mit 33,37 dt/ha einen um 1,28 dt/ha höheren Ertrag aus als die copulabasierte Schärtzung.

Abbildung 3: Empirische und parametrische Verteilungsfunktionen der Weizenproduktion (alle Regionen, Bezugsjahr: 2007)



a) Durchschnittsertrag (dt/ha) b) Gesamtertrag (100 Millionen Tonnen)

Tabelle 1: Kennzahlen der geschätzten Verteilungen der Weizenproduktion (alle Regionen, Bezugsjahr: 2007)

Quantile	Copulabasiert		Univariat (Weibull)		Empirisch	
	Abb. 3: a)	Abb. 3: b)	Abb. 3: a)	Abb. 3: b)	Abb. 3: a)	Abb. 3: b)
1%	32,09	3,26	33,37	3,39	n.a.	n.a.
5%	33,22	3,38	33,84	3,44	≈ 33,64	≈ 3,41
10%	33,74	3,43	34,12	3,47	≈ 33,91	≈ 3,45
50%	35,47	3,61	35,29	3,59	≈ 35,29	≈ 3,59
Erwartungswert	35,37	3,59	35,31	3,59	35,30	3,59
Varianz	1,47	0,0152	0,81	0,0084	0,86	0,0089

4 Schlussfolgerungen

Quintessenz der durchgeführten Analyse ist die Feststellung, dass sich Ertragsschwankungen bei Weizen in den wichtigsten Erzeugerregionen der Welt zu einem gewissen Teil ausgleichen. Eine vollständige Diversifikation der Ausfallrisiken ist allerdings nicht möglich, was natürlich auch kaum zu erwarten war. Es wurde deutlich, dass sich eine Angebotskonstellation, wie sie in 2007 aufgetreten ist, kein Jahrhundertereignis darstellt, sondern sich c.p. etwa alle 15 Jahre wiederholen kann. An dieser Stelle erscheint eine Relativierung des Ergebnisses geboten. Zunächst muss betont werden, dass keine Aussage über zu erwartende Knappheiten

am Weltweizenmarkt getroffen werden kann, allein, weil die Nachfrage und deren künftige Veränderungen überhaupt nicht betrachtet wurden. Aber auch die Analyse der Angebotsseite ist unvollständig, denn es wurden nur acht Länder analysiert und nicht die Gesamtproduktion der Welt. Des weiteren wurde nur eine jahresbezogene Betrachtung durchgeführt und eine zeitliche Diversifikation somit ausgeschlossen. Knappheiten sind vor allem dann zu erwarten, wenn mehrere „schlechte" Erntejahre in Folge auftreten und die weltweiten Lagerbestände nicht ausreichen, um solche Mindererträge zu kompensieren. Dies war beispielsweise in 2007/2008 der Fall. Weiterhin beziehen sich die Wahrscheinlichkeitsaussagen auf das durchschnittliche Ertragsniveau am Ende des Beobachtungszeitraums. Für Prognosen über künftige Angebotsengpässe müsste die durch technische Fortschritte getriebene Entwicklung des Ertragsniveaus fortgeschrieben werden. Schließlich sei auf die Endogenität der Anbauflächen hingewiesen, die Einfluss auf die Gesamterzeugung hat. So war zu beobachten, dass infolge des Preisbooms in 2007 die Weizenanbauflächen in 2008 weltweit ausgedehnt wurden. Einige der zuvor genannten Aspekte ließen sich im Rahmen partieller oder allgemeiner Gleichgewichtsmodelle berücksichtigen. Die Verknüpfung solcher Modelle mit der hier vorgestellten stochastischen Analyse stellt eine interessante zukünftige Forschungsaufgabe dar.

Literatur

ARTAVIA, M., A. DEPPERMANN, G. FILLER, H, GRETHE, A. HÄGER, D. KIRSCHKE und M. ODENING (2010): Ertrags- und Preisinstabilität auf Agrarmärkten in Deutschland und der EU: Betriebswirtschaftliche und agrarpolitische Implikationen. In: Rentenbank (Hrsg.): Auswirkungen der Finanzkrise und volatiler Märkte auf die Agrarwirtschaft. Schriftenreihe der Rentenbank Band 26, Frankfurt a.M. (im Druck).

CHARPENTIER, A., J.-D. FERMANIAN und O. SCAILLET (2007): The estimation of copulas: theory and practice. In Rank J. (ed.), Copulas: From theory to application in finance. Risk Publications, London: 35-60.

CHEN, S.-X. und T. HUANG (2007): Nonparametric estimation of copula functions for dependence modeling. In: The Canadian Journal of Statistics 35: 265-282.

CHERUBINI, U., E. LUCIANO und W. VECCHIATO (2004): Copula Methods in Finance. Wiley, Chichester.

EMBRECHTS, P., A. MCNEIL und D. STRAUMANN (2002): Correlation and dependence in risk management: Properties and pitfalls. In: Dempster, M.A.H. (ed.): Risk Management: Value at Risk and Beyond. Cambridge University Press, Cambridge: 176-223.

EMBRECHTS, P., F. LINDSKOG und A. MCNEIL (2003): Modeling Dependence with Copulas and Applications to Risk Managament. In: Rachev, S.T. (ed.): Handbook of Heavy Tailed Distributions in Finance: Elsevier, North-Holland.

FAO (2009): FAOSTAT database. Available at: http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor.

GENEST, C. und B. REMILLARD (2008): Validity of the Parametric Bootstrap for Goodness-of-Fit Testing in Semiparametric Models. In: Ann. Henri Poincare 44: 1096-1127.

HAERDLE, W., O. OKHRIN und Y. OKHRIN (2008): Modeling Dependencies in Finance using Copulae. In: Haerdle, W, N. Hautsch and L. Overbeck (eds): Applied Quantitative Finance (2nd ed.). Springer Verlag, Berlin.

HAERDLE, W. und O. OKHRIN (2009): De Copulis Non Est Disputandum, Copulae: An Overview. SFB 649 Discussion Paper (2009-031).

JOE, H. (1997): Multivariate Models and Dependence Concepts. Chapman&Hall, London.

LOTZE-CAMPEN, H. (2007): Managing regional climate risks in agriculture through diversified international trade relationships. Presentation at the 101st Workshop of the European Association of Agricultural Economists (EAAE) on Managing Climate Risks in Agriculture. Berlin, 05.July 2007. Miemo.

MORIONDO, M., M. BINDI und N. LUGERI (2009): Modelling Yield Distribution as Affect by Extreme Events. In: IOP Conference Series: Earth and Environmental Science 6 (2009): Nr.: 022011.

NELSEN, R.B. (2005): An Introduction to Copulas (2nd ed.). Springer, NewYork.

ODENING, M., E. BERG und C.G. TURVEY (EDS.) (2008): Management of Climate Risk in Agriculture. Agricultural Finance Review 68 (1) Special Issue.

OKHRIN, O. (2007): Hierarchical Archimedean Copulas: Structure Determination, Properties, Applications. PhD thesis at the Faculty of Economics, European University Viadrina, Frankfurt (Oder).

RETTBERG, U. (2005): Wenn Bauern zu Scheichs werden. Handelsblatt vom 04.Dezember 2005.

SARRIS, A. (2009): Factors Affecting recent and Future Price Volatility of Food Commodities. In: Berg, E. et. al (Hrsg.): Risiken in der Agrar- und Ernährungswirtschaft und ihre Bewältigung. Landwirtschaftsverlag Münster-Hiltrup: 29-48.

SKLAR, A. (1959): Fonctions de répartition à n dimensions et leurs marges. In: Publ. Inst. Stat.Univ. Paris 8: 299-231.

UPADHYAY, B.M. und E. G.SMITH (2005): Modeling Crop Yield Distributions from Small Samples. Selected Paper at the CAES annual meeting, July 6-8, 2005, San Francisco, California.

VEDENOV, D. (2008): Application of Copulas to estimation of joint crop yield distributions. Selected paper at the Annual Meeting of the AAEA, July 27-29, 2008, Orlando, Florida. Available at http://ageconsearch.umn.edu/handle/6264.

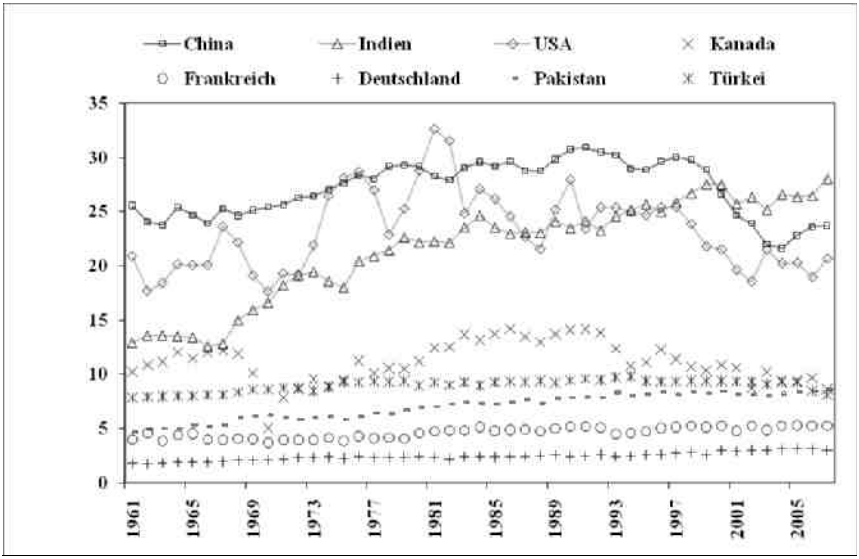
VON WITZKE, H. (2008): Weltagrarmärkte: Einige zentrale Änderungen der Rahmenbedingungen und deren Implikationen für die Landwirtschaft. Ländlicher Raum 10 (2008). 1-13.

WHELAN, N (2004): Sampling from Archimedean Copulas. In: Quantitative Finance 4: 339-352.

XU, W., G. FILLER, M. ODENING und O. OKHRIN (2010): On the systemic nature of weather risk. In: Agricultural Finance Review (in print).

Anhang

Abbildung A1: Ernteflächen für Winterweizen in Millionen Hektar



Quelle: FAO (2009)

Tabelle A1: Trendbereinigte Durchschnittserträge (in dt/ha) aufsteigend sortiert

Rang	Kanada		China		Frankreich		Deutschland		Indien		Pakistan		Türkei		USA		Gesamt	
	Ertrag	Jahr	Ertrag	Jahr	Ertrag	Jahr	Ertrag	Jahr	Ertrag	Jahr	Ertrag	Jahr	Ertrag	Jahr	Ertrag	Jahr	Ertrag	Jahr
1	6,13	1988	0,56	1977	***14,85***	***2007***	20,92	2003	4,66	2006	5,32	1987	8,68	1973	13,48	2002	9,39	2002
2	7,52	1961	1,67	2002	18,96	2003	***21,35***	***2007***	4,96	2005	5,62	1966	***8,90***	***2007***	15,05	2006	***9,43***	***2007***
3	8,97	2002	2,19	1980	20,73	2006	23,59	1976	***5,03***	***2007***	5,73	1984	8,97	1999	15,13	1989	10,13	2003
4	10,32	2001	2,34	2003	22,35	1976	24,80	2006	5,26	1980	5,74	1967	8,99	1994	15,13	1974	10,16	2006
5	10,85	1984	2,78	1973	23,36	1966	26,04	2002	5,50	1974	5,90	1994	9,08	1974	15,65	1991	10,20	1974
6	11,68	1989	2,85	2001	24,23	1961	26,75	1977	5,86	1964	6,31	1997	9,09	1961	15,70	1995	10,20	1977
7	11,86	1967	3,05	2000	24,26	2005	27,05	1966	5,88	1966	6,35	1996	9,28	1970	***15,72***	***2007***	10,36	2001
8	11,91	1974	3,48	1978	24,41	1975	27,28	1992	6,00	2003	6,42	2002	9,56	2001	15,77	1996	10,50	1961
9	12,19	1985	3,48	1981	24,71	2001	27,43	1989	6,00	1967	6,44	1978	9,58	2003	15,90	1967	10,66	1970
10	12,23	1979	3,55	1970	24,82	1963	27,44	1980	6,21	1977	6,56	1991	9,61	1968	16,07	1961	10,75	1980
11	***12,73***	***2007***	3,67	1974	25,12	1970	27,80	1970	6,56	2004	6,64	1985	9,67	1965	16,27	1988	10,80	2005
12	12,88	1980	3,81	1969	25,89	1977	27,80	1981	6,66	1978	6,67	1988	9,73	1964	16,46	1963	10,82	1988
13	12,92	1964	3,91	1971	27,75	1969	28,10	1965	6,68	1975	6,68	1999	9,94	2002	16,46	1966	10,87	1973
14	12,98	2003	4,01	1972	28,46	1971	28,48	1975	6,72	1987	6,74	2004	10,01	1995	16,56	1962	11,08	1978
15	13,07	1997	4,02	2005	28,60	1983	28,49	2005	6,73	1981	6,79	1990	10,04	1989	16,61	1964	11,17	1964
16	13,26	1968	4,07	1975	28,68	1986	28,57	1961	6,86	1982	6,82	1993	10,13	1969	16,68	1976	11,22	1975
17	13,35	1987	4,28	1998	28,69	1964	28,81	1972	6,93	1976	6,88	1971	10,29	1997	16,69	1977	11,28	1994
18	13,75	1994	4,55	1976	28,74	1987	28,87	1979	6,97	1973	7,05	1965	10,38	1996	16,87	1965	11,29	1967
19	13,96	1962	4,59	1991	28,77	1965	28,99	1983	6,97	1963	7,11	1964	10,55	1962	16,95	1978	11,30	1989
20	14,11	1998	4,66	2004	28,96	1997	30,38	1982	7,07	1979	7,28	2003	10,69	1966	17,01	1986	11,51	1969
21	14,14	1973	4,99	1968	29,32	1979	30,46	1969	7,10	1988	7,34	1974	10,85	1967	17,14	1975	11,53	1976
22	14,29	1972	5,21	1994	29,52	1981	30,47	1987	7,15	1998	7,35	1982	11,01	2004	17,15	1994	11,54	1995
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
45	17,34	2005	8,65	1983	35,74	1991	36,23	1991	9,12	2000	8,93	1986	13,82	1976	20,37	1971	13,09	1986
46	17,54	1976	9,32	1997	37,73	1998	36,60	2004	9,32	1995	***8,98***	***2007***	13,89	1979	20,41	1984	13,29	1983
47	17,60	1966	9,44	1984	40,88	1984	36,86	2001	9,56	1997	9,50	2000	14,59	1988	21,09	1983	13,60	1984

Quelle: FAO (2009)

Tabelle A2: Randverteilungen der Weizenerträge

Land	Verteilung	Parameter		χ^2 Test-Statistik	p -Wert
Kanada	Logistisch	Lage	0,0243	2,1910	0,9485
		Skala	0.1382		
China	Weibull	Form	3,4396	4,5740	0,7117
		Skala	0,6694		
		Verschiebung	-0,6022		
Deutschland	Normal	Mittelwert	0,0000	2,870	0,8966
		Standardabweichung	0,3895		
Frankreich	Normal	Mittelwert	0,0000	8,6600	0,2780
		Standardabweichung	0,4975		
Indien	Logistisch	Lage	0,0035	3,5530	0,8296
		Skala	0,0690		
Pakistan	Weibull	Form	3,4666	3,8940	0,7919
		Skala	0,3160		
		Verschiebung	-0,2842		
Türkei	Weibull	Form	1,7138	5,2550	0,6288
		Skala	0,3280		
		Verschiebung	-0,2934		
USA	Weibull	Form	3,7793	4,5740	0,7117
		Skala	0,6087		
		Verschiebung	-0,5498		