A Model of Enterprise-Level

Crop Yields under Climate Change

Proof of concept,

general computational strategy

and

partial implementation for the case of grain production

in Saxony-Anhalt, Germany

Dissertation

zur Erlangung des Grades

Doktor der Wirtschaftswissenschaft (Dr. rer. pol.)

der Juristischen und Wirtschaftswissenschaftlichen Fakultät

der Martin-Luther-Universität Halle-Wittenberg

Prof. Dr. Claudia Becker

Prof. Dr. Michael Grings

vorgelegt von

Diplom-Volkswirtin Denitsa Angelova

Halle (Saale),

24.11.2016

Erster Gutachter:	Prof. Dr. Claudia Becker
Zweiter Gutachter:	Prof. Dr. Michael Grings
Tag der mündlichen Prüfung:	23.01.2017
Erscheinungsjahr:	2017

Abstract

This thesis addresses the challenge of constructing a model explaining and predicting the changes in agricultural production under climate change. The thesis reviews existing economic models for climate change impact assessment: the Ricardian approach by Mendelsohn, Nordhaus and Shaw, an econometric panel data approach of Deschênes and Greenstone as well as a branch of models relying on mathematical programming. The thesis identifies a potential bias of the existing economic models on the enterprise level, thus establishing the need for a novel enterprise-level crop yield model under climate change.

A model for the approximation of enterprise-level crop yields under climate change is proposed, which integrates both economic and agronomic notions of crop production. It is based on the idea that the enterprise-level crop yields are representable as a function of biophysically determined crop yields and the optimal production aspirations. The model is potentially useful to policy makers since it could provide an estimate of the yield volume that can be expected to be produced under climate change. It can also be useful to farmers as a normative model recommending input commitment under a changing climate.

The proposed model is based on the state-contingent approach to production and decision-making under uncertainty introduced by Robert Chambers and John Quiggin (2000). The so-called states giving the approach its name refer to mutually exclusive scenarios for the development of the uncertain future. The theory is chosen due to its advantages, including the original unification of two economic subfields – the subfield of production economics and the subfield of decision-making under risk and uncertainty. More specifically, the state-contingent approach offers an explanation of production decisions under uncertainty with a production economic foundation. It can also deliver well-defined cost functions corresponding to stochastic production technologies.

A conceptual challenge is posed by the limitations of the state-contingent approach not being explicit on how agents construct expectations regarding the future levels of relevant variables under uncertainty. This thesis overcomes the challenge by combining

it with the hypothesis of adaptive price expectations. A general computational strategy for the proposed model is provided.

The strategy introduces a novel geometrical interpretation of the coefficient of risk-aversion, which is intrinsic to the farmer, as well as an approach to infer this coefficient of risk-aversion from farm-level accounting data. The identification strategy for the coefficient of risk-aversion relies on using the duality between the possible technology representations, a skillful data partitioning and the employment of a so-called grid-search, a machine-learning technique, which is underemployed as an instrument in the fields of agricultural economics and decision-making. The specific geometrical interpretation facilitates the identification. This result can also be transferred to the deterministic case: a similar interpretation of a coefficient expressing intertemporal preferences can be introduced and the proposed strategy for identification applied to elicit the coefficient.

A new indicator to detect the nature states, i.e. to infer the environmental conditions prevalent in a specific year, is proposed as a methodological contribution based on the previous work by the author. This detection is needed in order to estimate a primal technological representation like a production or a distance function. The indicator evaluates phenological observations rather than weather data, the latter being used by Nauges, O'Donnell and Quiggin (2011). Two types of statistical clustering algorithms to process the indicator are considered and their advantages and disadvantages investigated.

On the empirical side, the thesis offers an estimation of a state-contingent production function for the crop-producing sector in the Federal State of Saxony-Anhalt, Germany. Based on simulated data the estimation rejects the hypothesis of an output cubical technology, which would have indicated a technologically presupposed inability of the farmers to adapt to environmental changes by adjusting their input commitment.

Acknowledgements

The author acknowledges the support of the ScienceCampus Halle - plant-based bioeconomy. This research was possible within a subproject at the Campus with principle investigators Prof. Dr. Thomas Glauben (IAMO) and Prof. Dr. Michael Grings (Martin Luther University of Halle-Wittenberg), funded by the Leibniz Society and the Federal State of Saxony-Anhalt in 2012-2015. Within this project the author had the chance to get acquainted with the state-contingent approach, which proved to be a research tool bringing new challenges, insights and horizons.

Gratitude is expressed to Prof. Dr. Thomas Glauben for his guidance. Special thanks go to Prof. Dr. Michael Grings for the guidance and the long discussions, careful considerations and the patience throughout the joint work. The author is deeply indebted to Prof. Dr. Claudia Becker for her guidance in statistic modelling, the encouragement and her support during the thesis finalization phase. The author expresses most sincere appreciation to Prof. Dr. Heinrich Hockmann for fruitful discussions on the topic of empirical production analysis. Numerous conversations, guiding questions and critical remarks by Prof. Becker, Prof. Grings and Prof. Hockmann vitally assisted the author in thinking the concept of the proposed model through. Their guidance on empirical modeling was invaluable too.

Gratitude is expressed also to Dr. Ihtiyor Bobojonov for the valuable opinions on the calibration of farm-level mathematical programming models incorporating risk and uncertainty under climate change. The author would like to thank Dr. Thomas Chudy for his guidance in agronomic matters and for generating a map. She is indebted to Dr. Jörg Gersonde for his support in mathematical matters and to Brett Hankerson for a helpful conversation on the characteristics of different crops. The author sincerely appreciates a fruitful discussion with Falk Böttcher about the suitability of different climate simulation models depending on the time horizon of the specific climate change impact assessment model. The author expresses gratitude to the Landesanstalt für Landwirtschaft, Forsten und Gartenbau (LLFG) for assisting with the data acquisition. Gratitude is expressed to Prof. Dr. Christoph Wunder for useful discussions in econometric matters.

Warm appreciations go to Dr. Sven Grüner for his continuous help as a colleague and a friend as well as for his encouragement throughout the development of this project. Last but not least, the author is grateful to her entire family for their love, care and unconditional support over the last four years. Dr. Stefan Reinicke is another very special person deserving of a very special “thank you”.

Contents

List of Figures and Maps

List of Tables

List of Abbreviations

BELV	Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz
BVEL	Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft
COP	Conference of the Parties
IPCC	Intergovernmental Panel on Climate Change
LLFG	Landesanstalt für Landwirtschaft, Forsten und Gartenbau
LUF	Landwirtschaftliche Untersuchungs- und Forschungsanstalt
MELCY	Model of Enterprise Level Crop Yields
MELF	Ministerium für Ernährung, Landwirtschaft und Forsten des Landes Sachsen-Anhalt
MLU	Ministerium für Landwirtschaft und Umwelt des Landes Sachsen-Anhalt
MRLU	Ministerium für Raumordnung, Landwirtschaft und Umwelt des Landes Sachsen-Anhalt
UNFCCC	United Nations Framework Convention on Climate Change

List of Symbols

$\mathrm{c}(\boldsymbol{w},\boldsymbol{y})$: effort-cost function under linear input pricing

$\mathrm{C}(\boldsymbol{w},\boldsymbol{r},\boldsymbol{p})$: revenue-cost function under linear input pricing

λ: expectation adjustment parameter

η: risk aversion parameter

$\mathrm{O}(\boldsymbol{y},\boldsymbol{x})$: state-contingent output-distance function

$\boldsymbol{p}$: $M \times S$ matrix of state and output specific prices

$p_{ms}^{e}(t)$: expected price for output m in period t under s

$p_{ms}^{e}(t_s)$: expected price for output m under s in period t_s, the last period s occurred in

$p_{ms}(t_s)$: price for output m under state s in period t_s, the last period state s occurred in

π_s: probability of state s occurring

$\boldsymbol{r}$: S dimensional vector of state-contingent revenues

$\mathbb{R}_{++}^{N}$: N dimensional subspace of strictly positive real numbers

$\mathbb{R}_{+}^{N}$: N dimensional subspace of positive real numbers

$\mathbb{R}_{+}^{M \times S}$: $M \times S$ dimensional subspace positive real numbers

$s, s \in \Omega$: state of nature or nature-state

Ω: set of nature-states

$\Theta_{t-1}^{(k)}$: information set of producer k containing environmental, technological and market conditions up to period t

$\boldsymbol{w}$: N dimensional vector of input prices

$\boldsymbol{x}$: N dimensional input vector

$\mathrm{X}(\boldsymbol{y})$: set of input vectors, which can produce a matrix of state-contingent outputs $\boldsymbol{y}$

$\boldsymbol{y}$: $M \times S$ matrix of state-contingent outputs

$\boldsymbol{y}_{\boldsymbol{s}}$: M dimensional vector of outputs in state s

y_{ms}: quantities of output m in state s

$y_{ms}^{(k)}(t)$: crop yields of crop m of producer k in period t under state of nature s

$y_{ms}^{biophy}(t)$: biophysical crop yields of crop m in agricultural cycle t under state of nature s

$y_{ms}^{*(k)}\left(t|\Theta_{t-1}^{(k)}\right)$: production aspirations of producer k with regard to crop m in period t under state of nature s given the information set $\Theta_{t-1}^{(k)}$

$\mathrm{Y}(\boldsymbol{x})$: set of state-contingent output matrices, which can be produced with input vector $\boldsymbol{x}$

Glossary

Adaptation: “The process of adjustment to actual or expected climate and its effects. In human systems, adaptation seeks to moderate or avoid harm or exploit beneficial opportunities.” (IPCC, 2014, p. 1251)

Agricultural cycle: Annual agricultural cycle with respect to the growing and harvesting of crops.

Climate: “Climate in a narrow sense is usually defined as the average weather, or more rigorously, as the statistical description in terms of the mean and variability of relevant quantities over a period of time ranging from months to thousands or millions of years. The classical period for averaging these variables is 30 years, as defined by the World Meteorological Organization. The relevant quantities are most often surface variables such as temperature, precipitation and wind. Climate in a wider sense is the state, including a statistical description, of the climate system.” (IPCC, 2014, p. 1255)

Climate change: “Climate change refers to a change in the state of the climate that can be identified (e.g., by using statistical tests) by changes in the mean and/or the variability of its properties and that persists for an extended period, typically decades or longer. Climate change may be due to natural internal processes or external forcings such as modulations of the solar cycles, volcanic eruptions and persistent anthropogenic changes in the composition of the atmosphere or in land use.” (IPCC, 2014, p. 1255)

Climate model: “A numerical representation of the climate system based on the physical, chemical and biological properties of its components, their interactions and feedback processes and accounting for some of its known properties … Climate models are applied as a research tool to study and simulate the climate and for operational purposes, including monthly, seasonal and interannual climate predictions.” (IPCC, 2014, p. 1256)

Crop simulators (Crop simulation models): “Computer programs that mimic the growth and development of crops. Data on weather, soil, and crop management are processed to predict crop yield, maturity date, efficiency of fertilizers and other elements of crop production. The calculations in the crop models are based on the existing knowledge of the physics, physiology and ecology of crop responses to the environment.” (United States Department of Agriculture, 2016)

Cultivar: “A variety of plant that originated and persisted under cultivation.” (Dictionary.com, 2016)

ex ante: Based on anticipated changes or activity in an economy. (Dictionary.com, 2016b)

ex post: Based on analysis of past performance. (Dictionary.com, 2016c)

Normative model: A prescriptive model, which investigates what ought to be. (IPCC, 2014, p. 1267)

Positive model: A descriptive model, which investigates what is. (IPCC, 2014, p. 1268).

Production function

agronomic: Functional relationship between crop yields (response variable) and factors of the biophysical environment e.g. water availability and nitrogen content in the soil (explanatory variables).

economic: Functional relationship between an output e.g. crop yields (response variable) and factors of production or inputs e.g. labor, capital, land, intermediate inputs (explanatory variables).

States of nature (nature states): Mutually exclusive possible states of the world in the state-contingent framework e.g. hail or the absence of hail.

1 Introduction

The recent Conference of the Parties (COP) in the United Nations Framework Convention on Climate Change and the presidential elections in the United States of America once again establish the relevance of the topic of climate change as well as the significance of the economic implications of the phenomenon itself and of the measures taken to limit its extent (Scherer, 2016), (Wong, 2016), (United Nations, 2016).

Climate change is broadly defined in the reports of the Intergovernmental Panel on Climate Change (IPCC) as a lasting change in the statistical properties of a climate system (Core Writing Team 2014, pp. 120-121). The phenomenon is expected to have a significant impact on agricultural productivity (Gillis, 2013). Improving the understanding of the effects of a shift in climate parameters on agricultural productivity is a challenge explicitly identified by the latest report of the IPCC (Smith et al., 2014, p. 869). Constructing models explaining and predicting the changes in agricultural production, which could be expected with a shift in the parameters of a climate system, is important for policy recommendation purposes.

Creating such models, however, is a challenge. Their development requires taking into account multiple factors from: *(i)* the natural environment, *(ii)* the agronomic environment, *(iii)* the decision-making of the individual farmer and his technological considerations, *(iv)* the general economic environment responsible for price formation. The interconnections between these multiple factors should also be taken into account. Multiple theories and models within *(i)* to *(iv)* exist to explain phenomena from different perspectives. However, integrating the theories and models from different fields as well as a transfer of the insights is difficult: the theories and models might use concepts different in their nature. Even in the case of a matching concept the event might be observed at a different level of granularity. The integration of a variety of theories and concepts in a coherent whole is thus far from trivial and poses a challenge.

The thesis addresses the identified challenge. The work is inspired by a recent theoretical development, the so-called state-contingent approach to production analysis and decision-making under uncertainty developed by Chambers and Quiggin (Chambers and Quiggin, 2000). The so-called states, which give the approach its name, refer to mutually exclusive scenarios for the development of the uncertain future. Clearly innovative in the approach is the unification of two economic subfields – the subfield of production economics and the subfield of decision-making under risk and uncertainty. More specifically, the state-contingent approach offers an explanation of production decisions under uncertainty with a solid production economic foundation.

This work is motivated by the research, which started in the project ScienceCampus Halle - plant-based bioeconomy, in a subtask aiming at the investigation of the climate change effects on the agricultural production of Saxony-Anhalt. The state-contingent approach serves as a theoretical foundation for the research work due to its advantages. Statistical and econometric evaluation of data on agricultural production in the German Federal State of Saxony-Anhalt is performed in order to demonstrate the potential relevance of the state-contingent approach.

The ultimate goal of this thesis is to develop a model for enterprise-level crop yields under climate change, to provide a method for the calculation of the model and to demonstrate its computation using available data for the case of Saxony-Anhalt. Partly the work has to be theoretical due to the fact that the state-contingent approach is not formulated as a consistent conceptual model that explains all notions needed to understand enterprise-level yields under climate change. Partly the work has to be methodical, due to the necessity to design a procedure for the computation of the model, including the provision of a method for tuning the model to a specific empirical case. The work also needs to have an empirical component.

More specifically, the following objectives are set for the thesis:

i. To integrate economic and agronomic notions of crop production in a coherent whole.
ii. To develop a model based on the unified perspective and suggest a strategy to empirically substantiate the dynamics assumed for the model.
iii. To overcome a conceptual challenge posed by the state-contingent approach not being explicit on how economic agents form expectations on the level of relevant variables, more specifically output prices, under uncertainty.
iv. To check if the production technology of the crop-producing sector is output-cubical. An output-cubical technology would mean that the farmers cannot adapt in the state-contingent sense.

Three distinct research questions are set in the thesis. The first research question is the basic research question on whether or not the insights of the state-contingent approach can be adapted to the field of climate change impact assessment. The second research question is whether the conceptual challenge posed by the state-contingent approach not being explicit on how economic agents form expectations can be overcome. The third and purely applied research question is whether the crop production technology farmers in Saxony-Anhalt operate under is an output-cubical technology.

The research questions are framed as the following hypotheses:

i. It is possible to construct a computable model integrating both economic and agronomic notions of crop production under climate change.

ii. It is possible to overcome the conceptual challenge posed by the state-contingent approach not being explicit on how economic agents form expectations.

iii. The crop production technology of the farmers in Saxony-Anhalt is output cubical.

Figure 1 provides an overview of the scope and main concepts of the thesis, which approaches the prediction of enterprise-level crop yields under changing environmental conditions from both theoretical and empirical perspectives.

The thesis structure is as follows. Chapter 2 provides a literature overview of economic models for climate change impact assessment. Section 2.1 presents two popular econometric approaches to climate change impact assessment. Section 2.2 deals with mathematical programming models for climate change impact assessment. The existing models are subsequently discussed in section 2.3, knowledge gaps are identified and the necessity for the proposed model is established.

Chapter 3 discusses the theoretical foundations of this work: *(i)* the state-contingent approach to production and decision-making under uncertainty and *(ii)* the hypothesis of adaptive expectation formation. Essential elements and concepts of the state-contingent approach to production analysis and decision-making under uncertainty are depicted on the left hand-side of Figure 1 and presented in section 3.1: the notion of a state-contingent production function; a generalization of the notion of a stochastic production function, which, unlike the standard stochastic production function, has a corresponding cost function, the so-called effort-cost function. This effort-cost function serves in the theory as a basis for deriving a so-called revenue-cost function, a function serving effectively as a boundary describing feasible production alternatives expressed in monetary terms, which an agent could have a preference over. The hypothesis of adaptive expectation formation, which explains how economic agents form expectations on the level of relevant variables under uncertainty, is described in section 3.2. Section 3.3 analyzes the limitations to the state-contingent approach: the theory of state-contingent production relies on the assumption that output prices are known to the agent in order to make the transition from an effort-cost to a revenue-cost function. This assumption clearly limits the effective use of the concept of a revenue-cost function in empirical implementations.

Chapter 4 provides a literature overview of empirical studies employing the state-contingent approach in order to see how the theory has been operationalized before. Section 4.1 presents two studies in production analysis based on it. Section 4.2 presents a farm-level mathematical programming model based on the approach, which investigates farm-level adaptation to climate change. Section 4.3 discusses the advantages and limitations of these studies.

Chapter 5 introduces a model, which combines the two theoretical pillars presented in chapter 3. The model is constructed around the hypothesis that average crop yields per hectare in a specific agricultural period can be predicted as a function of two quantities: *(i)* biophysically determined crop yields, which would depend to a large extent on geographical location, soil type and weather, and *(ii)* a quantity referred to in the thesis as "output aspirations" – the output levels a rational producer should be striving to reach, which would depend largely on factors such as input and output prices as well as the past production experiences.

Chapter 6 focuses on an empirical challenge in implementing the model as illustrated in the right hand-side of Figure 1. The chapter presents a methodology for defining the states of nature and for detecting when a specific state of nature has occurred. A state-contingent production function is estimated for the grain producing sector in Saxony-Anhalt based on simulated data and the estimation results are discussed.

Chapter 7 puts the estimation results described in chapter 6 into the perspective of the model. It also discusses the data requirements necessary for the implementation of the model and outlines a sequence of steps to empirically substantiate the assumed model dynamics and eventually test the performance of the model in the statistical sense. It presents a novel methodology for identifying the appropriate weights for price expectation formation and degree of risk aversion from farm level accounting data.

Chapter 8 discusses the gained insights and the results, interprets the model in the context of the climate change impact assessment literature and concludes with remarks on future work.

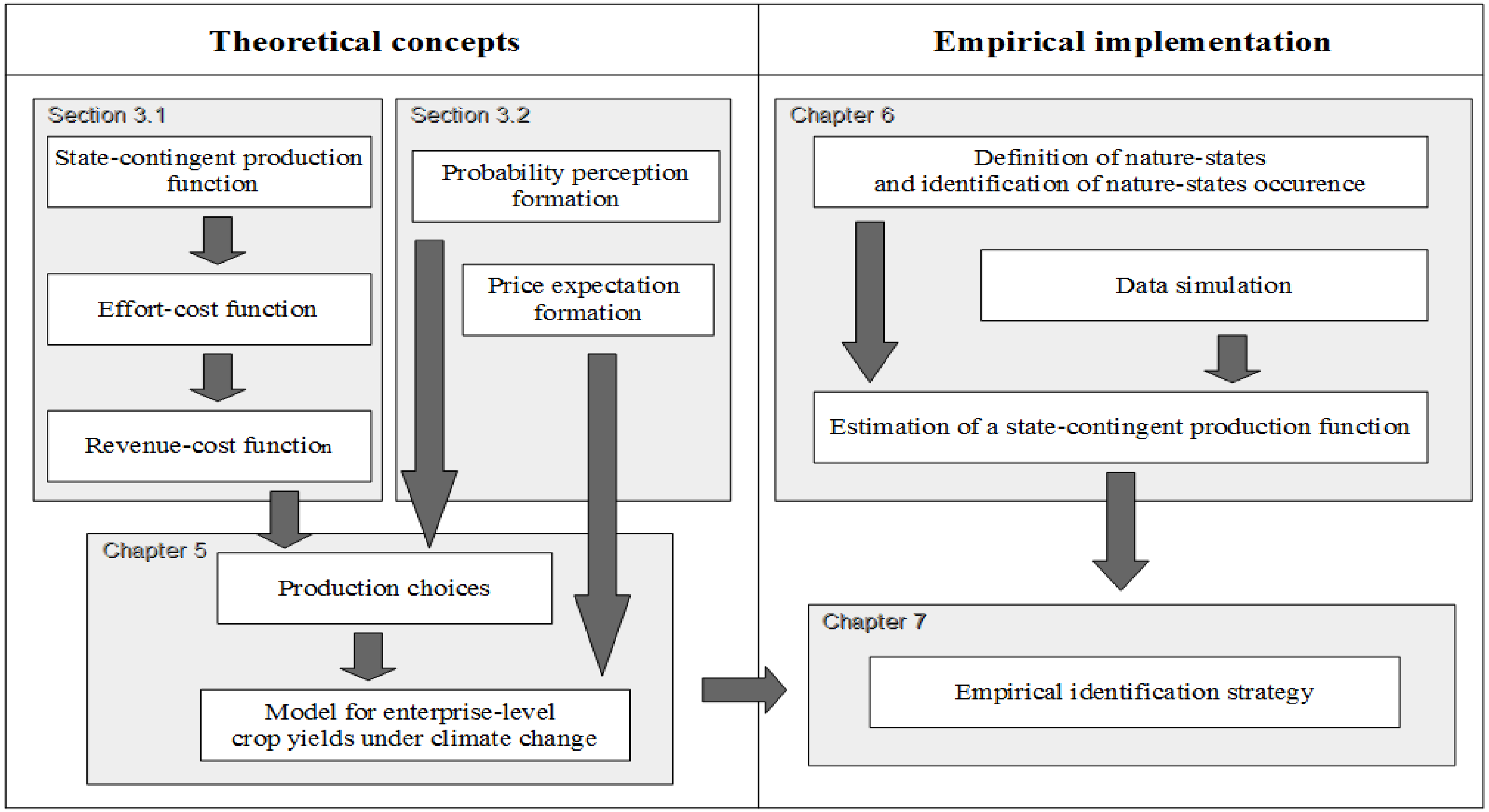


Figure 1 Scope of the thesis.
Note: Chapters 2 and 4 review literature.
Source: own illustration.

2 Literature overview

This chapter provides an overview of existing economic models for climate change impact assessment. The relevance of the models has been decided upon based on the prominence of the contributions in the field as outlined by the Handbook on Climate Change and Agriculture edited by Dinar and Mendelsohn (Dinar and Mendelsohn, 2011).

A brief critical evaluation of the presented models and modeling approaches is given after their presentation, while section 2.2 recapitulates the observations. Section 2.2 also identifies a potential bias in the climate change impact assessment models on the farm level and thereby outlines the necessity for the enterprise-level crop yield model proposed in this work.

2.1 Economic models for climate change impact assessment

This section reviews three branches of economic models. Two of the approaches, the Ricardian one described in subsection 2.1.1 and the sector level econometric evaluation described in subsection 2.1.2, rely on comparative statics. They provide a climate change impact estimate with respect to an aggregated quantity after farmers have adapted to the altered environmental conditions. None of the approaches is targeting an explanation of how farmers gradually adapt to environmental changes. The third branch of models, usually referred to as relying on mathematical programming techniques and presented in section 2.1.3, is vastly popular as a normative tool for decision-making support in agricultural production. The necessity to investigate the gradual farm level adaptation to climate change has led to the recognition of the branch as a positive tool for climate change impact assessment. As a positive tool for prediction the mathematical programming models depict the gradual adaptation of a specific farmer to changes in environmental conditions.

2.1.1 Mendelsohn, Nordhaus and Shaw: Ricardian farm land value analysis

The Ricardian approach to climate change impact assessment was introduced by Mendelsohn, Nordhaus and Shaw (Mendelsohn, Nordhaus and Shaw, 1994) in their analysis of land value sensitivity to climate and soil type in the case of the United States. It is also sometimes referred to as the hedonic approach (Deschênes and

Greenstone, 2007, p. 355). The study was initially conceptually contrasted by the authors to studies relying on the “traditional production function approach” (Mendelsohn, Nordhaus and Shaw, 1994; Passel, Massetti and Mendelsohn, 2014).

This fact is being explicitly mentioned here in order to sound a note of caution: climate change impact literature is rife with references to the “production function approach.” What is commonly meant by “production function” has agronomic origins and connects crop yields to environmental variables, for example average temperature and precipitation, as can be inferred from the comprehensive literature overview of Iglesias et al. on modeling crop productivity changes (Iglesias et al., 2011). The agronomic production function, commonly used to support policy making, is applied for example in the contribution of Iglesias, Quiroga and Diz (Iglesias, Quiroga and Diz, 2011) and should not be confused with the economic notion of a production function, which connects outputs and factor inputs.

The Ricardian approach to impact assessment relies on estimating the dependence of farmland value on climate, soil and socioeconomic variables such as population density or geographical proximity to output markets, while implicitly assuming that farmers have maximized profits by choosing the levels of output and input commitment under the environmental constraints they cannot control (Mendelsohn and Dinar, 2009, p. 54). Adaptation to environmental conditions in terms of agricultural producers optimizing a crop or cultivar strategy and choosing the optimal input quantities is thus assumed to have already taken place (Mendelsohn and Dinar, 2009, p. 54).

The significance of farm land value as a dependent variable is derived from the idea that under competitive markets the value of farm land should reflect the present value of the expected future earnings from a parcel of land (Mendelsohn and Dinar, 2009, p. 54). Farm-level data should thus allow a researcher to measure the current sensitivity of farmland value with respect to changes in climate data, which is considered exogenous in the model (Mendelsohn and Dinar, 2009, p. 54). The measured current sensitivity is then used to provide a crude estimate of the likely economic implications of a change in the level of an independent variable of the climate type, for example a marginal increase in mean surface temperature.

Formally, denoting the present value of farm land by V, and adapting from Mendelsohn and Dinar (Mendelsohn and Dinar, 2009, p. 54), the equation for discrete time reads:

$$V = \sum_{t=T+1}^{\infty} \frac{P_{LE}}{(1+\varphi)^t} = \sum_{t=T+1}^{\infty} \frac{\sum_i P_i Q_i(\boldsymbol{X},\boldsymbol{F},\boldsymbol{Z}) - \sum_j R_j X_j}{(1+\varphi)^t} \tag{1}$$

where

P_{LE}: net revenue per hectare,

P_i: price of crop i,

Q_i: output quantity of crop i per hectare,

$\boldsymbol{X}$: input vector,

$\boldsymbol{F}$: climate variable vector,

$\boldsymbol{Z}$: soil variable vector,

R_j: price of input j,

X_j: quantity of input j (e.g. labor, capital) per hectare,

t: a single time period,

T: current time period,

φ: discount rate.

It should be noted that the Ricardian approach to impact assessment relies on the hypothesis of informationally efficient markets, i.e. on the idea that farm level values fully reflect the opinion of market participants on the impacts of climate on the net revenues. This would not be the case if price rigidities are present. The approach thereby implicitly relies on the market participants being correct in their assessment of climate impacts. This would not necessarily be the case in the presence of valuation dynamics, which results in pricing bubbles.

Both of these issues are dealt with in the empirical regression analysis using control variables. The following empirical specification of the standard Ricardian model in vector form includes the so-called “quadratic formulation of climate” (Mendelsohn and

Dinar, 2009), which contains temperature and precipitation in linear and quadratic terms:

$$V = B_0 + \boldsymbol{B_1 F} + \boldsymbol{B_2 F}^2 + \boldsymbol{B_3 Z} + \boldsymbol{B_4 P} + e \qquad (2)$$

where

V: farm land values,

$\boldsymbol{F}$: temperature and precipitation variables,

$\boldsymbol{Z}$: soil and economic control variables,

$\boldsymbol{P}$: output market price variables,

e: error term.

Naturally, in empirical terms the applicability of the Ricardian approach is restricted by a potential lack of large, reliable datasets. A suitable dataset should cover a region geographically and climatically varying enough to deliver sound climate change impact estimates. Unsurprisingly, the first applications of the approach focused on the United States (Mendelsohn, Dinar and Shaw, 1996; Mendelsohn and Dinar, 2003).

Further notable implementations for the US exist (Schlenker, Hanemann and Fisher, 2005, 2006, 2007). While mean surface temperature was one of the climate variables used in the original study by Schlenker, Hanemann and Fisher in 2005, the more recent contributions in 2006 and 2007 have demonstrated the superiority of using growing season degree days instead of temperature.

The Ricardian analysis extrapolates on what future farmers are likely to do under specific climate conditions by assessing what current farmers do under geographically varying conditions (Mendelsohn, Dinar and Shaw, 1996). The approach assumes that the adaptation of farmers has already taken place.

In terms of equation (1) the Ricardian approach is not suitable for examining the gradual adaptation of farmers to environmental changes because it does not identify how farmers would arrive from one set of output choices Q_i and input commitments X_j to another set of output choices Q_i and input commitments X_j as the specific climate conditions $\boldsymbol{F}$ and soil conditions $\boldsymbol{Z}$ evolve. In the words of Massetti and Mendelsohn:

"… the Ricardian method does not explicitly measure adaptation. It does not identify how farmers adapt; it merely measures the resulting consequences of adaptation" (Massetti and Mendelsohn, 2011).

2.1.2 Deschênes and Greenstone: econometric approaches for sector level analysis

Another branch of sector level, empirically-statistically based approaches for climate change impact assessment is exemplified by the panel data evaluation introduced by Deschênes and Greenstone, whose initial analysis indicates a slight increase of US aggregate agricultural profits due to climate change (Deschênes and Greenstone, 2007, p. 356). A reinvestigation by the authors in 2011 arrives at the conclusion that climate change will have a negative, yet modest effect on US aggregate agricultural profits by the year 2100 (Deschênes and Greenstone, 2011, p. 113). The reevaluation of the original study (Deschênes and Greenstone 2012), following a comment by Fischer et al. in 2012 (Fischer et al., 2012) also indicates negative impacts of climate change on US cumulative agricultural profits.

The approach of Deschênes and Greenstone for climate change impact analysis consists in regressing county level agricultural profits on socio-economic indicators and presumably random fluctuations in growing season weather over the geographically heterogeneous territory of the US. Like the Ricardian approach, the approach by Deschênes and Greenstone extrapolates on the agricultural outcomes likely to happen in the future under specific climate conditions by assessing what is currently happening under geographically varying conditions, the random fluctuations in growing season weather in this case. Deschênes and Greenstone postulate that although the variable to be explained in the analysis proposed by them is not land value, the change in land value can be easily approximated by calculation based on the estimated change in agricultural profits (Deschênes and Greenstone, 2011, p. 115). The predictions delivered by the Ricardian approach and the approach of Deschênes and Greenstone are thus comparable.

The approach is suitable for an evaluation of the climate change impact after adaptation has occurred, to the extent that farmers are capable of using the whole range of adaptation options as a response to yearly weather fluctuations. The potential bias of the assessment due to an inability to unfold their adaptation options in the case of short-term weather fluctuations is acknowledged by Deschênes and Greenstone (Deschênes and Greenstone, 2007, p. 355).

The standard empirical specification (Deschênes and Greenstone, 2007, p. 367) reads:

$$y_{ct} = \alpha_t + \gamma_t + \boldsymbol{X}_{\boldsymbol{ct}}^T \boldsymbol{\beta} + \sum\nolimits_i \theta_i f_i(W_{ict}) + u_{ct} \tag{3}$$

where

y_{ct}: county level agricultural profits in county *c* in year *t*,

α_t: county level fixed effect,

γ_t: shock associated with year *t* and common for all counties (example: price shocks),

$\boldsymbol{X}_{\boldsymbol{ct}}^T$: row vector of socio-demographic characteristics of county *c* in year *t*,

i: index, associated with one of eight climate variables (e.g. growing season degree days),

W_{ict}: value of a climate variable *i* for county *c* and year *t*,

u_{ct}: stochastic shock, idiosyncratic for county *c* in year *t*.

The parameters θ_i capture the current sensitivity of agricultural profits to growing season weather fluctuations and the interaction between climate variables, which is modeled by the quadratic model $f_i(.)$. The latter includes the climate variables as well as squared values of the climate variables. Nonlinearities in the form of an interaction between the variables is, as with the Ricardian approach outlined in section 2.1.1, not included in the formulation.

As the authors remark, the effects of different combinations of climate variable fluctuations are relatively straightforward to calculate since the cumulative effect on agricultural profits is a linear combination of the parameters θ_i (Deschênes and Greenstone, 2007, p. 366). The impact of climate change on profits is subsequently assessed by a linear extrapolation using the obtained estimates of current sensitivity and the predictions of future climate developments generated with a climate model for specific scenarios. It should be noted that γ_t, the shocks common for all counties, seem to be omitted from the predictions.

The plausibility of the results of the linear extrapolation is safe-guarded by the fact that the authors work with a county-level panel covering the whole geographically and climatically heterogeneous US territory. The decrease in profits in one area (in the original example by the authors: the state of California) seems to be offset by an increase in profits in another one (in the original example by the authors: the state of South Dakota) so that aggregate profits seem unaffected in the original evaluation of 2007 (Deschênes and Greenstone, 2007, p. 367).

The evaluation following the comment by Fischer et al. uses corrected datasets as well as a more sophisticated version of the climate model (Deschênes and Greenstone, 2012, p. 3762). The model has been extended to include storage and inventory adjustments in response to yield shocks. The predicted impact of climate change on cumulative agricultural profits becomes negative. The magnitude of the impact is decreased by the storage and inventory adjustment options (Deschênes and Greenstone, 2012, p.3771).

The novelty of the approach of Deschênes and Greenstone as well as its heavy relience on low-aggregation level weather projections resulting from highly sophisticated climate models, might have restricted the number of studies utilizing the approach for the time being. As with the Ricardian approach described in section 2.1.1 however, the usability of the econometric technique of Deschênes and Greenstone with respect to the study of adaptation is restricted. This is related to the observation that, similarly to the Ricardian approach, the technique of Deschênes and Greenstone does not target to explain the gradual adaptation of the production choices of farmers.

2.1.3 Farm level optimization models

There are alternative models allowing for the study of adaptation options: the farm-level mathematical programming models, which rely on a variety of optimization techniques. As pointed out by Peck and Adams, the low level of spatial aggregation allows the explicit inclusion of multiple features of the production and decision-making environment farmers operate in, uncertainty included (Peck and Adams, 2011, p. 89). A programming model incorporating uncertainty in the decision-making environment could, for instance, rely on maximizing the total gross margin of the producer in a stochastic production environment. The impact of different management options, such as the decision to invest in an irrigation system or alter the crop rotation schedule, can be investigated in this matter (Peck and Adams, 2011, p. 89).

Discrete stochastic programming is a programming technique used in climate change impact assessment. The optimality criterion can be varied according to the management objectives under investigation, for instance maximizing the discounted net profits. An example for the application of non-linear programming in the context would be the model of Finger and Schmid presented in 2.1.3.1. Another example for the application of discrete stochastic programming in this context is the model of Dono and Mazzapicchio presented in 2.1.3.2.

2.1.3.1 Finger and Schmid: the expected value approach

Finger and Schmid integrate crop simulations in a farm level programming framework in order to investigate the impact of climate change on Swiss crop production. More specifically, the authors consider the farm-income stabilizing potential of adaptation options such as the variation in sowing dates, the production intensity and irrigation farming (Finger and Schmid, 2008, p. 26).

The representative farmer is assumed to maximize the certainty equivalent CE associated with managerial options $\boldsymbol{X}$ and I under deterministic input and output prices:

$$max_{\boldsymbol{X},I} CE = pE[Y(\boldsymbol{X})] - \boldsymbol{Z}^T\boldsymbol{X} - IK - \gamma p \sigma_Y(\boldsymbol{X}) \tag{4}$$

subject to

$$Y(\boldsymbol{X}) = \boldsymbol{\beta}^T\boldsymbol{X} \tag{5}$$

where

p: output price associated with a crop,

$E[Y(\boldsymbol{X})]$: expected yield of a crop given input vector $\boldsymbol{X}$,

$\boldsymbol{Z}$: column vector of input prices,

$\boldsymbol{X}$: column vector of inputs,

$I \in \{0; 1\}$: dummy variable $\{\text{no irrigation}; \text{irrigation}\}$,

K: cost of the adoption of an irrigation system,

γ: coefficient of risk aversion,

$\sigma_Y(\boldsymbol{X})$: variation of yields of a crop given input vector $\boldsymbol{X}$,

$\boldsymbol{\beta}$: column vector of production coefficients.

The irrigation option is presumably adopted if the certainty equivalent associated with adopting the system minus the cost K of adoption exceeds the certainty equivalent associated with the case of no irrigation (Finger and Schmid, 2008, p. 29):

$$CE(I = 1) - K > CE(I = 0) \tag{6}$$

Not one, but two technological conditions such as the condition illustrated in (5) constrain the maximization problem in (4) in terms of the model implementation (Finger and Schmid, 2008, p. 30). Those are econometrically estimated agronomic production functions, which reflect the managerial option to adopt an irrigation system or not. Irrigation water thus may or may not be considered a relevant input in production. The empirical specification, which corresponds to the case of both nitrogen and irrigation water being considered as relevant, reads:

$$Y = \alpha_0 + \alpha_1 N^{1/2} + \alpha_2 W^{1/2} + \alpha_3 N + \alpha_4 W + \alpha_5 (N * W)^{1/2} + e \tag{7}$$

where

Y: crop yields,

$\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$: production coefficients,

N: amount of the input nitrogen,

W: amount of the input irrigation water,

e: error term.

In terms of implementation the maximization problem in (4) still requires calibration with respect to the expected yields $E[Y(\boldsymbol{X})]$ and the yield variation $\sigma_Y(\boldsymbol{X})$. To obtain those the authors employ the crop simulation model Cropsyst. The expected yields $E[Y(\boldsymbol{X})]$ are assumed equal to the results from the production functions estimation.

The yield variation measure $\sigma_Y(\boldsymbol{X})$, not to be understood as a variance in the strict sense, is assumed equal to the absolute difference between yields generated by Cropsyst with varying weather conditions and the expected yields (Finger and Schmid, 2008, p. 29):

$$\sigma_Y(\boldsymbol{X}) = |Y(\boldsymbol{X}) - E[Y(\boldsymbol{X})]| \tag{8}$$

This work illustrates the potential of estimating a function and using the parameter estimates to form an empirically implied technological condition in a farm level optimization model. The contribution, however, empirically relies on a crop simulator and an agronomic production function, which considers nitrogen and irrigation water as the only relevant inputs in production. The crop yields are thus reduced to products of the biophysical environment, which might be appropriate on a field level. It should be investigated if adding up the predicted harvest of individual fields approximately matches the reported yield on enterprise level, where other factors, such as labor costs might be at play.

2.1.3.2 Dono and Mazzapicchio: assessment of farmers' vulnerability to changing precipitation patterns

Dono and Mazzapicchio use discrete stochastic programming techniques to identify farm types most vulnerable to changes in precipitation and water availability patterns due to their reliance on irrigation. The changes in precipitation patterns specifically dealt with in the study are the decreasing annual cumulative precipitation and the increasing monthly rainfall volatility (Dono and Mazzapicchio, 2010, p. 366). The changes in rainfall regime would presumably influence the availability of irrigation water collected in a dam, making the water level tendentially lower and less certain. The uncertainty regarding irrigation water availability is resolved after winter rainfall and thus in the midst of the agricultural cycle of the farmers, whose vulnerability is to be assessed, see Figure 2.

The first stage decisions of the farmer regard the autumn decisions on the amount of land devoted to the production of winter crops, which unlike spring crops do not depend on irrigation and therefore do not depend on the amount of irrigation water that would be eventually available after winter rainfall. The second stage decisions of the farmer regard the spring-summer decision round after the winter rainfall has occurred (Dono and Mazzapicchio 2010, p. 364).

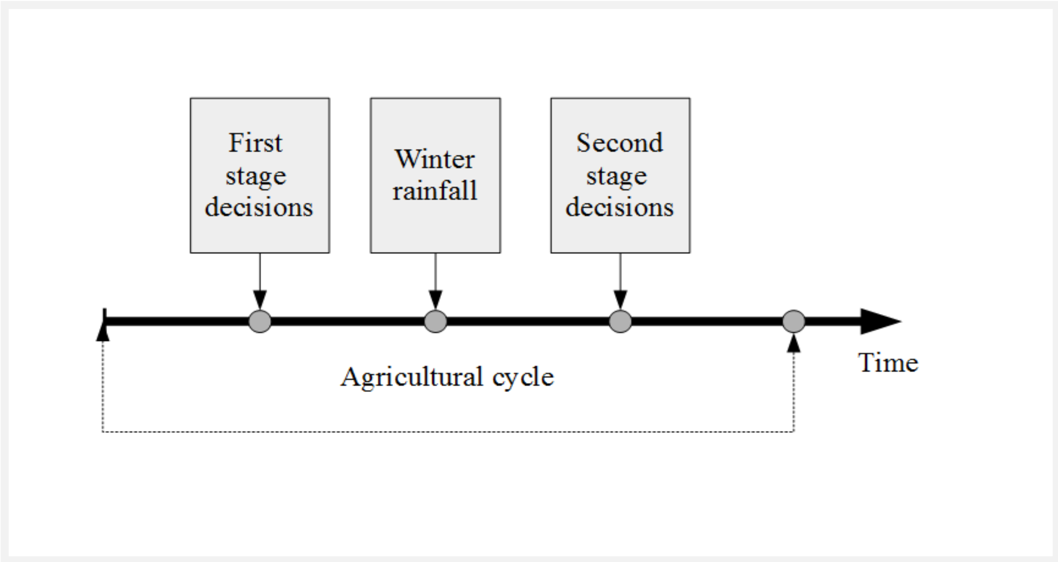


Figure 2 Timeline of the model Dono and Mazzapicchio.
Source: own illustration.

On the one hand the second stage decisions regard the decision on whether or not to cultivate the land eventually unused due to first stage decisions, which have led to *ex post* suboptimal results. On the other hand the second hand decisions regard the specific spring crop to grow.

The farmer is assumed to make an *ex ante* choice on both first and second state decisions by maximizing his total gross margin (adapted from Dono and Mazzapicchio, 2010, p. 365):

$$max_{\boldsymbol{X}}\, Z = \boldsymbol{Gl_1} * \boldsymbol{X_1} + \sum_{k=1}^{K} P_k * \boldsymbol{Gl_2} * \boldsymbol{X_{2,k}} \quad (9)$$

subject to:

$$\boldsymbol{A_1} * \boldsymbol{X_1} + \boldsymbol{A_2} * \boldsymbol{X_{2,k}} \leq \boldsymbol{b_k} \quad \forall k \quad (10)$$

$$\boldsymbol{X_1}, \boldsymbol{X_{2,k}} \geq \boldsymbol{0} \quad \forall k \quad (11)$$

where

Z: total gross margin,

k: state of irrigation water availability, $k \in \{1, \dots, K\}$,

$\boldsymbol{X_1}$: vector of first stage activities (for instance growing wheat),

$\boldsymbol{X_{2,k}}$: matrix of second stage activities,

P_k: perception of the farmer with respect to the occurrence probability of k,

$\boldsymbol{Gl_1}, \boldsymbol{Gl_2}$: unitary gross margins,

$\boldsymbol{A_1}, \boldsymbol{A_2}$: matrices of technical coefficients, which translate activities into production outcomes,

$\boldsymbol{b_k}$: vector of resource availability for each k,

1, 2: decision stages.

The maximization constraints given in (10) and (11) regard the expected land and labor availability, agronomical practices safe-guarding a lack of decline in crop yields, agricultural policies concerning set-asides and production quotas as well as production constraints regarding livestock and plant production. Constraints on irrigation water availability are defined for the presumably three possible states of water accumulation as well (Dono and Mazzapicchio, 2010, p. 365).

The maximization of the total gross margin in (9) is made by the farmer *ex ante*, which makes it necessary for the authors to make an assumption on how farmers construct their expectations regarding the occurrence probability of different water accumulation states. The expectations are presumably based on past irrigation water availability. The authors then investigate the effect these expectations have on the production decisions in the two stages, and thereby on farm income (Dono and Mazzapicchio, 2010, p. 361).

The influence of the expectations on farm income is especially pronounced if the expectations regarding water availability and the *ex post* water level in the dam diverge: the expectations influence the amount of land devoted to the production of spring crops, which heavily rely on irrigation. Through influencing the amount of land devoted to the production of the spring crops the expectations indirectly determine the amount of land devoted to the production of winter crops, which do not rely on irrigation (Dono and Mazzapicchio, 2010, p. 364). The uncertainty regarding water availability thus influences both first and second stage decisions.

The baseline scenario is assumed to consist of the average outcomes of the model based on expectations about water accumulation state occurrence as implied by the most current water accumulation series, with the outcomes weighted by the probabilities of state occurrence (Dono and Mazzapicchio, 2010, p. 366). A comparison of the baseline scenario with soil usage data available for the region confirms a more than 91.9% match, which leads the authors to conclude that the model reliably reflects the production choices made by farmers during the period (Dono and Mazzapicchio, 2010, p. 366).

This model clearly demonstrates the importance farmers' expectations regarding future resource availability could have in production decisions. It also highlights the potential economic losses associated with farmers' expectations on future water availability and the actual future water availability diverging. In order to further evaluate the model it is vital to consider how the technical coefficients, which translate first stage activities into production outcomes, are obtained. In case crop yields as the product of the first stage activity "growing wheat" are obtained by a crop simulator, the potential bias associated with reducing crop yields to mere products of the biophysical environment applies.

2.2 Remarks on the literature overview

The Ricardian approach and the panel data analysis proposed by Deschênes and Greenstone provide crude estimates of climate change impact after the farmers have presumably adapted. Both rely on making an inference about future agricultural outcomes by econometrically investigating current agricultural outcomes across a geographically and climatically diverse area.The reliability of both approaches rests on the geographical and climatic span of the specific study being sufficiently wide to guarantee the inclusion of extreme climate conditions in the dataset. Neither targets to explain the gradual adaptation of farmers to environmental changes, which could take the form of farmers arriving from one set of output choices and input commitments to another set of output choices and input commitments in order to accommodate the environmental changes.

The estimates of the Ricardian approach and the panel data analysis proposed by Deschênes and Greenstone can nevertheless provide a reference point with respect to the estimated changes in a quantity of interest, for instance cumulative agricultural profits in a county at a specific point of time in the future. The cumulative agricultural profits in a county are calculable for any point in the future by weighting and adding up the profits of multiple geographically and climatically diversified farm-level programming models in the county, which, as mentioned, have a limited spatial span, but a temporal dimension and can incorporate environmental uncertainty in the decision-making environment of the individual farmer. Changes in the cumulative agricultural profits in a county can thus be calculated using the ensemble of suitably diversified farm-level programming models as differences in the volume predicted by the ensemble across time.

Thus, *(i)* the estimated changes in cumulative agricultural profits in a county as given by the Ricardian approach, *(ii)* the estimated changes in cumulative agricultural profits as given by the panel data analysis of Deschênes and Greenstone and *(iii)* the estimated changes in cumulative agricultural profits in a county as provided by summing up the estimates of farm-level programming models, can be compared if the underlying assumptions about future price developments are kept identical. Comparability in the climate change impact estimates of different economic models across the scales of economic analysis can thus be established.

The farm-level programming models, as exemplified by the contributions of Finger and Schmid (2008) and Dono and Mazzapicchio (2010), accommodate price and environmental uncertainty in the decision-making environment of the individual farmer rather well. The models specifically target representing the gradual adaptation of farmers to environmental changes, which could take the form of farmers arriving from one set of output choices and input commitments to another set of output choices and input commitments in order to accommodate the environmental changes.

The farm-level programming models, however, utilize an agronomic notion of production. They consider crop yields as mere products of the biophysical environment, which can be inferred by the calibration of the models using crop simulators on the level of an individual field. The predicted crop yields on the enterprise level are thus obtained by adding up the predicted yields from individual fields. The models might thus be neglecting factors, which might be relevant to the formation of crop yields on the enterprise level, for instance labor.

The latter observation raises the question if there could be a normative or positive advantage in integrating different perspectives on the formation of crop yields. Combining the agronomic perspective of crop yields as mere products of the biophysical environment, on the one hand, and the agricultural production economic perspective of crop yields as outputs resulting from a farmer committing inputs to the production process, on the other hand, can prove useful.

In order to contribute towards addressing this issue, this thesis proposes a novel enterprise-level model of crop yields, which integrates the agronomic and production economic perspectives on crop formation. The model is described in chapter 5. The predictive performance and the climate change impact estimates of the proposed model are directly comparable to the predictive performance and impact estimates provided by a farm-level programming model. It should be kept in mind that the impact estimates of multiple geographically and climatically diversified models like the model proposed in chapter 5 can be summed up. The comparability to *(i)* the estimated changes in cumulative agricultural profits in a county as given by the Ricardian approach and *(ii)* the estimated changes in cumulative agricultural profits as given by the panel data analysis of Deschênes and Greenstone can thus be established.

3 Theoretical background

The theoretical background consists of two theoretical concepts: the state-contingent approach to production and decision-making under uncertainty, proposed by Robert Chambers and John Quiggin (Chambers and Quiggin, 2000) as well as the hypothesis of adaptive expectations. The state-contingent approach is the origin of the notions of a state-contingent production function, an effort-cost function and a revenue-cost function; the hypothesis of adaptive expectations introduces the concept of price expectation formation. Those concepts are illustrated in the left hand side of Figure 1.

3.1 The state-contingent approach to production analysis and decision-making

This section introduces the state-contingent approach to production and decision-making under uncertainty (Chambers and Quiggin, 2000). The description follows the brief overview given in a previous author's work (Angelova, 2014).

The approach involves describing the future after an uncertain event as production outcome vectors $\boldsymbol{y_s}$, each one assigned to one of a finite number of mutually exclusive so-called states of nature s (s belongs to the space of states of nature Ω, $\Omega = \{1, \dots, S\}$) (Chambers and Quiggin, 2000, p. 41).

A state-contingent production technology maps the vector of inputs $\boldsymbol{x} \in \mathbb{R}_+^N$ onto the output matrix $\boldsymbol{y} \in \mathbb{R}_+^{M \times S}$ where m (m = 1,..., M) identifies the output and s (s = 1,…, S) is a state of nature. An element of $\boldsymbol{y}$, y_{ms}, reflects the amount of output m that could be produced in s. A force referred to as Nature is assumed to make a draw from Ω, leading to a realization of a state of nature s. Only a single column of $\boldsymbol{y}$ then occurs corresponding to the state of nature s: $\boldsymbol{y_s} \in \mathbb{R}_+^M$.

A state-contingent production technology can be expressed in terms of sets by the formulation $\boldsymbol{x} \in \mathrm{X}(\boldsymbol{y})$. This should be interpreted as the input vector $\boldsymbol{x}$ belonging to the set of input vectors sufficient to produce the output matrix $\boldsymbol{y}$. Alternatively, the technology can be expressed as $\boldsymbol{y} \in \mathrm{Y}(\boldsymbol{x})$, which should be interpreted as the output matrix $\boldsymbol{y}$ belonging to the set of output matrices, which can be produced given the input vector $\boldsymbol{x}$.

The technology is functionally representable as a distance function:

$$\mathrm{O}(\boldsymbol{y}, \boldsymbol{x}) = inf\left\{\theta > 0 : \frac{\boldsymbol{y}}{\theta} \in \mathrm{Y}(\boldsymbol{x})\right\} \quad (12)$$

where

$\mathrm{Y}(\boldsymbol{x})$: set of output matrices, which can be produced with the input vector $\boldsymbol{x}$.

Each specific state of nature *s* is presumably subjectively perceived by an optimizing agent as occurring with probability π_s. The agent is assumed to be able to adjust the production efforts in order to *ex ante* maximize his utility given the technological and cost conditions he faces. A specific pattern of *ex ante* input commitment would produce a specific set of possible future outcomes. The adaptation consists in redistributing the *ex ante* input commitment in order to "substitute" potential future outcomes against each other.

An inability to adapt to uncertainty is accounted for as the extreme case of a so-called output cubical technology, a technology which would not allow for a substitution of state-contingent outputs by rearranging inputs *ex ante* (Chambers and Quiggin, 2000, p. 66). Thus, if faced with an output-cubical technology, the agent would be incapable of adapting. It can be argued that a stochastic production function formulation would sufficiently account for uncertainty in an empirical investigation if the agent is faced with an output cubical technology. A state-contingent formulation would nevertheless be formally correct in this case.

The state-contingent approach to production and decision-making under uncertainty, however, provides an additional modeling advantage compared to more traditional approaches to production analysis under uncertainty such as stochastic frontier analysis, namely being able to provide well-defined cost functions (Chambers and Quiggin, 2000, p. 9). This advantage is achieved by extending the formulation of production technology to a correspondence between inputs and *potential* outputs rather than understanding it as a correspondence between inputs and stochastically up- or downscaled outputs. The properties of the resulting cost functions are largely identical with those of cost functions in the deterministic case (Chambers and Quiggin, 2000, p. 125).

A so-called *effort-cost function*, which corresponds to a state-contingent production technology, however, has a slightly different interpretation than the cost function in the deterministic case. It is understood not as costs related to a certain output level, but rather as costs incurred in order to arrange *ex ante* for a specific set of possible future production outcomes (Chambers and Quiggin, 2000, p. 125).

In the case of linear input pricing the costs associated with arranging for a set of possible future production outcomes are:

$$\boldsymbol{w}^T\boldsymbol{x} \tag{13}$$

where

$\boldsymbol{x}$: input vector,

$\boldsymbol{w}$: vector of input prices.

The effort-cost function is defined as the optimal value function in the following minimization problem:

$$\mathrm{c}(\boldsymbol{w},\boldsymbol{y}) = \min_{\boldsymbol{x}}\{\boldsymbol{w}^T\boldsymbol{x}: \boldsymbol{x} \in \mathrm{X}(\boldsymbol{y}), \boldsymbol{w} \in \mathbb{R}^N_{++}\} \tag{14}$$

where

$\boldsymbol{y}$: output matrix.

As can be seen by the formal definition, it resembles a cost function and involves the choice of the optimal input vector under a technological condition. This input vector follows from the cost-minimization postulate and is thus independent of the risk preferences exhibited by the agent himself (Chambers and Quiggin, 2000, p. 127).

Apart from providing a cost function corresponding to a state-contingent production technology, the state-contingent approach contributes towards linking the subfields of production analysis and decision-making under uncertainty. It does so by developing the concept of a *revenue-cost function*, denoted by $\mathrm{C}(\boldsymbol{w},\boldsymbol{r},\boldsymbol{p})$, which involves production decisions over input vectors as well as potential output matrices to achieve at least a certain state-contingent monetary goal (Chambers and Quiggin, 2000, p. 143).

If diag is defined as an operation, which extracts the diagonal of a matrix and transforms it into a column vector, then the revenue-cost function is given by:

$$\mathrm{C}(\boldsymbol{w},\boldsymbol{r},\boldsymbol{p}) = \min_{\boldsymbol{y}}\{\mathrm{c}(\boldsymbol{w},\boldsymbol{y}) : \mathrm{diag}(\boldsymbol{y}^T\boldsymbol{p}) \geq \boldsymbol{r}\} \tag{15}$$

where

$\boldsymbol{p}$: $M \times S$ matrix of state and output specific prices,

$\boldsymbol{r}$: S dimensional column vector of state-contingent revenues.

The objective function in the minimization problem is the effort-cost function, which incorporates the costs associated with *ex ante* arranging for certain output matrix $\boldsymbol{y}$ and thus already takes into account the technological limitations posed by the state-contingent production technology.

As already stated, the agent is assumed to choose a monetary minimum, which he wishes to achieve when choosing the output matrix $\boldsymbol{y}$, while being aware of the matrix of state and output specific prices. This monetary minimum is expressed through the target revenue vector $\boldsymbol{r}$. The revenue-cost function $\mathrm{C}(\boldsymbol{w},\boldsymbol{r},\boldsymbol{p})$, which results from minimizing the effort-cost function with respect to the output matrix $\boldsymbol{y}$ under the target revenue condition, thus describes a set of production alternatives expressed in monetary terms. This set of production alternatives is determined by the production technology, the input regime and prices, the output prices in the different states of nature and the target revenues in each state.

The specific point in the set of production alternatives, which ends up being chosen, depends on the risk preferences of the agent. Two polar cases with respect to risk aversion in the case of two states of nature, assuming an efficient agent, are depicted in Figure 3. The common element between the two extremes, however, is the production decision being described by the point where the iso-cost curve and the indifference curve touch, but do not cross.

The first extreme is marked by a white dot, which depicts the production decision of the risk neutral producer. The indifference curve of the agent coincides with the so-called fair-odds line, which in the two dimensional case would be expressing the ratio between the subjective probabilities π_s of the occurrence of the two states of nature (Chambers and Quiggin, 2000, p. 166). The other extreme, marked by a black dot, represents the production decision of the extremely risk averse producer, whose preferences can be described by a maxi-min shape. Preference curves with this shape will always touch but not cross, the iso-cost curve, which engulfs the set of production alternatives at the point where the bisector crosses the iso-cost curve (Chambers and Quiggin, 2000, p. 177).

It is important to note that the extremely risk-averse producer *might* perceive the same fair-odds line the risk-neutral producer does. It is not necessarily the deviating perception of state occurrence probabilities, which drives him to choose points along the bisector where the state-contingent revenues are equal, but rather his extreme preferences with respect to the risk posed by state-contingent revenue fluctuations. Fully operationalizing the state-contingent approach, given the assumption that production choices are made *ex ante*, requires the involvement of two hypotheses. The first hypothesis is on how the producer constructs output price expectations. The second one is on how the producer constructs his perception of the likelihood of the nature-states occurring.

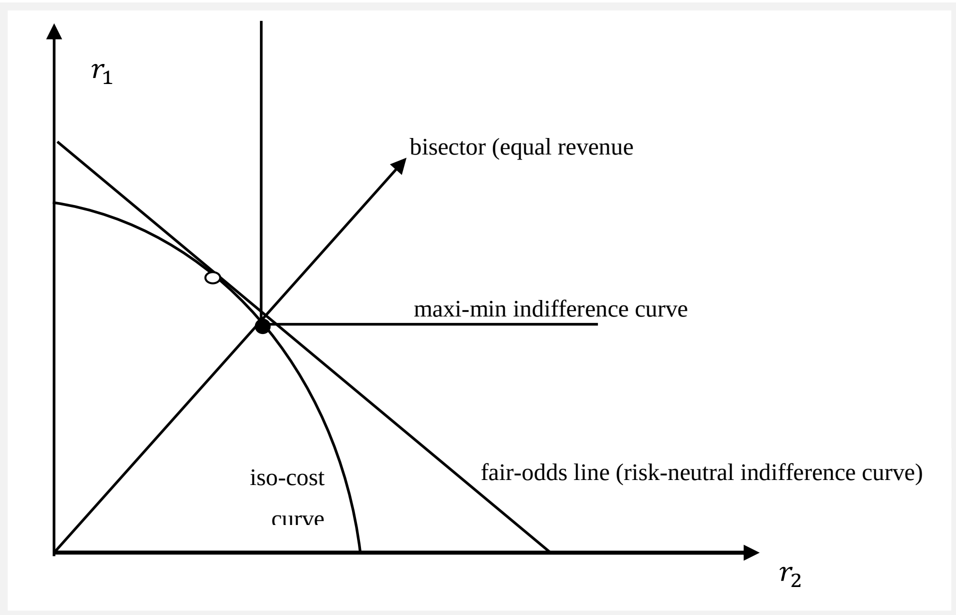


Figure 3 The production decision with two states of nature.

Note: The production decision of the extremely risk-averse producer is depicted by the black dot, the production decision of the risk neutral producer is given by the white dot.

Source: Chambers and Quiggin (2000), page 179.

3.2 The hypothesis of adaptive expectation formation

The hypothesis of adaptive expectation formation constitutes one of the prominent approaches to modeling expectation formation of economic agents, next to the naïve, or static, expectation hypothesis and the hypothesis of rational expectations (Evans and Honkapohja, 2001, p. 10). The article by Parkin (Parkin, 2008) in the New Palgrave Dictionary of Economics traces the origin of the hypothesis back to the early works of Fischer (Fischer, 1911). Evans and Honkapohja trace the origins of the hypothesis back to work of Fischer from the 1930s (Fischer, 1930), and the first mathematical formalizations of the hypothesis as appearing in the 1950s (Cagan, 1956), (Friedman, 1957), (Nerlove, 1958).

The idea is simple: economic agents form their predictions about the future levels of a certain economic variable by observing past realizations of the variable. The hypothesis of adaptive expectation formation is still relevant to macroeconomic price expectation modeling, as a recent contribution by Chow demonstrates (Chow, 2011).

In terms of price expectations the hypothesis reads (Evans and Honkapohja, 2001, p. 10):

$$p^e(t) = p^e(t-1) + \lambda\big(p(t-1) - p^e(t-1)\big) \tag{16}$$

where

$p^e(t)$: expected price in period *t*,

$p^e(t-1)$: expected price in period *t*-1,

$p(t-1)$: price in period *t*-1,

λ: expectation adjustment coefficient, $\lambda \in (0,1)$.

The expectation adjustment parameter λ reflects the significance the agent attaches to deviations between the price observed in the last period and his preceding expectations concerning the price in the last period. The parameter is assumed innate to the agent.

3.3 Remarks on the theoretical concepts

One of the advantages of the state-contingent approach consists in unifying the fields of production analysis under uncertainty and decision-making under uncertainty. This advantage motivates the adoption of the approach as a theoretical basis in the present work. There is, however, another advantage of the approach, which is essential from the perspective of production economics. It consists in delivering well-defined dual functions corresponding to stochastic production technologies. The significance of this fact requires clarification with respect to the evolution of the field of production economics.

Duality, broadly speaking, is a concept originating in mathematical optimization theory, which has been adopted in the field of production economics among others. The concept conveys the principle that optimization problems can be seen from two perspectives, a primal perspective and a dual perspective. The values of the solutions, which are delivered by the primal and the dual formulation of an optimization problem, are mathematically connected.

One of the first applications of the duality principle to the field of theoretical production economics was made by Shephard (Shephard, 1953), where he demonstrated the duality between input-distance functions and cost functions in a deterministic framework. As Färe and Primont note, such a relationship connects two models: the first (primal) model, exemplified by an input-distance function, represents the technological connection between inputs and outputs and is free of behavioral assumptions about the firm operating under these technological restrictions; the second (dual) model, the cost function, represents the costs of producing a specific vector of outputs under the technological restrictions, an input pricing regime and with the use of cost-minimizing quantities of inputs. The assumption of cost-minimizing behavior by the firm is thus implied in the second model (Färe and Primont, 1994, p. 2). The form of the cost function, however, vitally depends on the technological restrictions the firm operates under.

The contribution of Shephard is appreciated in applied production analysis since it allows empirical production analysts some flexibility regarding the specific model to estimate in the case of data limitations (Färe and Primont, 1994, p. 4). Applying duality theory to theoretical production analysis under uncertainty has been an open field of research in order to allow for similar flexibility in the presence of stochastic factors (Chambers and Quiggin, 1998). A solid theoretical framework, which applies duality theory to production analysis in the presence of risk and uncertainty, has been delivered with the state-contingent approach, hence the significance of the contribution.

This thesis, however, adopts the state-contingent approach as a tool for *both* production analysis under uncertainty and decision-making under uncertainty. Additional assumptions on two matters are thus required due to production choices being made *ex ante*. Firstly, a hypothesis on output price expectation formation is needed. The hypothesis on adaptive price expectation formation is chosen here for its conceptual and computational simplicity. Given the state-contingent nature of the output prices $\boldsymbol{p_{ms}}$, which form the output price matrix in the target revenue conditions described in (15), a slightly augmented version of the hypothesis in (16) is employed in order to accommodate state dependency:

$$p_{ms}^{e}(t) = p_{ms}^{e}(t_s) + \lambda\big(p_{ms}(t_s) - p_{ms}^{e}(t_s)\big), \tag{17}$$

where

$p_{ms}^{e}(t)$: expected price for output *m* in period *t* under *s*,

$p_{ms}^{e}(t_s)$: expected price for output *m* under *s* in period t_s, the last period state *s* occurred in,

$p_{ms}(t_s)$: price for output *m* under state *s* in period t_s, the last period state *s* occurred in,

λ: expectation adjustment coefficient, $\lambda \in (0,1)$.

The augmented definition does not take into account the time-lag between period *t* and the last period, in which state *s* was observed. It is worth noting that this is equivalent to assuming that the agent attaches the same significance to the monetary experiences he remembers regardless of when they happened. In other words, the agent either remembers the event or he does not. It is also worth noting that modifications of this rather strict assumption in the form of fading memories can be achieved by reducing the influence experiences in the distant past have on the agent's current decision-making. Such a tuning can be technically achieved by appropriate weighting.

Secondly, a hypothesis on the formation of nature-state occurrence probability perception is needed. Logical consistency requires a hypothesis on the way the agent constructs expectations regarding how often every specific nature state s occurs, which does not contradict (17). It is therefore assumed that the agent constructs his probability perception based solely on what he has observed in his past and attaches no significance to when in his past the states have occurred. Again, reducing the gravity which experiences in the distant past would have on the current decision-making, can be achieved by weighting. For the sake of simplicity, the current work assumes the agent perceives a nature-state as occurring with a probability equal to the relative frequency of the state occurring in his past.

4 Empirical studies based on the state-contingent approach

This chapter reviews empirical investigations based on the state-contingent approach to production and decision-making under uncertainty outlined in section 3.1. The relevance of the contributions in the literature overview has been decided upon based on the prominence of the studies in the field of empirical state-contingent production analysis, which is relatively easy to determine due to the novelty of the state-contingent modeling framework and the relative lack of implementations based on it. As Nauges, O'Donnell and Quiggin put it, the "…empirical implementation of the state-contingent approach is still in its infancy" (Nauges, O'Donnell and Quiggin, 2011, p. 3). A literature overview of the approaches is nevertheless essential in order to identify the type of models, which the state-contingent approach can support as a theoretical basis.

This literature overview is divided into two parts in order to reflect a difference in how the reviewed models utilize the state-contingent approach. On the one hand, there are models where the state-contingent approach is used as a tool for production analysis under uncertainty and optimization by the farmer is either disregarded or implied. Here these models are referred to as econometric studies, which are exemplified by the studies described in subsections 4.1.1 and 4.1.2. For a comparison with alternative approaches such as stochastic frontier analysis the literature review of Shankar can be consulted (Shankar, 2012). An overview of alternative methods for production analysis under uncertainty is not included here since the only alternative method, which can model environmental uncertainty (the stochastic frontier analysis) can be seen as a special case of the more general state-contingent approach.

There are, on the other hand, models where optimization by the farmer is explicit. Here these models are referred to as mathematical programming models, which are exemplified by the study presented in 4.2. In contrast to the contributions in subsections 4.1.1 and 4.1.2, the study in section 4.2 utilizes the state-contingent approach as a tool of decision-making under uncertainty.

4.1 Econometric studies

This section describes two production economic studies either implying or disregarding optimization by the farmer. One of them is the contribution by Chavas, providing a methodology for simulating non-realized potential outputs (Chavas, 2008). Such a simulation is necessary for the estimation of the parameters of a state-contingent dual, in this case a cost function, which would correspond to a state-contingent production technology. The cost function is estimated for US agriculture. The second study by Nauges, O'Donnell and Quiggin proposes a constant elasticity of substitution formulation of the state-contingent production technology (Nauges, O'Donnell and Quiggin, 2011). The contribution illustrates the necessity to detect the occurrence of the states of nature in order to estimate a primal representation of a state-contingent production technology. Their work examines the production in the Finnish agricultural sector.

4.1.1 Chavas: estimating an *ex ante* cost function

The study of cost-minimizing input decisions under uncertainty by Chavas constitutes one of the earliest empirical implementations of the state-contingent approach (Chavas, 2008). The article delivers a methodological contribution with respect to recovering the non-realized potential outputs, which is necessary for the estimation of the parameters of a dual representation of a state-contingent production technology, in this case the *ex ante* cost function.

In other words, the contribution addresses the fact that under a state-contingent technology, where inputs are linked to *potential* outputs occurring in mutually exclusive states of nature, only a single potential output vector is realized *ex post*, which leads to the corresponding cost and input demand functions being empirically intractable (Chavas, 2008, p. 438).

The basic idea of the methodology for the *ex ante* recovery of the parameters of a state-contingent production technology is to simulate the potential outputs that would have occurred in non-realized states of nature around realized observations of a firm's output $(y_{1t,} \dots, y_{mt})$ for T periods, $t = 1, \dots, T$. It is assumed that the unobservable variable e_{ms} captures the relative changes in output m between the states of nature (Chavas, 2008, p. 439). The *ex ante* output y_{mst} is expressed as a function of e_{ms}.

A presumably observable auxiliary variable, which is a function of the unobservable relative changes in the m'th output e_{ms}, is introduced. This functional relationship is presumably invertible. The unobservable relative changes e_{ms} can thus be expressed in terms of the observable auxiliary variable. This would allow the recovery of the realizations of output m in the non-realized states of nature, since the *ex ante* output y_{mst} can presumably be expressed as a function of e_{ms}.

The proposed scheme is used by Chavas to estimate input demand functions derived from a generalized Leontief cost function. A single output and two states of nature, ‘good weather’ and ‘bad weather’, are considered in the analysis of aggregated data from US agriculture for the period 1949 to 1999. A crop yield index is employed as an instrument for the observable auxiliary variable.

Based on the estimates the elasticity of transformation is calculated around the sample means and found to be near zero, which the author concludes indicates an output cubical technology, i.e. a technology, which does not allow for a substitution of the state-contingent outputs between the states of nature.

Chavas acknowledges the possible limitations to the analysis in stating that it holds independently of the nature of the state-contingent production technology only to the extent to which a specific parameter of the model can be said to reflect the nature of the technology and the economic risks associated with it (Chavas, 2008, p. 440). The contribution exemplifies the distributional assumptions needed in order to econometrically approach the dual, rather than the primal state-contingent problem. An econometric investigation of the latter is given by the study described in subsection 4.1.2.

4.1.2 Nauges, O’Donnell and Quiggin: estimating a state-contingent production function

The contribution of Nauges, O’Donnell and Quiggin is one of the first empirical studies dealing with the primal state-contingent problem (Nauges, O’Donnell and Quiggin, 2011). The authors propose a constant elasticity of substitution (CES) type functional form for the state-contingent production technology and subsequently investigate the production of the Finnish crop production sector. The hypothesis of an output-cubical state-contingent production technology is tested. The proposed functional form reads:

$$y_s = A_s \left[\theta^b x^b + \delta_s^b x_s^b + \sum_{k=1}^{K} \gamma_k^b z_k^b \right]^{\varphi/b} \tag{18}$$

where according to the description given by the authors

y_s: state-contingent output in state of nature $s, s \in \{1, \dots, S\}$,

A_s: state-contingent technical parameter,

x_s: state-allocable input,

x: sum of x_s over all states s, non-state allocable,

z_k: non-state allocable input, $k \in \{1, \dots, K\}$,

b: substitution parameter,

φ: coefficient reflecting the returns to scale.

In the original CES functional form, the coefficients associated with the inputs, namely θ, δ and γ_k for every k, should add up to unity (Arrow et al., 1961). b is a function of the elasticity of substitution and is referred to as substitution parameter by Arrow et al. (Arrow et al., 1961, p. 230). The magnitude of φ>0 reflects the returns to scale, with φ<1, φ=1, φ>1 reflecting the cases of increasing, constant and decreasing returns to scale respectively. In the formulation (18) by Nauges, O'Donnell and Quiggin the case δ_s=0 would constitute an output cubical production technology. The functional form is claimed to accommodate a variety of interesting functional forms as limiting cases. A slight correction is needed in order for this claim to be substantiated - the coefficients associated with the inputs in (18) should be factored by 1, instead of b:[1]

$$y_s = A_s \left[\theta x^b + \delta_s x_s^b + \sum_{k=1}^{K} \gamma_k z_k^b \right]^{\varphi/b} \tag{19}$$

For φ equal to 1 the functional form becomes:

$$y_s = A_s \left[\theta x^b + \delta_s x_s^b + \sum_{k=1}^{K} \gamma_k z_k^b \right]^{1/b} \tag{20}$$

[1] Gratitude is expressed to Jörg Gersonde and Michael Grings for uncovering inconsistencies in the formulation (18) by Nauges, O'Donnell and Quiggin.

(20) can be log-linearized to obtain:

$$ln(y_s) = ln(A_s) + {}^{1}/_{b}\, ln\left(\theta x^b + \ \delta_s x_s^b + \sum_{k=1}^{K} \gamma_k z_k^b\right) \tag{21}$$

Applying the rule of l'Hôpital yields:

$$\lim_{b \to 0} ln(y_s) = ln(A_s) + \theta ln(x) + \delta_s ln(x_s) + \sum_{k=1}^{K} \gamma_k ln(z_k) \tag{22}$$

For $b \to 0$ the functional form in (19) thus converges to a Cobb-Douglas form.

In the empirical analysis of the Finnish crop production sector Nauges, O'Donnell and Quiggin account for technical inefficiency, i.e. the eventuality that individual firms produce less than what is technologically feasible and the corresponding observations lie underneath the efficient frontier. The empirical part of the contribution accounts for the inputs labor, capital and intermediate inputs as non-state allocable inputs and the input land devoted to the production of a certain crop as the only state-allocable input. This fact seems more intuitive after the authors introduce the definition of nature-states underlying their empirical analysis. The issue of the nature-state definition will be addressed in this subsection after the presentation of the empirical specification.

The corrected empirical specification of Nauges, O'Donnell and Quiggin after log-linearization and after accounting for possible technical inefficiency and stochastic errors should read:

$$ln(y) = \sum_{s=1}^{S} e_s ln(A_s) + \left(\frac{\varphi}{b}\right) ln\left(\theta x^b + \sum_{s=1}^{S} e_s \delta_s \ x_s^b + \sum_{k=1}^{K} \gamma_k \ z_k^b\right) + v - u \tag{23}$$

where

y: observed output aggregate, which is synonymous with y_s once the state s is realized,

x_s: land devoted to producing a certain crop,

x: land devoted to crop production regardless of the crop,

z_k: labor, capital and intermediate inputs, $k \in \{1, \dots, K\}$,

e_s: dummy variable, taking the value of 1, should state s be realized, and 0 otherwise,

v: i.i.d. noise, $v \sim N(0, \sigma_v^2)$,

u: i.i.d. half-normally distributed variable, capturing technical inefficiency, $u \sim HN(0, \sigma_u^2)$.

Defining the states of nature and finding an indicator to detect their occurrence is central to the empirical part of the contribution of Nauges, O'Donnell and Quiggin. It plays a role when justifying the choice of land devoted to producing a certain crop as a state-allocable output and serves as a basis for constructing the dummy variables e_s. The states of nature are assumed synonymous with weather conditions beneficial to growing a specific crop. The crops being considered are wheat, oats and barley. Following the advice of Finnish grain experts the occurrence of states of nature in the sample is detected on the grounds of two meteorological indicators: the starting date of the growing season and the cumulative rainfall in June. The opinions of the experts on weather conditions favoring the production of a specific crop, thus constituting the states of nature in the empirical part of the contribution of Nauges, O'Donnell and Quiggin, are summarized in Table 1.

Table 1 Expert opinion on the weather conditions constituting the states of nature, which are defined as weather conditions favorable for a specific crop.

	Rainfall: low	Rainfall: average	Rainfall: high
Starting date: early	Barley-favorable	Oats-favorable	Barley-favorable
Starting date: average	Wheat-favorable	Wheat-favorable	Wheat-favorable
Starting date: late	Barley-favorable	Wheat-favorable	Wheat-favorable

Source: own illustration based on Nauges, O'Donnell and Quiggin, 2011, p. 8.

It is evident that the specification in (23) poses a non-linear estimation problem. The authors leave the estimation of the parameter b for further research and focus on several special cases that are thereafter compared to each other in terms of goodness of the statistical fit.

The bookkeeping records in the southern regions of Finland between the years 1998 and 2003 are matched with weather data concerning the precipitation and the starting date of the growing season, which is collected at geographically appropriate weather stations. The hypothesis of an output-cubical state-contingent technology (i.e. $\delta_s = 0, \forall s$) is rejected.

The analysis of Nauges, O'Donnell and Quiggin is tuned towards the question of whether or not the production technology in the Finnish grain producing sector is output-cubical, i.e. whether the technology of grain producers allows for adaptation through input reallocation. The contribution outlines the necessity to introduce a source of information to identify the occurrence of the states of nature in the econometric analysis of a primal state-contingent representation. The necessity arises because of the need to construct dummies for state occurrence in order to detect differences in the productivity of the state-allocable input in case the nature-state the input is allocated to occurs. The authors use weather data to detect the occurrence of the states of nature and construct the dummies.

4.2 Mathematical programming, Crean et al.: representing climatic uncertainty in agricultural models

This section considers in more detail a study, which explicitly accounts for the optimization of the farmer. It implements the state-contingent approach in the context of climate change impact assessment as a farm level adaptation model (Crean et al., 2013).

Another contribution, which is not presented explicitly to avoid redundancy, is an exploratory application of the state-contingent approach in the context of mathematical programming by Berg (Berg, 2012). This exploratory study compares the optimization results, based on the state-contingent approach, to the results obtained by the application of a more traditional decision-making heuristics such as expected utility maximization. The study clearly establishes the applicability of the state-contingent approach as a tool for decision-making under uncertainty.

The state contingent approach is also applied as a modelling framework in environmental issues, more specifically for the analysis of potential cooperation gains through coordination of water use along the stream of the rivers in the Murray-Darling basin in Southeastern Australia (Adamson, Mallawaarachchi, and Quiggin, 2007). The difficulties posed by conflicting demands for the already scarce water, environmental and consumptive, as well as the tension between the potential consumptive uses of water, are exacerbated by the geographically specific tendency towards dry land salinity and acidic soils. Using the water for irrigation purposes upstream injects salt loads into it, which in turn reduces the water's usability down the stream (Adamson, Mallawaarachchi, and Quiggin, 2007, p. 263). The gains of coordination or centralized resource management are demonstrated by means of comparing the solutions to a cascade of farm-level programming models, which depicts the river system as a directed network. The potential gains of cooperation are demonstrated by comparing different solutions of the optimization problem: a sequential solution, where optimization is seperately undertaken in every region starting with the region furthest up the river stream, and a global network solution, where the optimization problem is framed in a dynamic optimization fashion. The contribution demonstrates the applicability of the state-contingent approach as a tool for decision-making in a dynamic programming setting as well as its applicability to real world problems, such as natural resource management.

Downwards the focus lies on the application of the state contingent approach to climate change impact assessment. The contribution of Crean et al. explicitly acknowledges the sequential nature of risk in agricultural settings, which gives farmers the possibility to adapt to changes (Crean et al., 2013). It investigates the advantages of the state-contingent approach as a tool for decision-making in representing decisions under environmental uncertainty on a farm level compared to the representation delivered by the standard expected value approach. This is achieved by applying the state-contingent approach in a discrete stochastic programming setting by developing a two stage model of a representative mixed wheat-sheep farm. The optimal farm plans recommended by the model are subsequently compared to the recommended optimal farm plans an expected value model delivers (Crean et al., 2013, p. 360).

The discrete stochastic programming model represents a typical mixed production farm in the eastern part of New South Wales, Australia, with the farm's area presumably divided in fixed proportions between area for annual winter cropping, annual pasture and fallow land. The model is structured as a sequence of linear programming models, where the conditions in the second stage are dependent on the decisions in the first one and the state of nature in the first stage. The presumably risk-neutral farmer is assumed to maximize his expected net farm income $E[Y]$. To obtain the expected net farm income $E[Y]$ each state-dependent net farm income Y_s is calculated *ex post*. The expected net farm income $E[Y]$ is then calculated as the weighted sum of the net farm incomes Y_s:

$$max\ E[Y] = max\left(\sum_{s=1}^{S} \pi_s Y_s\right) = max\left(-\sum_{j=1}^{J} c_{1j} x_{1j} + \sum_{s=1}^{S} \pi_s \sum_{n=1}^{N} c_{2ns} x_{2ns}\right) \tag{24}$$

subject to

$$\sum_{j=1}^{J} a_{1ij} x_{1j} + \sum_{n=1}^{N} a_{2ins} x_{2ns} \leq b_i \quad \forall i, \forall s \tag{25}$$

$$\sum_{j=1}^{J} a_{1mjs} x_{1j} + \sum_{n=1}^{N} a_{2mns} x_{2ns} = 0 \quad \forall m, \forall s \tag{26}$$

$$x_{1j}, x_{2ns} \geq 0 \quad \forall j, \forall n, \forall s \tag{27}$$

where

π_s: probability of state s occurring,

c_{1j}: cost of growing crop or pasture j in stage 1,

a_{1ij}: quantity of resource i required by crop or pasture j in stage 1,

a_{1mjs}: quantity of output m produced by crop or pasture j in state s,

c_{2ns}: net revenue or cost from activity n in state s,

a_{2ins}: quantity of resource i required by activity n in state s,

a_{2mns}: quantity of output m required by activity n in state s,

b_i: availability of the resource i,

Y_s: farm income in state s,

x_{1j}: area of crop or pasture activity j grown in stage 1,

x_{2ns}: level of activity n chosen in state s in stage 2, e.g. selling crops or grazing crops.

The two stages of the model are evident from the objective function. The first term represents the variable cost associated with a certain activity x_{1j}, for example producing wheat. The second term reflects the expected state-contingent net revenue from the possible activities in the second stage, for example selling the wheat sawn in the first stage or grazing the wheat, which would not generate revenue but would eliminate second period costs associated with animal feed. The number of possible activities in the two stages of the model is different to reflect the fact that the possible activities are different. For instance, the option to sell crops is only relevant in the second stage of the model since the crops require time to grow.

The decisions between the two stages of the model are interrelated through the matrix of constraints with the decisions in the second stage dependent on the decisions from the first stage and the state of nature. Hence, the farmer would not have the option to sell wheat in the second stage, if he did not plant any wheat in the first stage.

Due to the mixed production nature of the farm two simulation models are employed to simulate the role of weather conditions, production and managerial decisions in forming outputs. A crop simulation model is run with nine sets of initial conditions, combinations of different planting dates (early, mid, late), which are considered a managerial decision, and starting soil moisture levels (low, average and high). Three states of nature have been defined based on the amount of growing season rainfall: dry, average and wet.

The advantages of state-contingent modeling are theoretically demonstrated with the help of a concept stemming from the field of stochastic programming, namely the concept of the value of a stochastic solution. Crean et al. also report that the discrete stochastic programming model based on the state-contingent approach provides a better fit to the actual farm data than the expected value model (Crean et al., 2013, p. 377).

The model clearly demonstrates the normative gains of the state-contingent approach to modeling decision-making under uncertainty in a farm level stochastic programming setting. While the study utilizes the state-contingent approach as a tool to decision-making rather than both decision-making and production analysis, the superior predictive power of the model, based on the state-contingent approach, with respect to real world farm data is noteworthy.

4.3 Remarks on the studies based on the state-contingent approach

The farm-level adaptation study of Crean et al. (2013) utilizes the state-contingent approach as a tool of decision-making under uncertainty. The contribution establishes the gains of modeling the specific farm-level optimization problem in a state-contingent way by calculating an index, which originates from the field of stochastic programming, namely the value of a stochastic solution. The state-contingent model introduced by Crean et al. also mirrors the farm-level data better than the alternative stochastic programming models applied by the authors.

Similarly to the study of Finger and Schmid (2008) and the study of Dono and Mazzapicchio (2010) described in subsection 2.1.3, the contribution Crean et al. relies on the understanding of crop yields as mere products of the biophysical environment. Again, the crop yields are calculated on a field level with the use of a crop model and obtained on the enterprise level by adding up the predicted yields from individual fields. The model of Crean et al. thereby suffers from the same potential bias, which results from neglecting factors that might be relevant to the formation of crop yields on the enterprise level, for instance labor.

The production economic studies of Chavas (2008) and Nauges, O'Donnell and Quiggin (2011), on the other hand, utilize the state-contingent approach as a tool for production analysis under uncertainty. These contributions illustrate the difficulties accociated with estimating the parameters of a state-contingent production technology. Both studies understand crop yields as outputs resulting from a farmer committing inputs to the production process under a specific production technology, thereby incorporating factors such as labor.

The immediate usability of the results of Chavas (2008) and Nauges, O'Donnell and Quiggin (2011) in the field of climate change impact assessment is restricted by one basic premise posed by the framework. It is the premise that the production technology, fully characterized by input and output quantities and the functional relationship between them, is objectively given and essentially invariant over time, with the possible exclusion of shifts induced by Hicks-neutral technical change (Hicks, 1932). It is evident, however, that from the production economic perspective in general the influence climate has on crop yields can only be accommodated in the functional relationship between inputs, e.g. land, and outputs, e.g. crop yields. Allowing for climate change would thus contradict the basic premise of a time-invariant production technology.

Accommodating the randomness introduced by climate change as a mere stochastic factor to the production technology, essentially falling back on stochastic frontier analysis, might seem intuitive in the case of such a contradiction. In the terms of section 3.3, however, this would lead to non-existent dual functions, e.g. non-existent cost functions. The latter implies an inability to give a recommendation on how farmers could minimize input commitment, for instance land input, in the case of climate change. This makes the approach less suitable for the field of climate change impact assessment.

The enterprise-level crop model proposed in chapter 5 overcomes the outlined challenges to the application of the state-contingent approach as a modeling tool for climate change impact assessment. This is achieved by integrating the agronomic perspective of crop yields as products of the biophysical environment, which underlies the study of Crean et al. (2013), and the production economic perspective of crop yields as outputs resulting from a farmer committing inputs to the production process under a specific production technology, which underlies the studies of Chavas (2008) and Nauges, O'Donnell and Quiggin (2011). The model fully operationalizes the state-contingent approach, i.e. uses it as a tool for *both* production analysis and decision-making under uncertainty, and delivers a theoretical contribution by unifying it with the theoretical concepts described in sections 3.2 and 3.3.

5 MELCY: Model for enterprise-level crop yields under climate change

A novel model for enterprise-level crop yields under climate change is described in this chapter. The model integrates the agronomic perspective of crop yields as products of the biophysical environment and the production economic perspective of crop yields as outputs resulting from committing inputs under a specific production technology. As it has been mentioned in section 4.3, the model delivers a theoretical contribution by unifying multiple theoretical concepts.

The model is constructed around the hypothesis that the enterprise level crop yields $y_{ms}^{(k)}(t)$ of crop *m* for a producer *k* in a certain period *t* under state of nature *s* can be meaningfully explained as a function of two quantities:[2]

- "Biophysical" hectare crop yields for crop *m* in period *t* under state of nature s. The *biophysical crop yields* are assumed to be reflective of the locally bounded climate effects, the combination of geographical location, soil type, long-term climate patterns and weather conditions, including weather extremes in *t*. Their formation is presumably free of anthropogenic influences.
- The amount of crop the optimizing producer *k* should be striving to achieve on an enterprise level in period *t* under state of nature *s* under a state-contingent production technology. The quantity is referred to hereafter as *production aspirations*. The production aspirations reflect *k*'s knowledge of environmental, technological and market conditions up to agricultural cycle *t* and the input prices in *t*. The knowledge of *k* is hereafter referred to as the information set $\Theta_{t-1}^{(k)}$. Some parts of the information set $\Theta_{t-1}^{(k)}$ are specific to producer *k*.

Formalizing the hypothesis and disregarding i.i.d. noise would yield:

$$y_{ms}^{(k)}(t) = f\left(y_{ms}^{biophy}(t), y_{ms}^{*(k)}\left(t|\Theta_{t-1}^{(k)}\right)\right) \tag{28}$$

where

$y_{ms}^{(k)}(t)$: crop yields of crop *m* of *k* in period *t* under state of nature s,

[2] A period is hereafter assumed synonymous with agricultural cycle. The latter is different to a calendar year.

$y_{ms}^{biophy}(t)$: biophysical crop yields of crop *m* in period *t* under state of nature s,

$y_{ms}^{*(k)}\left(t|\Theta_{t-1}^{(k)}\right)$: production aspirations of k with regard to crop *m* in period *t* under state of nature s given the information set $\Theta_{t-1}^{(k)}$,

$\Theta_{t-1}^{(k)}$: k 's information set containing environmental, technological and market conditions up to and including period *t-1*.

Producer k presumably perceives his crop production in the state-contingent sense: as outputs, which result from committing input factors under a state-contingent production technology. The production decision is assumed to be reached in a state-contingent fashion based on the information available on environmental, technological and market conditions at time t as illustrated in Figure 4. The information available at time t is contained in the information set $\Theta_{t-1}^{(k)}$.

The information set $\Theta_{t-1}^{(k)}$ over the past H periods contains:

$y_m^{(k)}(h),\ m = 1, \dots, M,\ h = t-1, \dots, t-H$: past observations of every output m,

$x_n(h),\ n = 1, \dots, N,\ h = t-1, \dots, t-H$: past observations of every input n,

$f_s, s \in \Omega$: relative frequency of nature-state occurrence,

$s(h),\ h = t-1, \dots, t-H$: knowledge of the state of nature, which has occurred in period h,

$p_{ms}(h),\ m = 1, \dots, M,\ s \in \Omega,\ h = t-1, \dots, t-H$: past state-dependent output prices,

$w_n(t),\ n = 1, \dots, N$: input prices for every input n in t.

The output and input quantities of every producer k are assumed to be common knowledge for all K producers in a region. The input and output prices are only known to the producer k himself. The expectations regarding the occurrence probability of the nature-states are presumably constructed by producer k based on the relative frequency of state occurrence f_s, which k has determined based on his past experience. The output price expectations of k are based on k's past observations of the state-dependent output price levels and are formed in the adaptive fashion described in section 3.3.

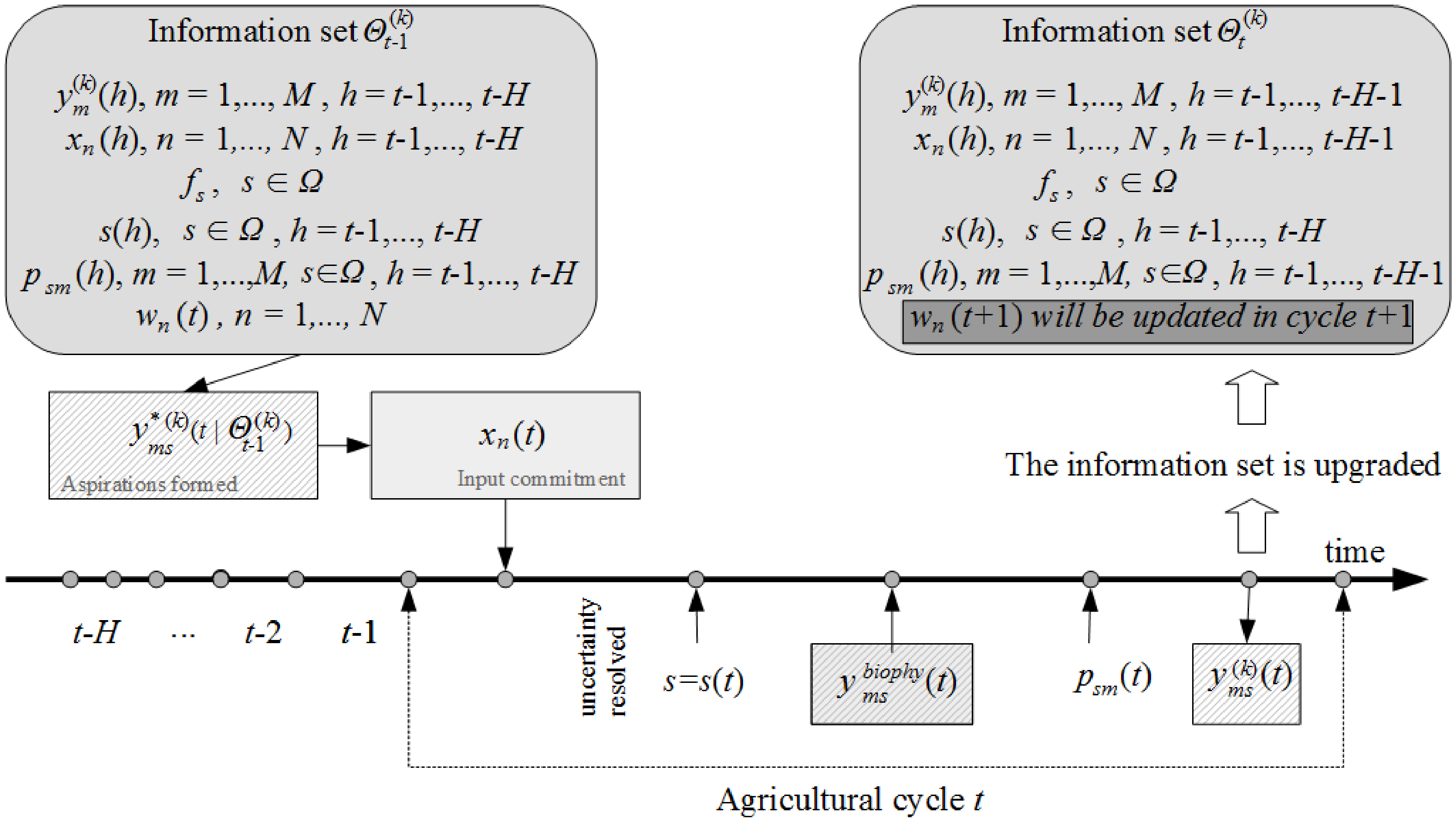


Figure 4 Assumed timeline in the model for agricultural cycle t.
Source: own illustration.

Based on the information set $\Theta_{t-1}^{(k)}$ producer k mentally constructs a set of possible production alternatives in a state-contingent fashion. This set would resemble the set of possible production alternatives illustrated in Figure 3. The exact location of the boundary of the set would depend on the information available to k with respect to past input and output amounts, the input and output price information and the monetary goals in the form of state-contingent revenues producer k has in mind.

The production decision would depend not only on the technologically feasible choices in terms of the set of production alternatives, but also on the environmental information available in the form of relative frequencies of the occurrence of the nature-states $f_s, s \in \Omega$ and on the individual risk preferences of producer k. These risk preferences can be captured by a risk-aversion coefficient.

The production aspirations $y_{ms}^{*(k)}\left(t|\Theta_{t-1}^{(k)}\right)$ are thus determined. The input amounts $x_n(t),\ n = 1, \dots, N$, committed to production are assumed equivalent to the cost minimizing amounts necessary to achieve the determined production aspirations $y_{ms}^{*(k)}\left(t|\Theta_{t-1}^{(k)}\right)$ before uncertainty is resolved, as illustrated in Figure 4.

Uncertainty is resolved for agricultural cycle t after Nature chooses a state of nature $s = s(t),\ s \in \Omega$. The biophysical crop yields y_{ms}^{biophy} thus occur. The yields are presumably free of anthropogenic intervention and only reflect geographical and climatic influences as well as the state Nature has picked. Uncertainty concerning the prices of the M outputs in cycle t, $p_{ms}(t), m = 1, \dots, M$, is resolved as well. The enterprise-level crop yields $y_{ms}^{(k)}(t),\ m = 1, \dots, M$, are realized and the information set of k is upgraded to $\Theta_t^{(k)}$.

6 Estimation of the parameters of a state-contingent production technology for the crop producing sector in Saxony-Anhalt

This chapter addresses an empirical challenge to the implementation of MELCY, the integrative enterprise-level crop yield model under climate change proposed in chapter 5. The model assumed that producer k mentally constructs a set of possible production alternatives based on the information set $\Theta_{t-1}^{(k)}$, which is available to him at the decision-making time. As expressed by (14) and (15) in section 3.1, the boundary of the possible production alternatives set depends on the production technological constraints among other things.

Approximating the boundaries of k's set of alternatives, as these boundaries would appear given the information set $\Theta_{t-1}^{(k)}$, thus requires a quantification of the limitations posed by the state-contingent production technology. This quantification is implemented here by an econometric estimation of the parameters of a state-contingent production technology. This work, similarly to the contribution of Nauges, O'Donnell and Quiggin (2011) described in subsection 4.1.2, estimates a primal technological representation in the form of a state-contingent production function. The choice is made after evaluating the restrictiveness of the distributional assumptions, which the contribution of Chavas (2008), described in 4.1.1, had to make in order to estimate a dual representation of a state-contingent production technology.

6.1 Motivation of the estimation from a production analysis perspective

Shankar comments in his literature overview on conventional production analysis failing to account for the possibly non-output-cubical nature of a production technology (Shankar, 2012, p. 23). The analysis in this chapter contributes towards closing the potential research gap created by the underuse of the state-contingent approach in empirical implementations.

Here a primal state-contingent technological representation, a state-contingent production function, is estimated for the grain producing sector in the German Federal State of Saxony-Anhalt. The functional form chosen closely resembles the functional form in Nauges, O'Donnell and Quiggin. This study also adopts the nature-state definition used in the contribution of Nauges, O'Donnell and Quiggin, the detection of which, as it has been mentioned in subsection 4.1.2 and section 4.3, is necessary in order to obtain coefficient estimates for a primal state-contingent technological representation.

In this empirical inquiry the hypothesis of an output cubical production technology, i.e. a technology, which does not allow the producers to substitute potential future outcomes against one another through input allocation, is tested for the case of crop production in Saxony-Anhalt. The analysis rejects the hypothesis based on simulated data and delivers estimates of the production coefficients based on the simulated data, which suggests that the conventional stochastic frontier analysis would be insufficient as a tool of analysis. The results of the analysis and the conclusions drawn from them, however, are valid to the extent to which the simulated data match the unobserved real data.

The empirical analysis in this chapter primarily delivers a methodological contribution by suggesting an alternative way to detect the occurrence of the nature-states, i.e. to infer the prevalent environmental conditions. Nauges, O'Donnell and Quiggin use expert opinion on the weather conditions, which would benefit the production of a specific crop, and match weather data to accounting records before estimating the production function. Based on previous work by the author this thesis suggests detecting the occurrence of the nature-states based on a new indicator, an index evaluating the yearly experimental crop yields observed at geographically and climatically relevant experimental stations (Angelova, Glauben, and Grings, 2014), (Angelova, 2015). The values of the index are statistically grouped and the results of the statistical evaluation matched to the accounting data before estimating the production function.

The alternative solution proposed in this chapter provides certain benefits, which make it relevant to empirical production analysis. On the positive side, the fact that the index is statistically evaluated contributes towards the objectivity of the evaluation results, since they are based on optimizing a mathematical measure of similarity or distance between the data points, which lacks the human subjectivity an expert opinion can exhibit. On the negative side, the statistical grouping might lead to mathematically optimal, yet agronomically meaningless results due to the mechanistic nature of the approach. Expert opinion and statistical grouping as ways to detect the occurrence of the states of nature should be therefore seen as complementary to each other.

A clear advantage of the statistical approach to detecting the nature-state occurrence proposed in this chapter is constituted by the index being constructed directly based on obtained crop yields. This is beneficial because crop yields result from a complex local interplay of weather conditions, soil type and systematic, region-specific nutrient applications (Schlenker and Roberts, 2008) rather than from weather only.

The analysis is described as follows: section 6.2 presents the quantitative methods for data clustering employed. Sections 6.3 and 6.4 are devoted to data description and data simulation respectively. Section 6.5 introduces the proposed index for nature-state detection, evaluates the data and assignes the yearly accounting records to one of the states. Section 6.6 deals with the functional form and the empirical specification of the production function. Section 6.7 presents and discusses the results.

6.2 Methods for data clustering

This section presents the quantitative methods employed in the empirical analysis, more specifically the cluster analysis methods used, the hierarchical clustering algorithm described in subsection 6.2.1 and the partitioning algorithm described in subsection 6.2.2.

Cluster analysis in general aims at grouping observations together into groups (clusters), which within themselves should be as homogeneous (dense) as possible, while simultaneously being as heterogeneous between themselves as possible (Härdle and Simar, 2003, p. 271). As Härdle and Simar also remark, the fundamental structure of any cluster analysis involves two steps: choosing a similarity or dissimilarity measure between the observations in order to decide how alike or unlike two observations are and choosing an algorithm to construct the clusters (Härdle and Simar, 2003, p. 271).

The main difference between hierarchical and partitioning algorithms, as Härdle and Simar further remark, consists in whether a specific observation can be reassigned to a different cluster during the application of the algorithm (Härdle and Simar, 2003, p. 277). While this is possible with partitioning algorithms, such as the one described in subsection 6.2.2, it is not possible with the hierarchical ones, such as the one described in subsection 6.2.1.

Both clustering approaches are presented here because their distinct underlying philosophy might provide an applied production researcher with flexibility with respect to determining the number of the nature-states and detecting their occurrences. Hierarchical clustering algorithms give hints with respect to potential meaningful groupings within the dataset and advice on the group membership of specific observations, without requiring the number of groups to be fixed a priori. Partitioning algorithms advise on the group membership of specific observations, where the number of groups is set a priori. If a researcher has a strong theoretical justification with respect to the number of nature-states, then the application of a robust partitioning algorithm would suffice. In this case, however, alternative interpretations of the dataset with respect to the number of nature-states would be lost. Should a researcher wish to empirically evaluate the dataset and contemplate on the number of states the data supports, then the application of a hierarchical algorithm is recommendable. However, evaluating dendrograms the binary tree structures, which visually display the results of a hierarchical clustering, introduces subjectivity to the analysis.

6.2.1 Agglomerative hierarchical clustering

Agglomerative hierarchical clustering is a method, which allows the investigation of the structure of a dataset consisting of N observations. The analysis aims at successively merging all observations into a single cluster starting from the finest possible partition, where each observation constitutes an own cluster.

The dissimilarity (or distance) matrix between all N observations is therefore computed.[3] As Härdle and Simar describe it, the algorithm subsequently proceeds in a two-step iterative fashion (Härdle and Simar, 2003, p. 277):

- The clusters, which are most similar as indicated by the distance matrix, are found and merged,
- The reduced distance matrix between clusters is recomputed based on the selected linkage criterion, which constitutes the way the distance between clusters is defined.

[3]Using a similarity matrix is also possible in principle. Since dissimilarity measures can be calculated from and converted to similarity measures without any loss of information, analyzing the dissimilarity matrix suffices.

The steps are repeated until all observations are merged into a single cluster. The analysis is concluded with the appropriate number of clusters being determined.

The choice of linkage criterion critically influences the clustering results. Two commonly used choices for the linkage criterion to determine the distance between two clusters which contain more than a single observation are the nearest neighbor criterion (the so-called single linkage criterion) and the farthest neighbor criterion (the so-called complete linkage criterion). The single linkage criterion relies on defining the distance between two clusters as the distance between the nearest two observations, which belong to the different clusters (Bartholomew et al., 2002, p. 19). The complete linkage criterion equates the distance between clusters to the distance between the farthest two observations in each cluster, which belong to the different clusters (Bartholomew et al., 2002, p. 20).

Both linkage criteria can be in their own way sensitive to outliers in the dataset since they both define the distance between clusters as the distance between two single observations, which belong to different clusters (Witten, Frank, and Hall, p. 275). As Witten, Frank, and Hall further remark, the two linkage criteria tend to produce clusters with different properties. They exemplify this remark by defining the diameter of a cluster as the largest distance between observations belonging to the cluster, in which case:

- The single linkage criterion tends to produce clusters with very large diameters.
- The complete linkage criterion tends to produce compact clusters with a small diameter, at the possible expense of an observation belonging to a certain cluster being more close to observations in a different cluster than it is to any other observation in its own cluster.

The choice of a linkage criterion critically influences the agglomeration results, i.e. potential meaningful groupings within the dataset and the group membership of specific observations. A similar agglomeration result across a variety of the linkage criteria should ultimately offer reassurance that some truth regarding the underlying structure in the dataset is indeed reflected, as Bartholomew et al. observe (Bartholomew et al., 2002, p. 21).

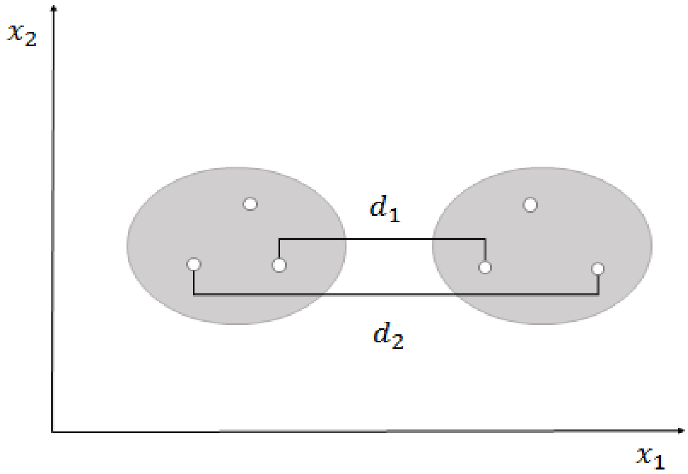


Figure 5 The different definitions of distance between clusters.
Note: Distance between clusters according to the linkage criterion: d_1 would be the distance according to the single linkage criterion; d_2 would be the distance according to the complete linkage criterion.
Source: own illustration.

The agglomeration results can be inspected with the help of a so-called dendrogram, which visually resembles a tree-like structure. The so-called "height" of the tree-like structure, as it is referred to in the illustrations generated by the R routine *hclust* (R Core Team and contributors 2016a) reflects the distance values, at which a merging of clusters is undertaken (Bartholomew et al., 2002, p. 23). The branches of the tree-like structure are cut at a level, where there appears to be sufficient 'space' to cut them.

6.2.2 Partitioning Around Medoids (PAM)

Partitioning around medoids is another clustering method, which allows the attribution of observations $x^{(n)},\ n = 1, \dots, N$, to a preselected number of clusters. It is described in the documentation regarding the R routine *pam* as a robust version of k-means (Maechler et al., 2015). PAM relies on clustering observations around so-called "medoids" - objects, which are considered representative of the groups with respect to revealing the underlying structure of the dataset. Unlike the clustering results from k-means, the clustering results from PAM are independent of the initial point (Kaufman and Rousseeuw, 2005, p. 104). PAM clustering is however more computationally intensive than k-means clustering.

Kaufman and Rousseeuw describe the algorithm in the following way: if $X^{(N)}$ denotes the set of observations $x^{(n)},\ n = 1, \dots, N,$ and $m^{(k)},\ k = 1, \dots, K,$ are the K medoids, then PAM searches to determine:

$$\left\{ m^{(1)}, \dots, m^{(K)} : \sum_{i=1}^{K} \sum_{j=1}^{N} \left| x^{(j)} - m^{(i)} \right| = min \sum_{i=1}^{K} \sum_{j=1}^{N} \left| x^{(j)} - x^{(i)} \right| \right\} \tag{29}$$

where

$x^{(j)}, x^{(i)}, m^{(i)} \in X^{(N)}$.

The PAM algorithm is implemented in two phases. The first phase, referred to as BUILD, consists in constructing the initial medoids. The second phase, referred to as SWAP, consists in improving the medoids, and, hence, the resulting cluster attribution.

The BUILD phase consists of successively selecting the K medoids $m^{(k)}$ which are considered representative to the K groups. A dissimilarity matrix between all N observations is computed. The first medoid selected is the observation which minimizes the sum of dissimilarities to all other observations. The subsequent medoids are chosen in the following way (Kaufman and Rousseeuw, 2005, p. 102):

- An observation $x^{(i)}$, which has not been chosen to be a medoid, is considered.
- Another observation $x^{(j)}$, which has not been chosen as a medoid, is considered. The dissimilarity D_j to the most similar previously selected medoid as well as the dissimilarity $d(j, i)$ to observation $x^{(i)}$ are calculated.
- The difference between D_j and $d(j, i)$, $D_j - d(j, i)$, is calculated. The following index, the partial gain towards choosing $x^{(i)}$ as a medoid brought about by the consideration of $x^{(j)}$, is calculated:

$$C_{ji} = max\left(D_j - d(j, i), 0\right) \tag{30}$$

  The significance of a partial gain C_{ji} index becomes intuitively clear when keeping in mind that D_j and $d(j, i)$ are measures of dissimilarity. Thus a positive difference $D_j - d(j, i)$ means that $x^{(j)}$ is more similar to $x^{(i)}$ than to the closest medoid. Thus a positive difference would mean that the consideration of $x^{(j)}$ provides an argument for observation $x^{(i)}$ being chosen as a medoid. If $x^{(j)}$ is closer to the closest medoid than to $x^{(i)}$, then the consideration of $x^{(j)}$ does not speak for observation $x^{(i)}$ being chosen as a medoid.

- The total gain obtained by choosing observation $x^{(i)}$ as a medoid is calculated by adding up the partial gains over all possible non-selected $x^{(j)}$:

$$\sum_j C_{ji} \tag{31}$$

- A new medoid is chosen among all observations $x^{(i)}$ as the $x^{(i)}$ which maximizes the total gain of being chosen, when all other non-selected observations $x^{(j)}$ are considered:

$$\max_i \sum_j C_{ji} \tag{32}$$

The BUILD phase ends when K medoids $m^{(k)}$ have been chosen among the observations.

The SWAP phase attempts to improve the medoid selection and, thus, to improve the resulting assignment to clusters, which is achieved by considering the potential gain of a swap between medoids selected in the BUILD phase and non-selected observations. The potential gain of a swap is evaluated from the point of view of the non-selected observations in the dataset by executing the following steps (Kaufman and Rousseeuw, 2005, p. 103):

- An observation $x^{(j)}$, which has not been chosen as a medoid, is considered. Also considered are a medoid $m^{(i)}$ and a non-selected observation $x^{(h)}$. The contribution of observation $x^{(j)}$ towards a swap C_{jih} between $m^{(i)}$ and $x^{(h)}$ is calculated. Four situations can feasibly arise:
  - Observation $x^{(j)}$ is further away from both $m^{(i)}$ and $x^{(h)}$ than to another medoid. The contribution C_{jih} is then set to zero.
  - Observation $x^{(j)}$ is closer to $m^{(i)}$ than to any other medoid and closer to $x^{(h)}$ than to the second closest medoid. The contribution towards a swap is set to

$$C_{jih} = d(j,h) - d(j,i) \tag{33}$$

  - Observation $x^{(j)}$ is closer to $m^{(i)}$ than to any other medoid and at least as far away from $x^{(h)}$ as from the second closest medoid. The contribution towards a swap is set to

$$C_{jih} = E_j - D_j \tag{34}$$

where E_j is the dissimilarity of observation $x^{(j)}$ to the second closest medoid and D_j is the dissimilarity of observation $x^{(j)}$ to medoid $m^{(i)}$.

- Observation $x^{(j)}$ is further away from $m^{(i)}$ than from at least one other medoid, yet closer to $x^{(h)}$ than to any other selected or non-selected observation. The contribution towards a swap is set to

$$C_{jih} = d(j,h) - D_j \tag{35}$$

The potential gain of a swap between $m^{(i)}$ and $x^{(h)}$ is subsequently calculated by adding up the contributions over all non-selected objects $x^{(j)}$ in the dataset:

$$T_{ih} = \sum_j C_{jih} \tag{36}$$

A decision to swap or not is based on selecting the pair ($m^{(i)}$, $x^{(h)}$), which minimized the potential gain of a swap T_{ih}. A negative minimum means a swap is carried out and the SWAP stage is repeated. In any other case the algorithm stops (Kaufman and Rousseeuw, 2005, p. 104). Kaufman and Rousseeuw report that PAM obtains reasonable clustering results in case studies (Kaufman and Rousseeuw, 2005, p. 92). The data used in these case studies is structurally similar to the data used for the empirical analysis presented in this chapter.

6.3 Data

The estimation of a state-contingent production function provides a way to meaningfully integrate “biophysical” and “economic” crop yield data. The biophysical crop yield data result from scientific experiments at agronomic stations, which are meant to primarily reflect the effects of the interplay of factors in the natural environment. This data could make for a suitable proxy for the biophysical crop yields of crop m in period t under state of nature s, $y_{ms}^{biophy}(t)$, in (28) in chapter 5. The application of fertilizers and fungicides during the experiments is kept to the amount typical of a well-managed enterprise in the region. Human intervention during the generation of the biophysical yield data is typically much lower than the amount of intervention during the generation of the economic yield data. This is understandable since the economic crop yield data stem from the accounting records of optimizing enterprises selling the crop harvest, which makes managerial decisions such as the use of machinery, labor and intermediate inputs vital. This economic crop yield data could make for a suitable proxy for the crop yields of crop m of k in period t under state of nature s, $y_{ms}^{(k)}(t)$, in (28) in chapter 5. The observations of both data types used in this thesis span between the years 1996 to 2007. The timeframe is chosen since both biophysical and economic data are available in this period.

6.3.1 Biophysical data

This subsection describes circumstances surrounding the formation of the biophysical crop yields at three experimental stations in Saxony-Anhalt, the location of which is marked on Map 1 in Appendix I. The term "biophysical crop yields" refers to a specific type of phenological observations, the mean yield levels of the crop experiments from multiple fields at the experimental stations in Saxony-Anhalt. The factors contributing to the formation of the yields are mainly climate and weather conditions, geographical location, soil type and other specifics of the fields. Managerial decisions also play a role to an extent, such as amount of seeds used, sewing dates and eventual use of fertilizers as well as plant protection products.

The crops to analyze, *winter wheat* and *winter barley*, have been chosen due to data availability reasons. The mean wheat and mean barley yields at the three stations for the years between 1996 to 2007 are plotted in Appendix II. The data was retrieved from the experimental reports of the Landwirtschaftliche Untersuchungs- und Forschungsanstalt (LUF), which outline the course of the experiments. An example of such an experimental report would be (LUF, 2001). An emphasis should be made on the anthropogenic influences, which contribute to the formation of the crop yields.

Crop yields result from open-field agronomic experiments, where human intervention is controlled and recorded. The goal of the experiments is primarily to discern average plant development under specific soil and weather conditions typical for the region given the planting practices. The experimental protocols describe the locations and the weather conditions in a specific year as well as the development of the plants under these conditions. Specifically, the protocols record when the phenological stages of plant development were reached, the occurrence of illnesses, the plant biomass as well as the crop yields.

The protocols also record the various management factors, which play a role in determining the resulting crop yields, for instance the sowing dates, the distance between the sowing rows as well as the specific plant sort and the amount of seeds used. A formal recommendation regarding the latter, which is issued in the guidelines of the Bundessortenamt, postulates that the amount of seeds used at the experimental stations should be approximately equivalent to the amount typically used in the region and for the soil type (Bundessortenamt, 2000, pp. 21-60).

The major anthropogenic factor, which determines crop yields and is regulated by the Bundessortenamt, is the use of fertilizers and plant protection products. The recommended amounts should be approximately the amounts used in a “well-managed” enterprise in the region (Bundessortenamt, 2000, p. 40). The definition of a “well-managed” enterprise is somewhat open. Fertilizers and plant protection products appear to have been used in moderate amounts according to the experimental protocols.[4] The biophysical yield values used in this work are obtained in an experimental setting with no fungicide applied in order to be as close as possible to the theoretical setting of the model proposed in chapter 5.

The issue of the acceptable level of anthropogenic influences for the biophysical yields used to detect the states of nature and the significance of different levels of anthropogenic influence for the estimation results will be raised again in subsection 6.8, which discusses the results.

6.3.2 Economic data

The analysis uses aggregations of accounting records from three agricultural regions in the Federal State of Saxony-Anhalt, Germany. The farm level accounting data was collected by the statistical division of the Landesanstalt für Landwirtschaft, Forsten und Gartenbau (LLFG), a local agency, which annually informs the general public on relevant statistics in several sectors in the Federal State. The publicly accessible reports show aggregates, i.e. mean values for so-called “agricultural regions” (“Agrarregion”). These agricultural regions are small areas thought of as more or less homogeneous entities by agro-economic criteria such as general landscape, soil types and the predominant way arable land is used in the region (Heyer, 2010, p. 14). The aggregation of accounting records in principle is necessary in order to satisfy the requirements of the German data privacy legislation.

[4] Gratitude is expressed to Thomas Chudy for helping with the evaluation of the experimental protocols.

A full time series of publicly accessible accounting averages are available only for three agricultural regions in Saxony-Anhalt, the location of which is marked on Map 1 in Appendix I. The reason for the incomplete time series in the other regions is the borderline number of enterprises active in the regions: only in some years is the number of active enterprises sufficient to obtain aggregated values, which would satisfy the data privacy legislation requirements and inform the general public. Therefore the three agricultural regions considered are Altmark, Schwarzerde and Heiden.

The aggregated data is on sole proprietorships ("Einzelunternehmen") specialized in market crop production. The observations span between the years 1996 to 2007 and are retrieved from print versions of the public records (MRLU, 1997), (MELF, 1998), (MELF, 1999), (MRLU, 2000), (MRLU, 2001), (MLU, 2002), (MRLU, 2003), (MLU, 2004), (MLU, 2005), (MLU, 2006), (MLU, 2007), (MLU, 2008). The number of active enterprises in the agricultural regions Altmark, Schwarzerde and Heiden as well as the mean number of enterprises in each region and for each year are indicated in Table 2.

Table 2 Number of specialized sole proprietorships.

	Years												
	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	Mean
Altmark	24	35	35	29	28	30	29	24	29	34	34	34	30.42
Schwarzerde	79	92	80	87	98	101	103	95	150	157	154	122	109.83
Heiden	15	18	17	14	19	18	13	12	11	13	13	11	14.33
Mean	39.3	48.3	44	43.3	48.3	49.7	48.3	43.7	63.3	68	66.3	55.7	51.53

Note: Number of sole proprietorships specialized in crop production in each one of the three agricultural regions of Saxony-Anhalt.
Source: (MRLU, 1997), (MELF, 1998), (MELF, 1999), (MRLU, 2000), (MRLU, 2001), (MLU, 2002), (MRLU, 2003), (MLU, 2004), (MLU, 2005), (MLU, 2006), (MLU, 2007), (MLU, 2008).

The output and input variables for the regression analysis are described in Table 3. The public records provide 36 observations per variable in Table 3, 12 yearly observations for each of the three agricultural regions. Some are monetary quantities, which have to be converted to real values using an appropriate price index in order to estimate the parameters of the production technology:

- Output aggregate has been deflated with the index of prices received by farmers (Index der Erzeugerpreise landwirtschaftlicher Produkte nach Erzeugnissen),

- Depreciations have been deflated with the corresponding subindex of prices paid by farmers (Index der Einkaufspreise landwirtschaftlicher Betriebsmittel),
- Intermediate inputs have been deflated with the corresponding subindex of prices paid by farmers (Index der Einkaufspreise landwirtschaftlicher Betriebsmittel).

The common base of the price indices is set to the year 2005 (BVEL, 2001), (BVEL, 2004), (BELV, 2009), (BELV, 2012).

Table 3 Regression variables and their short description.

Variable	Description
y_t	Volume of aggregate agricultural output in t: deflated aggregate output per enterprise
x_1	Land devoted to wheat production in ha per enterprise
x_2	Land devoted to barley production in ha per enterprise
x	All arable land in ha per enterprise
z_1	Labor in standardized labor units per enterprise
z_2	Capital: deflated depreciations and maintenances for machinery and buildings per enterprise
z_3	Volume of intermediate inputs: deflated sum of intermediate inputs for seeds, fertilizers, pesticides, electricity, etc. per enterprise
t	Year of the observation in a four-digit format divided by the number of observations

Source: own illustration.

6.4 Simulation

The publicly available accounting aggregates ("LLFG mean") provide 36 observations for the seven variables in Table 4, Table 5, Table 6, 12 observations for each region. The number of observations is too small to estimate the coefficients mentioned later in (42) and (43) according to the rule of thumb given by Green, which recommends a number of observations of at least 104 plus the number of regressors (Green, 1991). Making data abundant through simulation, and hence making a distributional assumption on how observations on individual enterprises are distributed around the publically available averages, might provide a solution in case of an insufficient database. This would make the results of the subsequent analysis reliable insofar as the data generating process is justifiable. A multivariate normal distribution has been chosen for the task at hand due to its extensive use in the field of statistics and the appealing mathematical properties (Härdle and Simar, 2003, p. 147).

The means of the seven-dimensional multivariate normal distributions, one for each year in each agricultural region, are assumed equivalent to the publicly available yearly accounting aggregates of the output and input variables described in Table 3. These are indicated by $\hat{\boldsymbol{\mu}}$ in Figure 6. An estimate for the variances $\hat{\sigma}$ in the estimated matrix $\hat{\boldsymbol{\Sigma}}$ is obtained through the temporal variation across each dimension as Figure 6 indicates. The off-diagonal elements in the estimated matrix $\hat{\boldsymbol{\Sigma}}$ are assumed equal to zero for computational simplicity with the possibility of estimating the covariances from the sample discussed in section 6.7. It is important to note that while simulating data could make observations abundant, it introduces assumptions about the underlying statistical population, in this case on how observations for specific enterprises are distributed around the publicly available yearly accounting aggregates. Results based on the data are thus valid insofar as these assumptions are justified.

Thus 36 multivariate normal distributions are obtained, one multivariate normal distribution for each region and each year. A number of data points simulated for each region per year approximately corresponds to the average number of enterprises active in the respective region, as suggested by Table 2:

- 30 data points per year in Altmark,
- 110 data points per year in Schwarzerde,
- 14 data points per year in Heiden.

The task is achieved with the R routine *mvrnorm* in the package MASS (Venables and Ripley, 2002). A symmetric truncation has been undertaken around the means after the simulation, which results in the values spanning between 0.6*mean and 1.6*mean, with 1848 observations in total. The truncation is necessary, given the nature of the data, to avoid both negative values, which should be impossible, and unrealistically high values. The truncation ensures the absence of outliers and allows for the use of non-robust regression techniques such as ordinary least squares.

The observed yearly mean of each variable and the corresponding empirical mean of the simulated observations for the variable are contrasted for Altmark, Schwarzerde and Heiden in Table 4, Table 5 and Table 6 respectively.

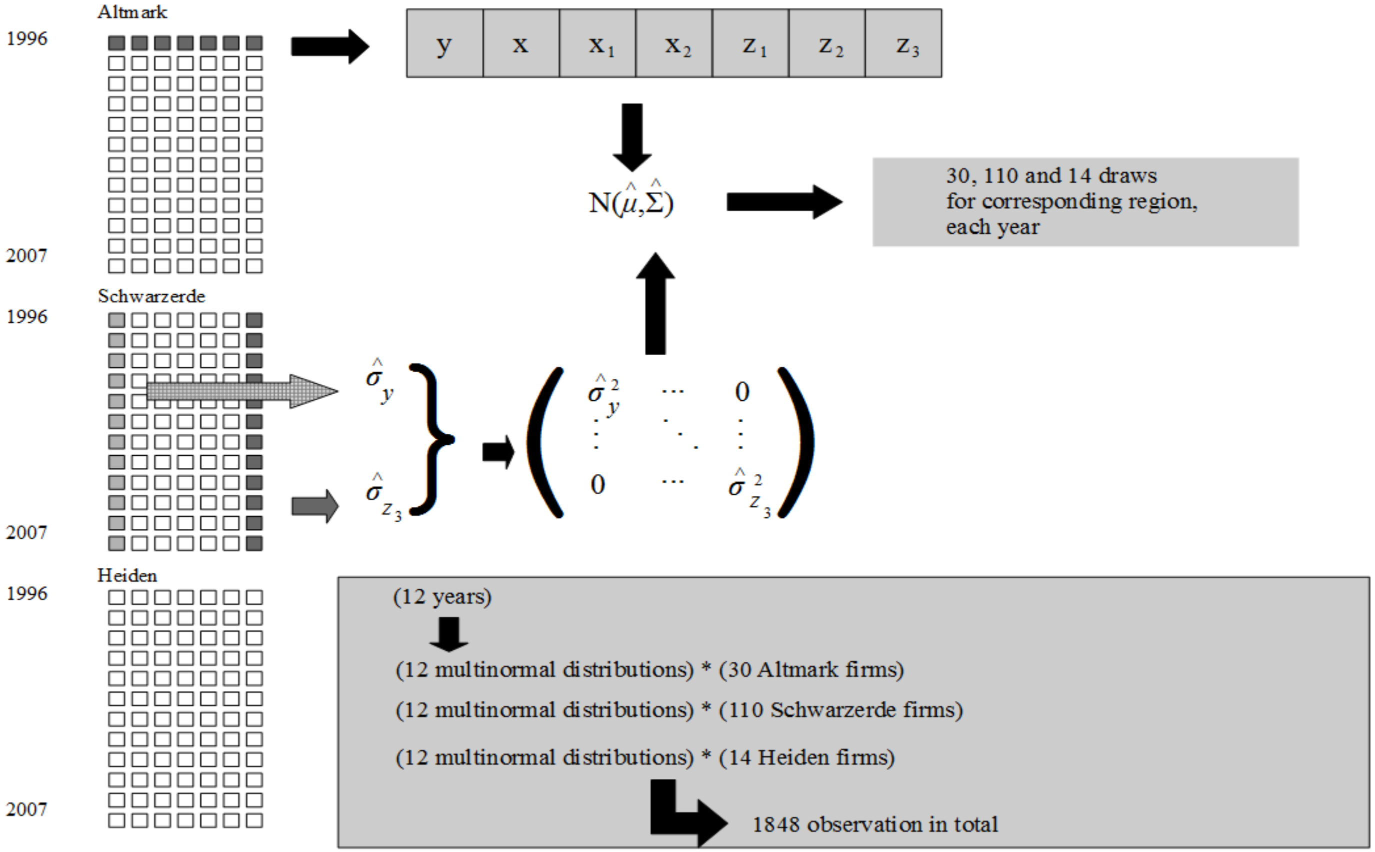


Figure 6 Distributional assumptions behind the data simulation process.
Source: own illustration.

Table 4 Comparison between observed means and simulated means for Altmark.

Year t	Type	y	x	x_1	x_2	z_1	z_2	z_3
1996	LLFG mean	147281,4751	187,82	47	26	2	49648	58361,5491
	Sim. Mean	147456,0535	190,2779	45,2249	25,5918	2,008	49508,0139	58731,4749
1997	LLFG mean	98469,8867	177	44	24	2,1	46069	71992,3382
	Sim. Mean	99751,1646	177,9404	43,828	23,431	2,1371	46352,2455	70833,4411
1998	LLFG mean	107139,5271	169,7	42	28	1,8	44045	74021,5314
	Sim. Mean	106151,4937	167,8516	42,2413	28,927	1,795	44450,2704	73369,2612
1999	LLFG mean	108378,9849	182	42	27	1,9	47608	83353,2342
	Sim. Mean	109464,0282	182,2743	40,7658	25,5338	1,9264	48151,7711	84875,8844
2000	LLFG mean	115660,6128	172,3	38	22	1,6	45744	76251,8484
	Sim. Mean	115030,3903	173,4294	38,4562	22,6894	1,5543	45093,0078	76511,9606
2001	LLFG mean	91572,1572	179,5	46	22	1,7	48127	72602,1151
	Sim. Mean	92763,764	179,8867	48,253	21,9783	1,6921	49616,6635	73346,5627
2002	LLFG mean	111123,0884	185,2	52	22	1,8	46588	76943,9996
	Sim. Mean	109634,2658	184,502	49,7368	21,5045	1,825	46951,8891	79956,1428
2003	LLFG mean	83362,8327	186	50	22	2	43865	73867,348
	Sim. Mean	79337,0961	181,8321	50,9408	23,5988	2,0293	43711,8816	71109,1772
2004	LLFG mean	105199,3069	238,3667	66	25	2,1	53562	95099,3747
	Sim. Mean	101359,5112	239,284	64,9067	24,8363	2,1067	54237,4223	95579,5076
2005	LLFG mean	116734,0825	225,5	62	29	2,02	40151	82780,0276
	Sim. Mean	116631,459	229,0527	64,8842	28,8626	1,9969	39958,5739	82550,6741
2006	LLFG mean	109371,4004	205	66	27	1,93	40848	76690,11
	Sim. Mean	104910,0768	207,0066	63,3711	27,2352	1,9106	41267,4999	76572,9348
2007	LLFG mean	108465,1571	219,3333	65	45	2,15	45969	79732,2716
	Sim. Mean	106037,8904	223,5602	62,5033	45,7614	2,1809	45544,7769	83037,4159

Note: Yearly comparison of the observed means ("LLFG mean") and the empirical mean of the simulated observations ("Sim. mean") for each variable.
Source: own illustration.

Table 5 Comparison between observed means and simulated means for Schwarzerde.

Year t	Type	y	x	x_1	x_2	z_1	z_2	z_3
1996	LLFG mean	293555,4446	269,05	126	42	2,64	72153,7858	96191,7665
	Sim. Mean	291973,531	269,3046	125,9505	41,8502	2,6456	71076,752	97347,2892
1997	LLFG mean	245009,5114	274	127	42	2,7	81816,5916	128515,8046
	Sim. Mean	249471,2481	273,6615	126,7419	42,1856	2,7138	79612,5311	128345,9868
1998	LLFG mean	253979,6432	269,2	123	45	2,7	81721,8933	131907,1701
	Sim. Mean	248949,2967	269,0722	123,2383	45,1321	2,7052	81811,1653	131805,6254
1999	LLFG mean	237856,2242	272	123	46	2,8	77319,2483	134656,2557
	Sim. Mean	237969,3845	272,9632	123,0584	45,692	2,7984	77108,9416	135527,0136
2000	LLFG mean	249916,5033	263	119	38	2,6	78413,8929	127104,4798
	Sim. Mean	253122,5037	262,3878	119,2612	37,5385	2,6174	78071,1858	127985,8324
2001	LLFG mean	226965,9976	261,8	128	41	2,7	73498,1996	130785,6217
	Sim. Mean	224643,4297	261,1624	128,2089	40,8828	2,7156	73216,6468	130466,8913
2002	LLFG mean	165893,7859	262,5	123	35	2,5	61342,2719	107411,9718
	Sim. Mean	163979,8036	261,0252	123,3361	35,3447	2,4896	61370,568	108546,5443
2003	LLFG mean	186322,5774	272,8	121	35	2,5	64781,213	124010,4241
	Sim. Mean	182994,104	272,1662	121,1912	33,9123	2,5042	64383,7332	123872,5923
2004	LLFG mean	215508,7812	267,1667	124	34	2,3667	66070,337	128235,3633
	Sim. Mean	216932,3179	266,6204	124,6722	34,9161	2,3661	65334,071	128409,992
2005	LLFG mean	223513,4796	264,9	128	30	2,27	62728,1832	126645,2296
	Sim. Mean	229005,9814	265,7649	127,6555	29,7609	2,2444	61719,077	125503,4322
2006	LLFG mean	208111,3632	278,3333	132	31	2,19	63330,6483	124556,6227
	Sim. Mean	208791,5713	277,9978	132,7407	31,9617	2,2392	63565,3542	124530,3886
2007	LLFG mean	229725,4112	283	133	39	2,4267	65711,8608	128215,4667
	Sim. Mean	230213,9509	283,2101	133,1248	38,3498	2,421	65071,8299	128024,7105

Note: Yearly comparison of the observed means ("LLFG mean") and the empirical mean of the simulated observations ("Sim. mean") for each variable.
Source: own illustration.

Table 6 Comparison between observed means and simulated means for Heiden.

Year t	Type	y	x	x_1	x_2	z_1	z_2	z_3
1996	LLFG mean	74982,1925	156,39	13,96	16,31	1,83	33889,9186	46194,0872
	Sim. Mean	83904,185	154,9019	14,6752	16,4039	1,8617	31770,4711	45018,1572
1997	LLFG mean	55178,7283	175	15,14	17,68	1,9	34116,4395	59775,2092
	Sim. Mean	55138,0522	182,7685	15,7268	19,0653	1,8763	33834,2393	57825,086
1998	LLFG mean	66000	195	10,99	21,1	2,11	37126,8965	64998,0348
	Sim. Mean	69268,9476	189,6819	11,1936	20,9138	2,1556	36202,9846	61993,3385
1999	LLFG mean	72880,8838	213,4	19,43	23,47	2,3	43104,5168	72292,1016
	Sim. Mean	85513,9251	218,2874	21,1593	24,66	2,2961	44409,0829	71610,9318
2000	LLFG mean	69905,9105	199	15,95	19,79	2	41011,897	68421,0112
	Sim. Mean	72684,4562	196,452	17,3414	20,6471	1,9339	40845,3764	73148,9589
2001	LLFG mean	89196,3545	200,7	31,68	22,55	1,9	48905,1468	83976,0513
	Sim. Mean	85497,542	204,5247	32,0418	23,0067	1,837	47298,6653	86838,5919
2002	LLFG mean	76535,8385	188,9	10,52	16,04	1,7	38793,738	69708,3928
	Sim. Mean	75858,3986	183,9105	10,7278	16,7893	1,6939	39123,4499	73148,9166
2003	LLFG mean	75268,2031	203,1	21,21	29,29	1,8	40860,8888	73556,3828
	Sim. Mean	83389,7088	213,7773	26,3516	31,0264	1,7475	40141,6958	72154,0645
2004	LLFG mean	82666,1579	222,6	29,02	21,66	1,9	39471,9672	75419,3907
	Sim. Mean	79917,2781	231,4189	26,4691	21,9664	1,9318	38965,3889	80013,4442
2005	LLFG mean	91666,2679	240,7	23,56	24,99	2,22	41845,0561	79863,7365
	Sim. Mean	84406,7342	240,5508	25,6385	25,9149	2,2235	39067,9706	77390,0289
2006	LLFG mean	114574,1844	252	43,88	22,99	2,01	47313,986	91096,0382
	Sim. Mean	114558,7359	257,8205	43,0375	21,7404	2,0306	45037,9304	94296,2838
2007	LLFG mean	132042,1211	271	37,31	34,92	2,08	51040,8433	102493,5806
	Sim. Mean	135194,2712	273,2147	39,8666	36,4608	2,1148	52225,6941	100783,8995

Note: Yearly comparison of the observed means ("LLFG mean") and the empirical mean of the simulated observations ("Sim. mean") for each variable.
Source: own illustration.

6.5 States of nature: definition and detection

The definition of nature-states used in the empirical analysis matches the definition of Nauges, O'Donnell and Quiggin, the contribution described in subsection 4.1.2. By that definition a nature-state is a combination of environmental conditions, which favor the production of a certain crop. This thesis proposes an alternative way to detect the occurrence of the nature-states and to infer the environmental conditions, the benefits of which have already been addressed in section 6.1.

The method for detection proposed by Nauges, O'Donnell and Quiggin consists of asking expert opinion on the weather conditions, which constitute a state of nature, and build a multidimensional index based on weather data. The dimensions of the index are summarized in Table 1 in subsection 4.1.2.

The method of detection proposed here relies on constructing a ratio of crop experimental results and statistically grouping them, as it has been suggested by Angelova (Angelova, 2015). In other words, transition is made from using weather data as an indicator for the prevalent environmental conditions to using phenological data as an indicator for the prevalent environmental conditions. An index, which sets the phenological observations in relations to one another, is constructed and statistical methods to evaluate the index proposed.

Following Angelova (Angelova, 2015), the three agricultural regions of Saxony-Anhalt are matched with three experimental stations, where field experiments on a variety of crops are performed. The location of both agricultural regions and experimental stations is indicated on Map I in Appendix I.

The three stations are:

- Beetzendorf (Altmark),
- Magdeburg (Schwarzerde),
- Gadegast (Heiden).

The four barley observations for Gadegast corresponding to the years 1997, 2001, 2003, and 2005 are missing. These observations have been replaced with the average barley yield observed at Gadegast, which was computed from the available observations. The means of the winter wheat and winter barley experiments (see Appendix II) have been set in relation to one another for every experimental station, constructing the following indicator for each year t, that will be hereafter referred to as biophysical yield ratio or simply yield ratio:

$$\frac{mean(barley)_t}{mean(wheat)_t} \tag{37}$$

As in Angelova (Angelova, 2015) the phenological data used in this work are the mean values of crop experiments at experimental stations and the index is single-dimensional. Environmental conditions more suitable for the cultivation of one crop over another crop would presumably be reflected in a relatively high mean yield for the first crop compared to the mean yield for the second crop, thus determine the magnitude of the index. Environmental conditions more suitable for the cultivation of barley than for the cultivation of wheat would thus be reflected by higher values of the index in (37). Environmental conditions more suitable for the cultivation of wheat than for the cultivation of barley would be reflected by lower values of the index in (37).

Angelova, Glauben and Grings proposed a similar index, which uses phenological observations to infer the prevalent environmental conditions (Angelova, Glauben and Grings, 2014). The proposed index is multidimensional and evaluates the mean values of crop experiments at experimental stations, the minimum values of crop experiments at experimental stations and the maximum values of crop experiments at experimental stations. The multidimensional index proposed requires more sophisticated statistical methods to evaluate - the possibility of the dimensions of the index “contradicting” each other with respect to the result of the evaluation is present. This multidimensional index also requires much more detailed data, which has been unavailable in the present analysis. The single-dimensional index in Angelova (2015) is thus employed here.

The index in (37) is calculated for the three experimental stations over the twelve years. The dendrograms resulting from a hierarchical clustering of the biophysical yield ratios by single and complete linkage for each experimental station are displayed in Appendix IV. The dendrograms for the experimental station Beetzendorf indicate two clusters regardless of the linkage criterion. The branches of the dendrograms are cut where there appears to be sufficient space to do so. In the case of the single linkage dendrogram for the station Beetzendorf a suitable place to cut would be at a height between 0.07 and 0.27, for instance at a height of 0.15. A suitable place to cut the complete linkage dendrogram would be at a height of 0.35. The biophysical crop ratio for the year 2003 seems to belong to one cluster and the remaining observations to another cluster.

The dendrograms for the experimental station Gadegast also indicate two clusters regardless of the linkage criterion. Cutting the branches of the dendrograms at a height of 0.085 and 0.35 for single and complete linkage respectively indicates that the biophysical crop ratios for the years 2000, 2002 and 2007 belong to one cluster and the remaining observations to another cluster.

The dendrograms for the experimental station Magdeburg are less unambiguous. Cutting the branches of the single linkage dendrogram at a height of 0.055 results in three clusters and in identifying the biophysical crop ratio for the year 1996 as an outlier. There is ambiguity regarding the point at which the branches of the complete linkage dendrogram are to be cut. The height 0.25 would result in two clusters, while cutting the branches at a height of 0.1 would result in four clusters.

A partitioning algorithm is used to validate the attribution of the biophysical yield ratios to two clusters as recommended by the hierarchical clustering. The robust PAM is used to group the biophysical yield ratios at all stations into two groups. The attribution of the yearly yields ratios to one of the two states of nature is visualized in Figure 7 using the symbols shown in Table 7. The symbols represent the affiliation of a specific year to one of the states of nature, the barley-favorable state or the wheat favorable state.

Table 7 Symbols representing the states of nature.

	States of nature	
State	Barley-favorable	Wheat-favorable
Symbols	☆	○
Manifestation	High barley yield, low wheat yield	Low barley yield, high wheat yield

Note: The star represents the barley-favorable state, the circle represents the wheat-favorable state.
Source: own illustration.

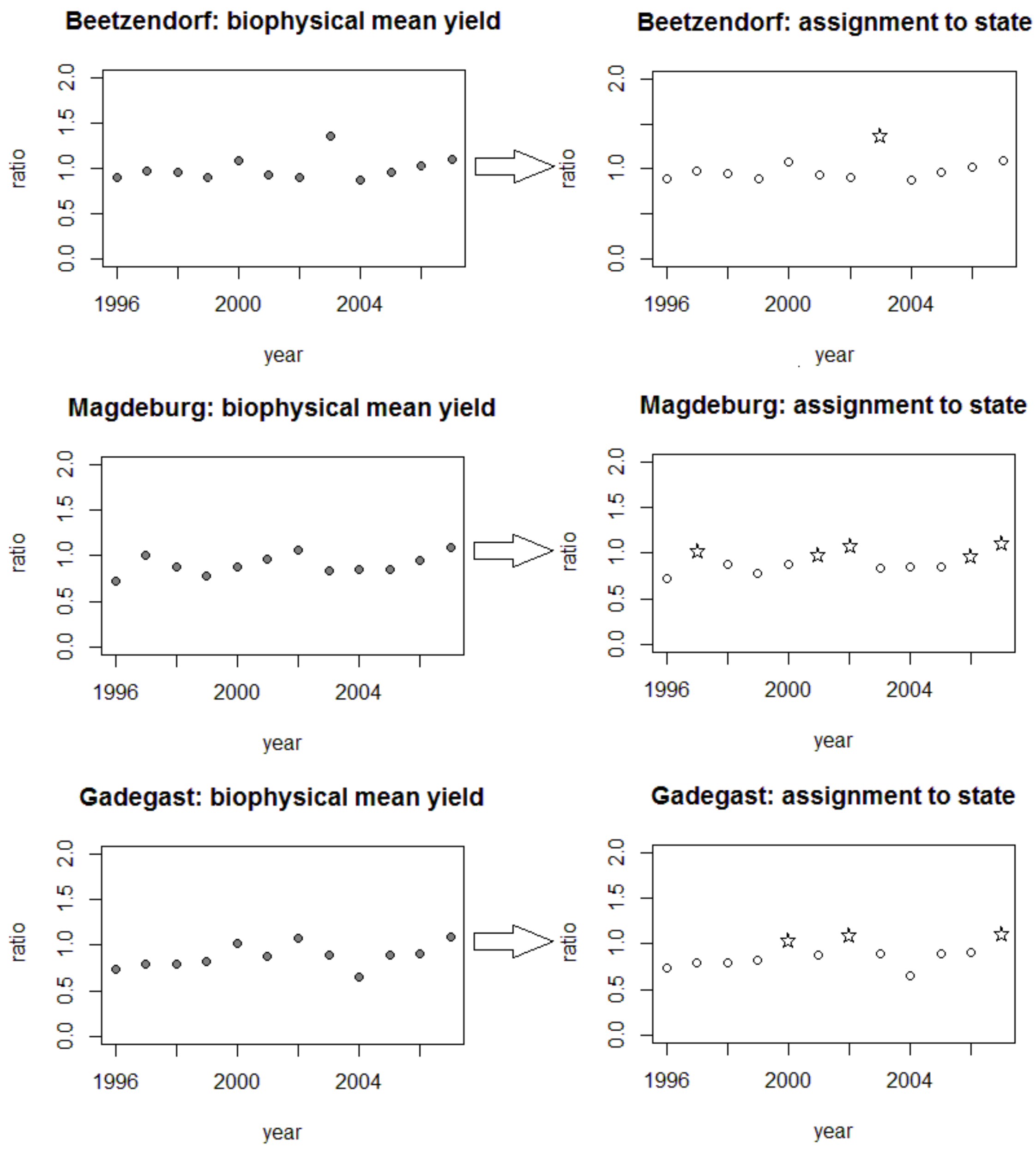


Figure 7 Clustering results.
Note: The yearly biophysical mean yield ratios (left hand side) are attributed to one of the two states of nature (right hand side) using the symbols introduced in Table 7.
Source: own illustration.

As can be inferred from Figure 7 PAM validates the attribution of the biophysical yield ratios recommended by the agglomerative hierarchical clustering. In the case of the experimental station Magdeburg PAM validates the results of the agglomerative hierarchical clustering using the complete linkage criterion. Disregarding the ambiguity surrounding the observations from the experimental station Magdeburg, two states of nature are decided upon:

- Barley-favorable state of nature, which is presumably consistent with rather high barley yields and rather low wheat yields compared to the average values in the agricultural region over the observed years,
- Wheat-favorable state of nature, which is presumably consistent with rather low barley yields and rather high wheat yields compared to the average values in the agricultural region over the observed years.

The attributions are summarized in Table 8. These attributions of a specific year in a specific region to a state of nature are adopted in the further production analysis in order to construct dummies for the nature-state occurrence. The yearly dummies signify the environmental conditions prevalent in the year and account for the environmental factor in the econometric analysis of accounting records.

Table 8 Attribution of the yearly observations of the biophysical yield ratio at the three experimental stations in Saxony-Anhalt to one of the two states of nature.

	Beetzendorf	Magdeburg	Gadegast
1996	Wheat-favorable	Wheat-favorable	Wheat-favorable
1997	Wheat-favorable	Barley-favorable	Wheat-favorable
1998	Wheat-favorable	Wheat-favorable	Wheat-favorable
1999	Wheat-favorable	Wheat-favorable	Wheat-favorable
2000	Wheat-favorable	Wheat-favorable	Barley-favorable
2001	Wheat-favorable	Barley-favorable	Wheat-favorable
2002	Wheat-favorable	Barley-favorable	Barley-favorable
2003	Barley-favorable	Wheat-favorable	Wheat-favorable
2004	Wheat-favorable	Wheat-favorable	Wheat-favorable
2005	Wheat-favorable	Wheat-favorable	Wheat-favorable
2006	Wheat-favorable	Barley-favorable	Wheat-favorable
2007	Wheat-favorable	Barley-favorable	Barley-favorable

Source: own illustration.

6.6 Functional form and empirical specification of the state-contingent production function

The coefficients of a state-contingent production function based on functional forms (18) and (23) in subsection 4.1.2 are to be estimated. The output aggregate for a specific state of nature s and period t as a function of the inputs is given in (38).

$$y_{st} = (1+\rho)^t \left[\theta x^b + \sum_{s=1}^{S} \delta_s {x_s}^b + \sum_{k=1}^{K} \gamma_k z_k^b\right]^{\varphi/b} \tag{38}$$

where

y_{st}: aggregate output in state s in period t,

ρ: rate of Hicks-neutral technical change,

t: temporal indicator,

θ: coefficient associated with the non-state-allocable input x,

δ_s: coefficient associated with the state-allocable input x_s,

γ_k: coefficient associated with the non-state-allocable input z_k,

x_s: a state allocable input,

x: non-state allocable input,

z_k: non-state-allocable input,

b: substitution parameter,

φ: coefficient reflecting the returns to scale.

The formulation in (38), unlike the formulation in (18), accommodates Hicks-neutral technical change. Features of the model of Nauges, O'Donnell and Quiggin are kept, for instance modeling land devoted to the production of a certain crop as the state-allocable input x_s and allowing the non-state allocable input x to play a role in the formation of the state-contingent aggregate output y_{st} even if there is no land dedicated to any of the analyzed crops. Yet, the most notable feature inherited from the model of Nauges, O'Donnell and Quiggin is the incorporation of the parameter φ. This parameter, which constitutes a departure from the typical CES form of a production technology, provides additional modeling flexibility compared to the pure CES form by allowing for increasing or decreasing returns to scale.

This parameter φ scales the production elasticities in (38) for $b \neq 0$, which are given by

$$\eta_s = \frac{\partial ln(y_s)}{\partial ln(x)} = \frac{\varphi\theta x^b}{\theta x^b + \sum_{s=1}^{S}\delta_s {x_s}^b + \sum_{k=1}^{K}\gamma_k z_k^b} \quad (39)$$

$$\xi_{ss} = \frac{\partial ln(y_s)}{\partial ln(x_s)} = \frac{\varphi[\theta x^{b-1}x_s + \delta_s x_s^b]}{\theta x^b + \sum_{s=1}^{S}\delta_s {x_s}^b + \sum_{k=1}^{K}\gamma_k z_k^b} \quad (40)$$

$$\xi_{sk} = \frac{\partial ln(y_s)}{\partial ln(z_k)} = \frac{\varphi\gamma_k z_k^b}{\theta x^b + \sum_{s=1}^{S}\delta_s {x_s}^b + \sum_{k=1}^{K}\gamma_k z_k^b} \quad (41)$$

The empirical specification in (42) corresponding to the functional form in (38) resembles the empirical specification (23) in subsection 4.1.2:

$$ln(y_t) = tln(1+\rho) + \left(\frac{\varphi}{b}\right) ln\left(\theta x^b + \sum_{s=1}^{2}\delta_{1|s}\, d_s {x_1}^b + \sum_{s=1}^{2}\delta_{2|s} d_s {x_2}^b + \sum_{k=1}^{3}\gamma_k z_k^b\right) + e \quad (42)$$

where

y_t: observed deflated output aggregate per enterprise in period t, which is synonymous to y_{st} in (38) once the state s is realized,

t: year the observation was made in divided by the number of observations,

x: entire land per enterprise,

x_1: arable land devoted to wheat production,

x_2: arable land devoted to barley production,

z_1: labor force in standardized labor units per enterprise,

z_2: deflated depreciations for machinery and buildings per enterprise,

z_3: deflated sum of intermediate inputs per enterprise,

$ln(1+\rho)$: rate of Hicks-neutral technical change,

$\delta_{1|1}$: coefficient associated with land devoted to wheat production in the wheat-favorable state of nature,

$\delta_{1|2}$: coefficient associated with land devoted to wheat production in the barley-favorable state of nature,

$\delta_{2|1}$: coefficient associated with land devoted to barley production in the wheat-favorable state of nature,

$\delta_{2|2}$: coefficient associated with land devoted to barley production in the barley-favorable state of nature,

d_1: a state dummy, 1 if the wheat-favorable state occurs and 0 otherwise,

d_2: a state dummy, 1 if the barley-favorable state occurs and 0 otherwise,

e: i.i.d. noise, $e \sim N(0, \sigma^2)$.

The simulated nature of the data means that no adjustment needs to be made in (42) for the eventual technical inefficiency of a specific enterprise, which constitutes a difference to (23) in the description of the study by Nauges, O'Donnell and Quiggin. (42) is estimated by non-linear least squares (Madsen, Bruun, and Tingleff, 2004, p. 5) using the Levenberg-Marquardt algorithm (Gavin, 2011, p. 2). The R package employed is *nlsLM* (R Core Team and contributors 2016b).

After initiating non-linear least squares with the Levenberg-Marquardt algorithm with a couple of combinations of starting values as suggested in Ritz and Streibig (Ritz and Streibig, 2008, p. 23), convergence has been achieved at a point approximately consistent with the theoretical case of a Cobb-Douglas functional form, where φ is around 0.75 and b is around 0.023. Since the algorithm first chooses the values for φ and b, the estimates of the coefficients associated with the inputs are exceptionally high. Non-linear least squares seems to fail to obtain sensible estimates at the initial initialization points due to possible numerical issues (Fox and Weisberg, 2011, p. 11), which merits further investigation. A numerical investigation with respect to possible initialization points could nevertheless result in non-linear least squares providing realistic coefficient estimates.

An empirical specification corresponding to a Cobb-Douglas function (43), which is similar to one of the models estimated by Nauges, O'Donnell and Quiggin (FLEX0), is estimated by ordinary least squares. The notation in (43) is equivalent to the notation in

(42). Again, the empirical specification in (43) resembles the empirical specification (23).

$$ln(y_t) = tln(1+\rho) + \theta\, ln(x) + \sum_{s=1}^{2} \delta_{1|s}\, d_s ln\,(x_1) + \sum_{s=1}^{2} \delta_{2|s}\, d_s ln\,(x_2) + \sum_{k=1}^{3} \gamma_k ln\,(z_k) + e \quad (43)$$

The estimates of the coefficients θ, $\delta_{1|1}$, $\delta_{1|2}$, $\delta_{2|1}$, $\delta_{2|2}$, γ_1, γ_2, γ_3 in the empirical specification (43) are expected to be positive and in the range between zero and one. For estimation purposes it should be noted that $ln(1+\rho)$ is approximately equal to ρ for small values of ρ.

6.7 Results and discussion from a production analysis perspective

The attribution of the years to states of nature influences the regression result through the dummies. A discussion on the indicator used to detect the state-occurrence, hence infer the environmental conditions, and the way it is evaluated is necessary. The type of clustering algorithm, hierarchical or partitioning, has to be set in accordance with the goal of the researcher, as mentioned in the preamble of section 6.2.

As in (Angelova, 2015) the phenological data used in this work are the mean values of crop experiments and the index is single-dimensional. The single-dimensional indicator seems intuitive, since it sets the mean biophysical crop yields in a relation to one another. As mentioned in section 6.5, environmental conditions more suitable for the cultivation of barley than for the cultivation of wheat would presumably be reflected in a relatively high mean yield for barley compared to the mean yield for wheat. Environmental conditions more suitable for the cultivation of wheat than for the cultivation of barley would be reflected by lower values of the index.

Some remarks on the level of anthropogenic influences, which have contributed to the biophysical crop yields used to detect the states of nature in Angelova (2015), are warranted. This study considers biophysical crop yields obtained with the low-level intensity of fertilizer applied without the application of fungicides in order to maintain theoretical consistency with MELCY, the model proposed in chapter 5. Clustering the theoretically consistent biophysical crop yields might provide a suboptimal empirical fit in (42) and (43) as measured by AIC or BIC compared to clustering biophysical crop yields, which emerged after high-level intensity of fertilizer and fungicide use.

Comparing the information criteria associated with the regressions with eventually deviating dummy variable matrix for state occurrence can thus prove to be a way for choosing a suitable model.

Section 6.1 already outlines the benefits of using the proposed approach to attribute yearly observations to states of nature. The resulting attribution should nevertheless be compared to the attribution obtained as Nauges, O'Donnell and Quiggin propose. An immediate comparison cannot be made here due to the difference in studied regions and granularity of the data used: Nauges, O'Donnell and Quiggin use individual farm level data from Finland, while this analysis is run on simulated data based on aggregate averages from agricultural regions in Germany. Both approaches should be compared using the same economic data. If data on observed biophysical crop yields from experimental stations are missing, some simulations can be delived by an appropriately calibrated crop model.

The estimation results for the Cobb-Douglas functional form are displayed in Table 9. In the Cobb-Douglas case the estimated coefficients associated with the inputs are directly interpretable as output elasticities, i.e. the ceteris paribus responsiveness of the output with respect to a change in an input variable. It is extremely important to note that the estimation results, and the conclusions based on them, are valid to the extent to which the simulated data successfully reconstruct key characteristics of the underlying statistical population. The estimated coefficients in the Cobb-Douglas case, which represent estimates of the log-linearized functions' first derivatives with respect to the inputs, are in the expected range of zero to one.

All coefficients, including the coefficients associated with the state-allocable inputs are significant at a 5% level. The value of the adjusted R-squared statistic associated with the regression is 0.9997, which is to be expected due to the simulated nature of the data. The levels of standard error, while not negligible, appear reasonable.

Table 9 Estimated rate of Hicks-neutral technical change and estimated output elasticities.

Coefficient	Estimate	Standard error
$log(1+\rho)$ (rate of Hicks-neutral technical progress)	0.017	0.002631
θ (output elasticity with respect to total land)	0.237	0.049853
γ_1 (output elasticity with respect to labor)	0.248	0.049184
γ_2 (output elasticity with respect to capital)	0.456	0.034858
γ_3 (output elasticity with respect to intermediate inputs)	0.085	0.036284
$\delta_{1\|2}$ (output elasticity with respect to land devoted to wheat production in the barley-favorable state)	0.256	0.023007
$\delta_{1\|1}$(output elasticity with respect to land devoted to wheat production in the wheat-favorable state)	0.313	0.018425
$\delta_{2\|2}$ (output elasticity with respect to land devoted to barley production in the barley-favorable state)	0.134	0.035504
$\delta_{2\|1}$(output elasticity with respect to land devoted to barley production in the wheat-favorable state)	0.080	0.025279

Note: For instance, the coefficient $\delta_{1|2}$ stands for the elasticity of the output with respect to land devoted to wheat production in the state of nature favorable for barley. The significance level is fixed at 5%.

Source: own calculation.

The estimated rate of Hicks-neutral technological change is around 1.7%. The elasticity of output with respect to total land might seem low compared to the results of the FLEX0 model estimated by Nauges, O'Donnell and Quiggin. The same can be said about the elasticity of output with respect to labor. The elasticity of output with respect to capital seems rather high compared to the results of Nauges, O'Donnell and Quiggin. The elasticity of output with respect to intermediate inputs might seem high compared to the results of Nauges, O'Donnell and Quiggin. The standard error coefficient in the output elasticity with respect to intermediate inputs is, however, significant. Overall, the estimate differences could be explained by different proxies used to quantify the amounts of labor, capital and intermediate inputs used in production.

The output elasticities with respect to land devoted to specific crops are higher in the state of nature favorable for the specific crops than in the states of nature favorable for the cultivation of the other crop. The estimation results indicate that the devotion of land to wheat production would marginally raise the aggregate output more than the devotion of land to barley production regardless of the state of nature. Therefore, in this analysis based on simulated data, the hypothesis of an output cubical production technology is rejected: a substitution of potential state-contingent outputs seems technologically feasible for the crop farmers in the Federal State of Saxony-Anhalt based on the statistically reconstructed data.

A desirable modification of the analysis is the estimation of a distance function rather than a production function. While both are primal representations of a production technology, the distance function representation allows for the estimation of the parameters of a multi-output, multi-input technology. Each of the outputs, in this case crops, is accounted for separately rather than melt together in the form of an aggregate, as is the case in the estimation of a production function in (42) and (43).

Another challenge worth considering is how to circumvent the so-called endogeneity problem, a commonly cited critique in the field of empirical production analysis with respect to estimating primal representations of a production technology such as production or distance functions rather than dual representations such as cost or profit functions (Coelli 2000). The issue concerns the fact that the input quantities in estimating the production and distance functions are treated as exogenous rather than as cost-minimizing or profit-maximizing quantities derived from the calculus of the producers. This, it is argued, can lead to biased estimates of the technological parameters.

The bias can be avoided by estimating the production or distance function within a system of equations, which also includes input demand functions and (or) output supply functions, by the method of seemingly unrelated regressions e.g. (Antràs 2004). The extent to which the endogeneity bias would be problematic here should be investigated with respect to the possibility of ordinary least squares providing unbiased estimators as is the case in the model of Zellner, Kmenta, and Drèze under normality assumptions (Zellner, Kmenta, and Drèze, 1966).

6.8 Remarks

This analysis demonstrates that an estimation of a primal representation of a state-contingent production technology is possible. The significance attached to a potential bias due to endogeneity is understandably high in production analysis studies due to the nature of the research question, e.g. would a percentage change in the input factor capital or a percentage change in the input factor labor be more important in perceptually raising output. MELCY, however, interprets the estimated primal representation as an empirically implied technological condition in the optimization problem, the solution of which is the effort-cost function. The estimated primal representation is thus presumably the technological limitation the agent *perceives*, an interpretation consistent with the concept of the revenue-cost function (15) in the description 3.1 of the state-contingent approach.

Thus, a bias is not of primary importance because the model proposed in chapter 5 targets to replicate production decisions. The farmer could perceive his production technological limitations in a biased way and make suboptimal production decisions due to this perception bias. Avoiding a bias in the estimates is thus secondary, should the predictive performance of MELCY be deemed satisfactory. Evaluating the effects of the biased perception of production technological constraints on the production decisions and, ultimately, farm income, might be a fruitful topic of further research.

7 Empirical substantiation of the assumed dynamics of MELCY

This chapter outlines an approach to substantiate the assumed dynamics of MELCY, the model proposed in chapter 5. Such a substantiation involves an appropriate partitioning of the data as well as the application of a combination of econometric and statistical methods in order to identify important, yet not directly observable, parameters of MELCY, which presumably influence one of the proposed predictors of enterprise-level crop yields.

As mentioned in chapter 5, the model relies on the idea that enterprise-level crop yields $y_{ms}^{(k)}(t)$ in an agricultural cycle t can be represented as a function f of:

- biophysically determined yields, $y_{ms}^{biophy}(t)$, which would be determined by Nature choosing the weather conditions prevalent in t, and
- $y_{ms}^{*(k)}\left(t|\Theta_{t-1}^{(k)}\right)$, the target crop yield the producer should be striving for based on the environmental and market information available to him up to this point in time.

The first argument of the function f, the biophysically determined crop yields $y_{ms}^{biophy}(t)$, can be seen as directly inferable. They can either be approximated by crop experimental results at relevant agricultural experimental stations or by simulated values resulting from an appropriately calibrated crop simulation model. A proxy for future yields can be obtained by an appropriately calibrated crop simulator, which uses the weather predictions of a suitable climate model.

The second argument of the function f, the production aspirations $y_{ms}^{*(k)}\left(t|\Theta_{t-1}^{(k)}\right)$, are presumably formed based on the environmental and market information available to the producer and cannot be observed directly. As indicated in Figure 4, the producer is assumed to perceive and remember:

- insights on his production and the production of his geographical neighbors,
- information on biophysical yields which have occurred,
- information on the states of nature which have occurred,
- information on input and output price levels he himself was confronted with.

The producer is furthermore assumed efficient and utility maximizing in the state-contingent sense. He is presumably sensitized to the possibility of variations in his production due to environmental changes that could alter the extent to which state-allocable inputs influence output. This is assumed to prompt him to evaluate and re-evaluate his input choices and production outcomes in the manner exemplified in chapter 6. He presumably adapts to perceived changes in the production technology as well as to changes in the biophysical and market environments by adjusting his production aspirations and input commitment.

Calculating the production aspirations $y_{ms}^{*(k)}\left(t|\Theta_{t-1}^{(k)}\right)$ is more challenging. Inferring essential parameters, which presumably contribute towards their formation, from enterprise level accounting data is vital in order to be able to reconstruct $y_{ms}^{*(k)}\left(t|\Theta_{t-1}^{(k)}\right)$. A reconstruction of the production aspirations $y_{ms}^{*(k)}\left(t|\Theta_{t-1}^{(k)}\right)$ is necessary to approximate future levels of the production aspirations $y_{ms}^{*(k)}\left(t+i|\Theta_{t+i-1}^{(k)}\right)$ and to make a prediction about the enterprise-level crop yields $y_{ms}^{(k)}(t+i)$.

Two critical parameters for the production aspirations determination are the expectation adjustment parameter λ in (17) and η, a risk-aversion parameter, whose definition is first proposed in this thesis in equation (45) and will be defined here based on Figure 3. Both parameters are intrinsic to the farmer and time-invariant but none of them is directly observable. How the farmer is assumed to evaluate his production-technological constraints is demonstrated in chapter 6. Chapter 7 is dedicated to inferring the expectation adjustment parameter λ and the risk-aversion parameter η. Section 7.1 outlines the data requirements. Section 7.2 outlines the identification procedure assuming two crops and two states of nature.[5] Section 7.3 discusses the procedure and remarks on challenges and potentials.

7.1 Data requirements

The model seeks to meaningfully integrate the type of data described in Table 10. The weather data is needed in order for the crop model to simulate observations of biophysical crop yields (code 04). Observed weather data (code 01) can be used as a

[5]The procedure is generalizable with respect to the number of crops and nature-states.

feed for a crop model in case agronomic experimental stations, and thus observed biophysical yields (code 03), are missing. Simulated weather data (code 02) under a changing climate can be used to obtain projections for the development of the biophysical yields, which are necessary if MELCY is to be used for predictive purposes. Simulated weather data can be generated under different climate change scenarios in order to compare the predictions of MELCY across scenarios. Simulated input and output prices (code 07) are also needed in order to generate MELCY predictions. For illustration purposes the prices could be modeled as geometric Brownian motions, the stochastic process used for instance in (Black and Scholes, 1973).

The observed biophysical yields (code 03) are needed in order to

- estimate the primal representation of the state-contingent production technology,
- identify the states of nature and
- determine the likelihood for their occurrence.

The estimation of the primal production representation also requires farm-level accounting data (code 05).

Table 10 MELCY data requirements.

Data	Type	Origin	Use	Code
Weather data	Observed	Weather stations	Input for crop models in case observed biophysical crop yields (code 03) are missing due to a lack of experimental stations	01
Weather data	Simulated	Climate model	Input for crop models in order to simulate biophysical crop yields (code 04) under different climate change scenarios	02
Biophysical crop yields	Observed	Agronomic experimental stations	Needed for (*i*) the estimation of a primal production technology representation for the identification of nature-states. Also needed for (*ii*) the determination of the relative frequency of state occurrence, which presumably determines the probability perception of the farmer	03
Biophysical crop yields	Simulated	Crop model	Needed for MELCY predictions. The biophysical crop yields simulated using 01 can be employed instead of 03 if 03 is missing.	04
Farm-level accounting data	Observed	Farm-level accounting data collectors	Needed for the estimation of a primal production technology representation and the determination of the expectation adjustment parameter λ and the risk-aversion parameter η	05
Input and output prices	Observed	Data collectors	Needed for the determination of the expectation adjustment parameter λ and the risk-aversion parameter η	06
Input and output prices	Simulated	Price model	Needed for MELCY predictions	07

Source: own illustration.

7.2 Steps in the identification procedure

This section describes the steps in the general computation strategy developed to infer the expectation adjustment parameter λ and the risk-aversion parameter η. Inferring the parameters would allow to approximate future levels of the production aspirations $y_{ms}^{*(k)}\left(t+i|\Theta_{t+i-1}^{(k)}\right)$, which are necessary in order to make a prediction on future levels of enterprise-level crop yields $y_{ms}^{(k)}(t+i)$.

The observed data, whose type is described in Table 10 with codes 01, 03, 05 and 06, is therefore separated into two datasets, as illustrated in Table 11:

- an initialization dataset, which contains the available observations from periods 1 to τ,
- a tuning dataset, which contains the available observations from periods $\tau+1$ to T.

The agricultural periods refer to the term of an agricultural cycle introduced in chapter 5 and illustrated in Figure 4.

Table 11 Partitioning of the data into datasets.

	Agricultural Periods
Initialization	1 to τ
Tuning	$\tau+1$ to T
Prediction	$T+q, q=1,\dots,\infty$

Source: own illustration.

Table 12 summarizes the phases and steps of the implementation as well as the data requirements at each step using the codes introduced in Table 10. For convenience Table 12 contains in its fourth column pointers to the subsections, where the corresponding step is presented.

Table 12 Steps in the procedure and their data requirements.

Phase	Step	Data types Required	Subsection
Pre-processing	Partitioning of the data into initialization and tuning datasets.	01, 03, 05, 06	7.2.1
	Simulation of the biophysical crop yields under future climate conditions, which is needed for the prediction of enterprise-level crop yields.	02, 04	7.2.2
	Selection of the functional form of the production technology and derivation of the revenue-cost function	None	7.2.3
Initialization	Nature-state identification	03 or 04	7.2.4
	Estimation of the coefficients of a primal representation of the production technology	03, 05	7.2.5
	Determination of the relative frequency of state occurrence	03	7.2.6
Tuning	Calibration of the revenue cost function	06 and estimates from 7.2.5	7.2.7
	Determination of the expectation adjustment and risk aversion coefficients	Results from 7.2.5, 7.2.6 and 7.2.7	7.2.8
	Fitting of the enterprise-level crop yields	Results from 7.2.8 and 03	7.2.9
Prediction	Prediction	02, 04, 07	7.2.10

Source: own illustration.

7.2.1 Data partitioning

Available observations of the data-types 01, 03, 05 and 06, as described in Table 10, are aligned along a timeline and separated into initialization and tuning datasets. The initialization dataset is to be used for the nature-state identification (subsection 7.2.4), the initial estimation of the coefficients of a primal representation of the state-contingent production technology (subsection 7.2.5) and the initial determination of the relative frequency of nature-state occurrence (subsection 7.2.6). The first period in the tuning dataset is to be used when determining the expectation adjustment parameter λ and the risk-aversion parameter η (subsection 7.2.8). The tuning dataset is also used to fit the enterprise-level crop yields (subsection 7.2.9).

7.2.2 Simulation of biophysical crop yields based on the predictions of a climate model

Predictions on future biophysical crop yields are obtained under the predicted weather conditions, which are simulated using a climate model for a climate change scenario of interest. This effectively constitutes the assumption that there is no feedback of agricultural production to the climate system. This assumption is implicit in the farm-level programming models presented in subsection 2.1.3 and section 4.2 and appropriate for enterprise level analysis due to the relatively small output volume.

7.2.3 Selection of the functional form and derivation of the revenue cost function

A functional form for the primal representation of the state-contingent production technology is chosen, thereby selecting the basis for the empirical specification to be estimated. The primal representation of the state-contingent production technology is substituted as the production technological condition in the minimization problem given in (14), the solution of which results in the effort-cost function.

Depending on the functional form chosen, the minimization problem might be solvable by means of the Lagrangian multiplier method. If there does not exist a closed form solution, a numerical method has to be used instead. This thesis, for instance, uses a Cobb-Douglas formulation, which implies self-duality (Coelli, 2000), (Chambers, 1988, p. 91).

If a distance function is chosen as the primal functional representation and the corresponding empirical specification estimated, then every output can be taken into account separately as mentioned in section 6.7. The revenue-cost problem in (15) is considered. Reaching the revenue-cost function might be achievable by means of a modification of the simplex method once the revenues and prices are set, since the target revenue restrictions are linear. Reaching the revenue-cost function is important in order to reconstruct the set of production alternatives as it is presumably perceived by the producer.

7.2.4 Identification of the states of nature

The biophysical crop yields in the initialization dataset are set in relation to one another year-wise, e.g. the biophysical mean of barley for the period 1996 is divided by the biophysical mean of wheat for the period 1996. The ratios are constructed for every period $\tau - h,\ h = 0, \dots, \tau - 1$ in the initialization dataset. In the case of two crops:

$$v_{\tau-h} = \frac{mean(crop_1)_{\tau-h}}{mean(crop_2)_{\tau-h}} \tag{44}$$

The goal is to isolate the states of nature defined as being favorable for the production of a specific crop. It is assumed that the farmers attribute their past production experiences to one of the states according to the year the production experiences occurred in. These ratios $v_{\tau-h},\ h = 0, \dots, \tau - 1$ are grouped using a clustering algorithm like the partitioning algorithm described in subsection 6.3.2.

Once the grouping results are deemed satisfying, the assignment of periods to groups, which would correspond to states of nature, is undertaken. The objects representative of the clusters, $m^{(1)}, \dots, m^{(S)}$, are assumed to comprise the view of the farmer on what constitutes the states of nature. The farmer's perception of the states of nature is assumed complete at this point.

Any biophysical crop yield ratio observed (after the perception of the nature-states is set) is presumably attributed to one of states. This attribution can be completed via discriminant analysis since the groups are already known a priori (Härdle and Simar, 2003, p. 289).

7.2.5 Estimation of the state-contingent production function

The empirical specification corresponding to the functional form of the state-contingent production technology chosen in subsection 7.2.3 is fitted using the initialization dataset. It should be determined whether the estimated function has the theoretical properties of a production or a distance function. This is important because the estimated coefficients will be used to calibrate the revenue-cost function. The convexity of the set of production alternatives described by the calibrated revenue-cost function cannot be guaranteed if the theoretical properties are not fulfilled. In case the estimated function does not exhibit the expected theoretical properties the dataset should be thoroughly inspected for data anomalies.

7.2.6 Determination of the relative frequency of nature state occurrence

The relative frequency of detected nature-state occurrence $f_s,\ s \in \Omega$ is determined within the initialization dataset. This relative frequency is assumed synonymous to the initially perceived probabilities of state occurrence, which would constitute the initial probability perception of every farmer and determine the initial slope of the indifference curves of a farmer in case of risk-neutrality (section 3.1). The perceived probabilities of state-occurrence are assumed to develop over time as new biophysical crop yield ratios are considered.

7.2.7 Calibration of the revenue cost function

In this step the revenue-cost function in the form analytically derived in subsection 7.2.3 should be calibrated using:

- the estimated coefficients of the primal state-contingent technology representation obtained in subsection 7.2.5 (in this thesis the coefficients are reported in Table 9, section 6.8),
- the input prices, which can be set to the values $w_n(\tau),\ n = 1, \dots, N$ observed in τ,
- the target revenues, which can be set for instance at the level of fixed costs last observed in the specific state of nature s,
- the initial state-specific output price expectations, which can be set equal to the state-specific output prices $p^e_{ms}(1) = p_{ms}(1)$,

- the state-specific output prices $p_{ms}(t_s)$, which can be set to the last values observed in each state of nature s.

After a comparison of the so calibrated revenue-cost function with (14), (15) and (17) it can be seen that the only coefficient preventing the plotting of an iso-cost curve like the one in Figure 3 is the expectation adjustment coefficient λ, which plays a role in the output price expectation formation. An approach to avoid making assumptions about the target revenues and the initial state-specific output price expectations is discussed in section 7.3.

7.2.8 Determination of the expectation adjustment and risk aversion coefficients

Expectations regarding the future levels of the state-contingent output prices are presumably constructed in the manner described in (17). The expectation adjustment coefficient can be presumed non-state specific for the sake of computational simplicity.[6] The appropriate value of the expectation adjustment parameter λ and the risk aversion parameter η introduced below in (45) will be determined jointly.

The risk aversion parameter η is graphically based on Figure 3. In order to facilitate reading Figure 3 is introduced again as Figure 8. The expectation adjustment parameter λ plays a role in shaping the iso-cost curve in Figure 8.

The indifference curve of a risk neutral producer in Figure 8 coincides with the 'fair-odds' line, the slope of which gives the relation between the subjective probabilities of state occurrence. This slope is assumed determined by the relative frequency of nature-state occurrence detected in the initialization dataset (subsection 7.2.6).

As it has been mentioned in section 3.1, the decision of the extremely risk averse producer would still lie along the bisector, where state-contingent revenues are equal in an effort to avoid any fluctuations in them, even if he perceives the same probability of state occurrence as the risk neutral producer. A producer with some degree of risk aversion will choose a point on the iso-cost curve which is bounded between the production choices of the risk neutral and the extremely risk averse producers.

[6] Possible relaxations of this assumption and their computational implications are discussed in section 7.3.

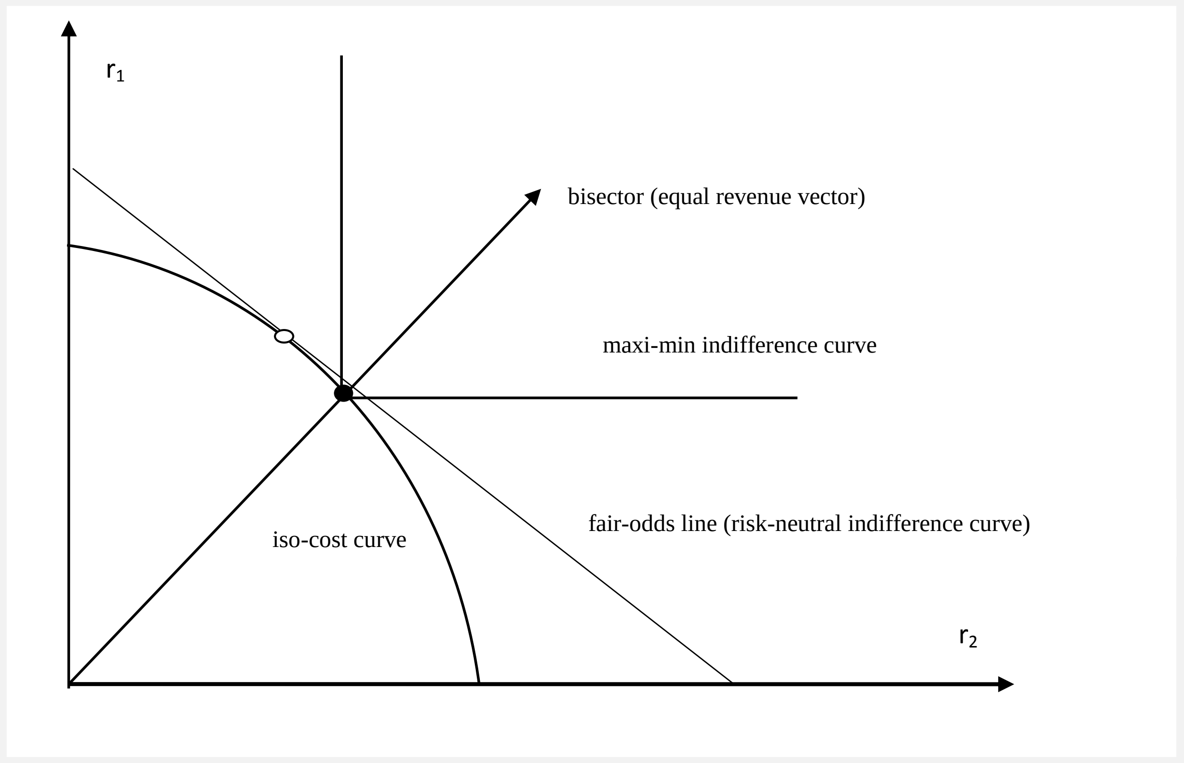


Figure 8 The production decision with two states of nature.
Note: The production decision of the extremely risk-averse producer is depicted by the black dot, the production decision of the risk neutral producer is given by the white dot.
Source: Chambers and Quiggin (2000), page 179.

It is assumed that the set of production alternatives described by the calibrated iso-cost curve (subsection 7.2.7) is a convex one, which is safe-guarded by the estimated function in subsection 7.2.5 exhibiting the desirable theoretical properties. Then, the production decision of a risk averse producer (grey dot on the iso-cost curve in Figure 9) can be represented as a combination of the decisions of the risk-neutral one (white dot on the iso-cost curve in Figure 8 and Figure 9) and the extremely risk-averse producers (black dot on the iso-cost curve in Figure 8 and Figure 9).

If rays connect the origin and the points representing the production decisions in Figure 9, then the risk-aversion coefficient is defined as $\eta \in [0,1]$:

$$\gamma = (1-\eta)\alpha + \eta\beta \tag{45}$$

where γ is the polar angle consistent with the production decision of a producer with a coefficient of risk-aversion η; α is the polar angle consistent with the production decision of a risk-neutral producer and β is the polar angle consistent with the production decision of an extremely risk-averse producer.

A specific value of η signifies a point on the iso-cost curve between the production decisions consistent with risk-neutrality (given by the white dot and consistent with α in Figure 9) and extreme risk-aversion (given by the black dot and consistent with β in Figure 9). This section of the iso-cost curve between the white and black dots can be grid-searched in order to identify η. The points, however, have to be compared to an observable, non-stochastic entity.

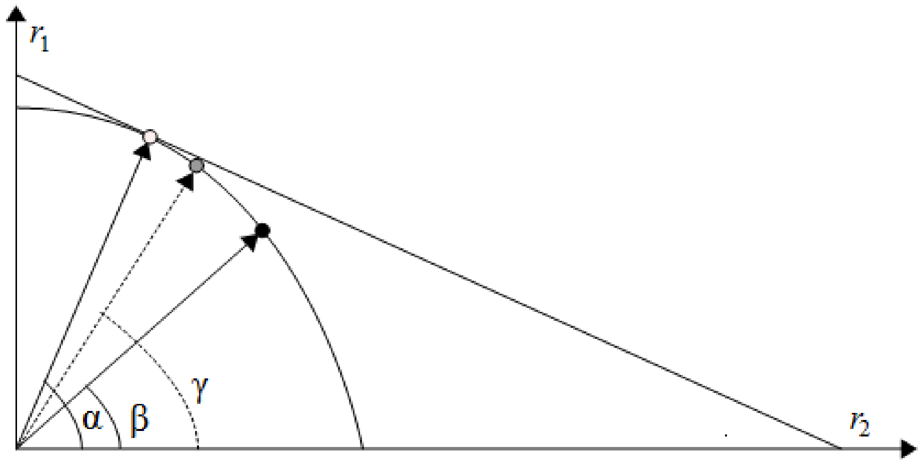


Figure 9 Polar angles consistent with diverse degrees of risk-aversion of the producer.
Note: Similarly to Figure 8 the production decision of the extremely risk-averse producer is depicted by the black dot and the production decision of the risk neutral producer is given by the white dot.
Source: own illustration.

The search of an observable, non-stochastic entity leads to the following observation: each point of the iso-cost curve corresponds to a mix of state-contingent revenues. This mix of state-contingent revenues is connected to a matrix of state-contingent outputs $\boldsymbol{y}$ through (15); $\boldsymbol{y}$ is connected to a cost-minimizing input vector $\boldsymbol{x}$ through (14). Thus, the cost-minimizing input vector $\boldsymbol{x}$, obtained on the basis of the information set $\Theta_\tau^{(k)}$ and some assumed values of the expectation adjustment parameter λ and the risk-aversion parameter η, can be compared to the input vector $\boldsymbol{x}$ observed in $\tau + 1$. In other words, what producer k should have committed in $\tau + 1$ given specific values of λ and η is compared to what producer k actually committed in $\tau + 1$. A choice of a distance measure should be made a priori.

To put it in a different way, the numerical identification of the appropriate values of the risk-aversion and the expectation adjustment coefficients requires considering that the revenue-cost function is calibrated based on an estimation, which takes into account the initialization dataset. The strategy for identification consists of grid-searching a list of possible values for the two parameters, determining the production aspirations consistent with each specific parameter value pair and determining the cost-minimizing input amounts to produce the production aspirations consistent with each specific parameter value pair. The cost-minimizing input amounts can be compared to the observed input amounts in the tuning dataset. A specific parameter value pair is selected based on how well the calibrated model reproduces the input commitments in the tuning dataset.

For given specific values of λ and η, the aspirations of the producer are a fully-defined point on the iso-cost curve in the space of state-contingent revenues illustrated by Figure 8.

These aspirations would take the form:

$$((r_1^*, r_2^*)|(\lambda, \eta)) \tag{46}$$

Having certain values for the state-contingent output prices p_{ms} would lead to the aspirations in the form of state-contingent output aspirations:

$$\{\left(\left(y^*_{crop_1,1}, y^*_{crop_2,1}\right), \left(y^*_{crop_1,2}, y^*_{crop_2,2}\right)\right)|(\lambda, \eta)\} \tag{47}$$

where

$y^*_{crop_1,1}$: aspirations of the producer for crop 1 in the first state of nature,

$y^*_{crop_2,1}$: aspirations of the producer for crop 2 in the first state of nature,

$y^*_{crop_1,2}$: aspirations of the producer for crop 1 in the second state of nature,

$y^*_{crop_2,2}$: aspirations of the producer for crop 2 in the second state of nature.

Regardless of the form of the production aspirations they can be cost-effectively produced by a certain combination of inputs, which is determined by the input prices and the production technology, when the production technology might be expressed either in terms of a production or a distance function. This fact can be used for the numerical identification of the appropriate values of the risk-aversion and the expectation adjustment coefficients.

For example, the value pair (0.9,1) would correspond to an extremely risk-averse producer, who attaches a great significance to past deviations between expected and realized state-contingent output prices. The production aspirations in terms of state-contingent outputs, based on the data in the initialization dataset are identified as:

$$\left\{\left(\left(y^*_{crop_1,1}, y^*_{crop_2,1}\right), \left(y^*_{crop_1,2}, y^*_{crop_2,2}\right)\right) \middle| (0.9, 1)\right\}_\tau \tag{48}$$

These aspirations can be produced in a cost-minimizing manner with the input quantities $(x^*_n(\tau+1)|(0.9,1))$, as the effort-cost function would convey, i.e. the quantities $(x^*_n(\tau+1)|(0.9,1))$ are computable and can be compared to the observed input quantities $x_n(\tau+1)$.

Following this brute force strategy, scanning the intervals $[0,1]$ for λ and $[0,1]$ for η with a given increment the production aspirations can be calculated:

$$\left\{\left(y^*_{crop_1,1}, y^*_{crop_2,1}\right), \left(y^*_{crop_1,2}, y^*_{crop_2,2}\right) \middle| (\lambda, \eta)\right\}_\tau \tag{49}$$

The corresponding input quantities $(x^*_n(\tau+1)|(\lambda,\ \eta))$ are derived. A distance z between the predicted cost-minimizing input quantities and the observed input quantities is calculated:

$$z(\lambda, \eta) = \sqrt{\sum_{n=1}^{N}((x^*_n(\tau+1)|(\lambda,\eta)) - x_n(\tau+1))^2} \tag{50}$$

Although λ and η are calculated for period $\tau+1$, as a maintained hypothesis these parameters are considered constants for all periods. A number $z(\lambda,\eta)$ is juxtaposed to each pair (λ,η). The value pair for λ and η, which minimizes the selected distance measure between predicted and observed input commitments, is chosen. The optimal value pair is referred to as (λ^*, η^*), i.e. $(\lambda^*, \eta^*) = arg\ min_{(\lambda,\eta)}\ z(\lambda,\eta)$.

7.2.9 Fitting of the enterprise-level crop yields

Once the optimal estimated value pair (λ^*, η^*) has been determined for each producer k, a proxy for the production aspirations of each producer k can be obtained for the remaining observations in the tuning dataset. To achieve this steps 7.2.5, 7.2.6 and 7.2.7

are repeated, adding observations of the data types 03, 05 and 06 period-wise. Observations of the biophysical yield ratio are attributed to the nature-states determined in 7.2.4 via discriminant analysis.

Thus the production aspirations for each period i in the tuning dataset are calculable and take the form:

$$\left\{\left(y^*_{crop_1,1}, y^*_{crop_2,1}\right), \left(y^*_{crop_1,2}, y^*_{crop_2,2}\right)\middle|(\lambda^*,\ \eta^*)\right\}_{\tau+i} \tag{51}$$

This way the second argument of the function f is also determined and all the functional values and the values of both arguments are known for the periods in the tuning dataset.

The ideal relationship corresponding to MELCY in chapter 5 can then be modeled:

$$y^{(k)}_{ms}(\tau+i+1) = f\left(y^{biophy}_{ms}(\tau+i+1), y^{*(k)}_{ms}\left(\tau+i+1\middle|\Theta^{(k)}_{\tau+i}\right)\right) \tag{52}$$

where

$y^{(k)}_{ms}(\tau+i+1)$: yields of crop m for a producer k in period $\tau+i+1$ under state s,

$y^{biophy}_{ms}(\tau+i+1)$: biophysical yields of crop m in period $\tau+i+1$ under state s,

$y^{*(k)}_{ms}\left(\tau+i+1\middle|\Theta^{(k)}_{\tau+i}\right)$: production aspirations of the producer k with regard to crop m in period $\tau+i+1$ under state s.

Plotting the obtained values for the biophysical crop yields and the production aspirations could help determine the functional form appropriate for modeling. The function can be interpolated with some well-known class of functions, e.g. splines. A multiplicative form of the Cobb-Douglas type would however provide a first suggestion – a period with a natural disaster (i.e. biophysical yields approaching zero) should lead to zero outputs, regardless of what the producer is striving to achieve; similarly, a beneficial period (i.e. high biophysical yields) should lead to zero outputs, if the producer is not trying to produce anything (i.e. aspirations equal to zero). Once the functional form is chosen the coefficients can be estimated econometrically.

7.2.10 Prediction

The functional form estimated in subsection 7.2.9 serves as a basis to simulate enterprise-level crop yields in the future. This is possible since the production aspirations are constructed based on information already available from previous periods and the biophysical crop yields are independent of production choices, should crop rotation be disregarded. A relaxation of the letter condition and the implications are discussed in section 7.3.

The production aspirations and the cost-minimizing input commitments absorb the biophysical as well as the market information available up to this point in time – the probability perceptions with respect to state occurrence, the way the production technology is perceived by the producer, the input prices, the output price expectations and the target revenues. The biophysical crop yields account for the changes in the natural environment.

7.3 Remarks

This chapter presented a sequence of steps to substantiate the model MELCY proposed in chapter 5. It elicits the unobservable production aspirations among others through a partitioning of the available data into an initialization dataset and a tuning dataset. The procedure thereby uses the fact that from the point of view of the researcher all the observed data is known, while the realizations in period $\tau + 1$ were unknown to the farmer making production choices in period τ.

The model selected in subsection 7.2.8 could be tested given that sufficient amount of data is available for some longer period of time. While a good predictive performance cannot ultimately confirm the validity of the model, it provides empirical evidence in its favor. It should be investigated how possible simulation errors, which MELCY would absorb through the use of data types 02, 04 and 07, could influence the simulated output of MELCY. To test the empirical performance of MELCY the observed data can be separated into not two, but three datasets: initialization dataset, tuning dataset and testing dataset. In terms of number of observations belonging to a dataset the number is somewhat arbitrary. Another discipline, machine learning, can offer guidelines: 50% of the data is used to train an algorithm, another 25% of the data are used for validation and the remaining 25% for testing (Hastie, Tibshirani, and Friedman, 2011, p. 222). In the case of MELCY this would mean committing 50% of the data to initialization,

another 25% to tuning and the last 25% to testing. As Hastie, Tibshirani, and Friedman remark the specific number of observations assigned to a dataset should depend on the task (Hastie, Tibshirani, and Friedman, 2011, p. 222). In the case of MELCY the tuning dataset should be sufficient in order to elicit enough observations on the production aspirations in order to merit the econometric analysis in subsection 7.2.9. The number of observations needed depends on the functional form chosen and the number of coefficients, which are to be thereby estimated.

Subsection 7.2.7 poses assumptions on the target state-contingent revenues and the initial state-specific output price expectations. It is possible in principle to include those values in the grid search introduced in subsection 7.2.8. While grid-searching the optimal value of the expectation adjustment parameter λ relies on searching along the same iso-cost curve, grid searching possible values for the target state-contingent revenues and the initial state-specific output price expectations would rely on searching different iso-cost curves. The optimal values for the state-contingent revenues can be compared to the accounting records of the data type 05 in order to see what magnitude the elicited values resemble. Such a modification would require insuring that the tuning dataset is large enough: every state-contingent parameter, including a state-specific expectation adjustment parameter λ, requires at least one observation belonging to the state of nature to identify. Hopefully, suitable larger datasets will be accessible in the emerging era of big data and open public data.

Further development of the model should include first and second round of input commitments within the same agricultural period, in the manner the model presented in subsection 2.1.3.2 describes. Figure 10, which is based on Figure 4, illustrates the suggestion, with the second round of input commitment indicated by the dark box. Thought should be devoted to the way the second round input commitments are identified. While the first round input commitments are optimal *ex ante*, the second round input commitments are optimal *ex post*, i.e. after the uncertainty about the state of nature is resolved. Further research and experimentation with real data are needed in order to tackle the issue effectively.

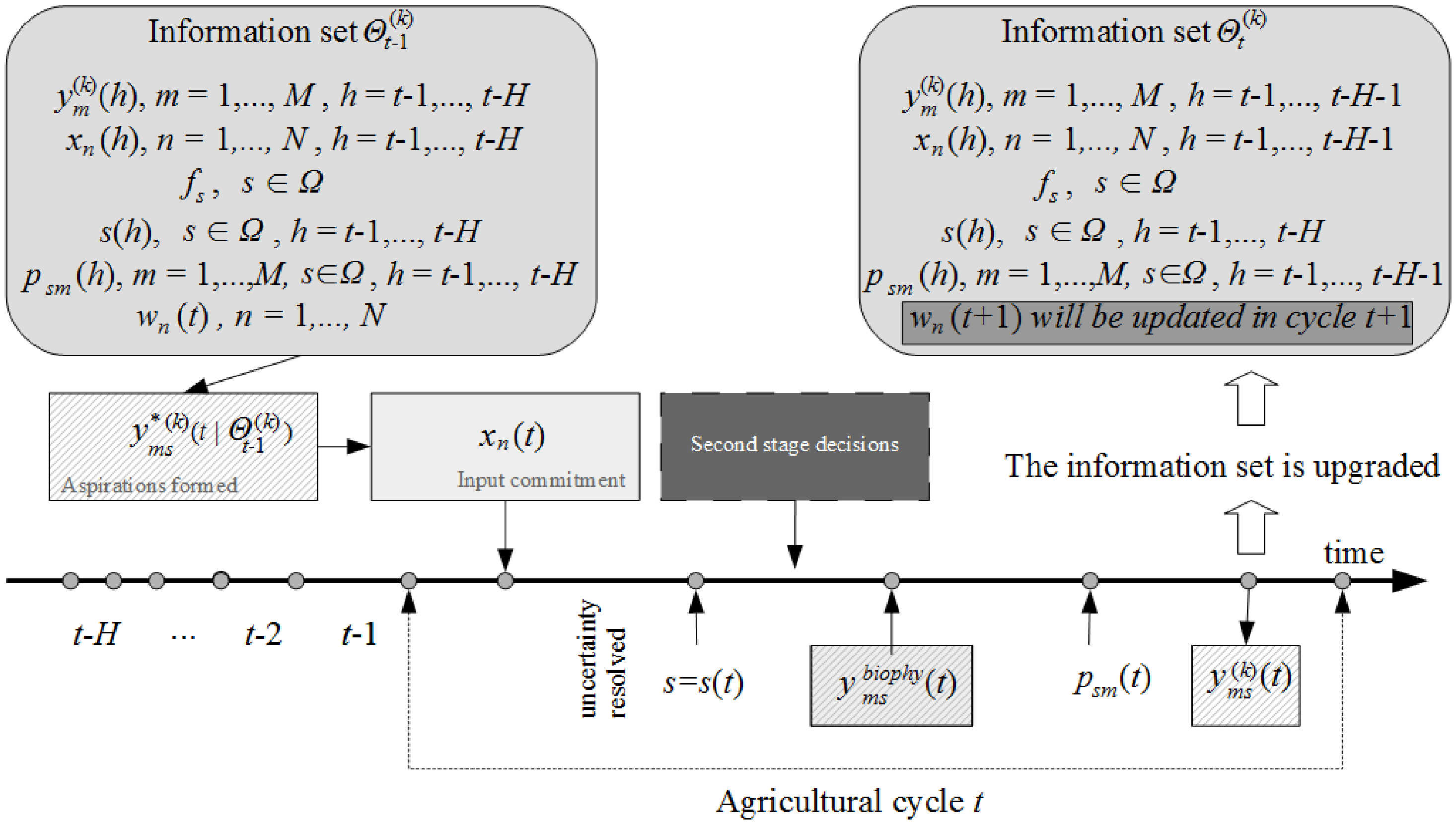


Figure 10 Suggestion for further model development.
Source: own illustration.

8 Conclusion and outlook

As it has been established in chapter 2, assessing the impacts of climate change on agricultural production requires models on different scales of economic activity. The approaches described in subsections 2.1.1 and 2.1.2 rely on spatial comparisons across geographically and climatically diverse areas and extrapolate the findings to probable future agricultural outcomes. Crude estimate of the impacts of climate change after adaptation has occurred are delivered by these approaches.

However, assessing climate change impacts requires understanding the ways and extent to which farmers can adapt to changing environmental conditions, which none of these two approaches targets to explain explicitly. This assessment relies on farm-level mathematical programming models incorporating uncertainty like the models described in subsection 2.1.3 and section 4.2. Crop yields are obtained on the field level through crop simulators.

While reducing crop yields to biophysical environment products is appropriate on the field level, summing up these yields to the enterprise level might lead to distortions, since it neglects essential production determinants, e.g. the wage associated with labor. Such potential biases have to be addressed in order to make the predictions of impact and adaptation models comparable across all scales of economic activity.

The essential production determinants like wage are accounted for in the production analysis studies in the field of agricultural economics, like the studies based on the state-contingent approach described in section 4.1. The direct applicability of the results of the production analysis studies to the field of climate change impact assessment is, as it has been established in section 4.3, restricted by basic premises posed by the production economic framework. Production economic elements can, however, be helpful in addressing the identified potential biases across scales of economic activity.

Hence, MELCY - an enterprise-level model for crop yields under climate change - is proposed in chapter 5. Similarly to Crean et al. (2013), the mathematical programming farm-level model described in section 4.2, which reports positive gains from state-contingent modeling, MELCY is based on the state-contingent approach presented in section 3.1. Unlike a mathematical programming model, which explicitly shows the decision-making of the farmer and subsequently simulates field crop yields, MELCY represents the enterprise level crop yields as a function of average crop yields per field and optimal production aspirations under production-economic constraints. It thus integrates the agronomic and economic perspectives of crop yields.

With respect to the empirical implementation, a new indicator - the biophysical yield ratio - is proposed to detect the states of nature as well as two statistical approaches to evaluate it. As section 6.3 observes the proposed hierarchical clustering algorithm proposed is suitable for the nature-state detection if the researcher has not a priori decided on the number of nature states and wishes to explore potential meaningful groupings within the dataset containing the biophysical yield ratios. If, on the other hand, a researcher has a strong theoretical argument concerning the number of nature-states, then the application of a robust partitioning algorithm is appropriate. Interpreting the results of hierarchical clustering introduces an element of subjectivity into the analysis, while the use of a robust partitioning algorithm would reveal an objectively optimal grouping, yet will not reveal alternative interpretations of the dataset.

This chapter is divided into two parts. Section 8.1 reports the results in a manner consistent with the research questions and hypotheses set in chapter 1. Section 8.2 provides an outlook and plans for future work.

8.1 Results and discussion

The ultimate goal of this thesis - developing a model for enterprise-level crop yields under climate change, providing a method for the calculation of the model and demonstrating the model implementation using available data for the case of Saxony-Anhalt - is achieved.

The first specific objective of this thesis - proposing a model for enterprise-level crop yields under climate change, which integrates the economic and agronomic notions of crop production in a coherent whole - is reached with the introduction of MELCY in chapter 5. A general strategy for the computation of MELCY, which demonstrates its empirical tractability, is provided in chapter 7. MELCY overcomes the posed conceptual challenge by combining the state-contingent approach with a hypothesis of how economic agents form expectations on the level of relevant variables under uncertainty, namely the hypothesis of adaptive price expectations. A partial implementation for the case of crop production in Saxony-Anhalt, an estimation of the coefficients of the primal representation of the state-contingent production technology, is presented in chapter 6.

Chapter 5 demonstrates that the insights of the state-contingent approach are applicable to the field of climate change impact assessment, as it has been demonstrated by the construction of MELCY. The latter also demonstrates that it is possible to overcome the conceptual challenge posed by the state-contingent approach not being explicit on how agents form expectations about the future levels of output prices. The analysis outlined in chapter 6 rejects the hypothesis of the crop producing farmers in Saxony-Anhalt operating under an output-cubical technology to the extent to which the simulated data match the unknown real data.

With respect to the framing as research hypotheses the following has been achieved:

- Chapter 5 demonstrates that the construction of a model, which integrates both economic and agronomic notions of crop production under climate change, is possible. The constructed model is computable, as demonstrated in chapter 7.
- The conceptual challenge posed by the limitations of the state-contingent approach can be overcome, as chapter 5 demonstrates.
- The hypothesis of an output-cubical production technology in the case of the crop production sector in the Federal State of Saxony-Anhalt is rejected. This result holds to the extent to which the simulated data successfully reconstructs key characteristics of the underlying statistical population.

Overall, this thesis delivers three distinct contributions:

- It establishes the need for a novel enterprise-level crop yield model under climate change. It delivers such a model, which fully operationalizes the state-contingent approach as a tool of production analysis and decision-making under uncertainty by combining it with the hypothesis of adaptive price expectations. The thesis outlines a general computational strategy for the model.
- This work introduces a new indicator to statistically detect the occurrence of nature-states, an alternative to the approach proposed by Nauges, O'Donnell and Quiggin (2011) based on an earlier work by the author. It also proposes two statistical approaches to process the introduced indicator and explores their advantages and disadvantages.

- A “geometrical interpretation” of the degree of risk-aversion, which is intrinsic to the farmer, is offered. An approach to infer the degree of risk-aversion from farm-level accounting data is proposed. The transformation into a polar coordinate system facilitates the identification of the parameter.

The proposed model is computable and would

- provide recommendations to crop producers with respect to optimal input quantities under the changing environmental conditions and
- make a prediction with respect to enterprise-level crop yields under climate change.

MELCY fills in the identified gap in existing models in climate change impact research: it accounts for crop yields as the products of the biophysical environment and input factors on an enterprise level, rather than selecting few economic production factors and reducing the crop yields to field-level products of the biophysical environment like the mathematical programming models as described in subsection 2.1.3 and section 4.2.

The function predicting the crop yields (52) accounts for the changes in climate via the biophysical crop yield term for each crop separately, which accounts for the fact that crops react to changes in climate differently. Through the introduction of the notion of production aspirations the function (52) accounts for changes in the prices of the enterprise-level factors such as wage. Such changes are generally unaccounted for in mathematical programming models.

Another specific aspect of the proposed model MELCY is that it accounts for an evolution in the perceived probabilities of state-occurrence based on the relative frequency of occurrence of climatic events as encountered by the producer. This evolution in the perceived probabilities of state-occurrence builds onto the non-evolutionary expectation formation accounted for in the mathematical programming model of Dono and Mazzapicchio presented in subsection 2.1.3.2. A non-evolutionary formulation can nevertheless be accounted for by keeping the perceived probabilities at the values of the relative frequencies $f_s, s \in \Omega$ initially determined in subsection 7.2.6. A comparison between the two formulations, an evolutionary and non-evolutionary one, could then be used to investigate the significance of expectation formation in shaping the production aspirations of farmers and, thereby, for the resulting enterprise income.

The proposed model MELCY relies on using a coefficient of risk-aversion estimated from farm-level accounting data and thus tuned to the specific case, unlike the mathematical programming models presented in subsection 2.1.3 and section 4.2.

MELCY also relies on an estimate of an expectation-adjustment coefficient. The model, however, does not account for output price volatility. Another limitation of the model is the fact that it can only make a prediction about crop yields in the original portfolio of crops. The introduction of new crops in the original production portfolio might prove problematic due to difficulties in obtaining the initial estimates for the production coefficients in subsection 7.2.5. Obviously the full implementation of MELCY will require further research and extensive study of its application performance, in case the necessary data become available.

8.2 Outlook and plans for future work

With respect to future work, the methodical contribution introduced in section 6.6 concerning the detection of nature-states should be systematically compared with the approach to detection proposed by Nauges, O'Donnell and Quiggin (2011). Using phenological observations rather than weather data to infer the prevalent environmental conditions, as in (Angelova, 2015), seems by itself promising. This claim is supported by the recent results of Dalhaus and Finger (Dalhaus and Finger, 2016), who apply a similar method to detect the occurrence of a phenomenon insured against in the context of weather index-based insurances.

A distance function, rather than a production function, should be estimated, which would allow for the outputs to be accounted for separately, rather than melt together in the form of an aggregate. The potential bias, which results from the endogeneity problem, should be quantified for the stochastic case. The possibility of accounting for a second round of managerial decisions within the same agricultural period, as suggested in section 7.3, should be considered.

Testing the predictive performance of the model with respect to past enterprise-level crop yields and input commitment is a central topic for future research. Comparing the performance of MELCY to the performance of a mathematical programming model is another vital topic for future research. Both would rely on the data described in Table 10 be complete. Fulfilling the data requirements for a region in order to construct a complete picture is far from trivial since it involves simultaneously having weather and agronomic data as well as farm-level accounting data for the same span of time.

Last, but not least, another central topic for future research should be addressed: comparing the results of the different models for climate change impact assessment as suggested in section 2.2. The point estimates provided by the Ricardian approach and the panel data analysis proposed by Deschênes and Greenstone (2007) can be used to obtain a figure for the change in cumulative agricultural yields in a county at a specific point of time in the future under a scenario of interest. Changes in the cumulative agricultural yields in a county can, on the other hand, be calculated using multiple suitably diversified farm-level mathematical programming models or enterprice-level models like MELCY for the same climate change scenario. Ensuring comparability of the results as well as conducting an in depth comperative analysis of the monetary assumptions underlying the approaches are both fruitful topics for future work. Both are highly relevant in order to ensure the quality of the recommendations given to policy makers.

References

Adamson, D., T. Mallawaarachchi, and J. Quiggin. 2007. "Water Use and Salinity in the Murray-Darling Basin: A State-Contingent Model." *The Australian Journal of Agricultural and Resource Economics* 51(3): 263–281.

Angelova, D. 2014. "The State-Contingent Approach to Production and Choice under Uncertainty : Usefulness as a Basis for Economic Modeling." In *Proceedings of the FACCE MACSUR Mid-Term Scientific Conference. 01-04 April 2014, Sassari, Italy*. http://ocs.macsur.eu/index.php/Hub/Mid-term/paper/viewFile/233/22 (accessed on November 23, 2016).

Angelova, D. 2015. "State-Contingent Production: The Case of Grain N Saxony-Anhalt." In *Proceedings of the IAMO Forum 2015: Agriculture and Climate Change in Transition Economies, 17-19 June 2015*, Halle an der Saale, Germany. http://projects.iamo.de/fileadmin/veranstaltungen/iamo_forum/2015/Presentations/IAMO_Forum_2015_E3_3_Angelova.pdf (accessed on November 23, 2016).

Angelova, D., T. Glauben, and M. Grings. 2014. "Statistical Identification of Nature-States within the State-Contingent Framework." In *Abstract Book of the FACCE MACSUR CropM International Symposium and Workshop: Modelling Climate Change Impacts on Crop Production for Food Security. 10-12 February 2014. Oslo, Norway*.

Antràs, P. 2004. "Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution." *Contributions in Macroeconomics* 4(1): 1–34.

Arrow, K. J., H. B. Chenery, B. S. Minhas, and R. M. Solow. 1961. "Capital-Labor Substitution and Economic Efficiency." *The Review of Economics and Statistics* 43(3): 225–250.

Bartholomew, D., F. Steele, I. Moustaki, and J. Galbraith. 2002. *The Analysis And Interpretation Of Multivariate Data For Social Scientists*. Florida, USA: Chapman&Hall/CRC.

BELV [Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz]. 2009. *Statistisches Jahrbuch über Ernährung, Landwirtschaft und Forsten der Bundesrepublik Deutschland 2009*. Bremerhaven, Germany: Wirtschaftsverlag NW.

BELV [Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz]. 2012. *Statistisches Jahrbuch über Ernährung, Landwirtschaft und Forsten der Bundesrepublik Deutschland 2012*. Münster-Hiltrup, Germany: Landwirtschaftsverlag.

Berg, E. 2012. "Der zustandsabhängige Ansatz in der mathematischen Optimierung / The state-contingent approach in the context of mathematical optimisation." *German Journal of Agricultural Economics* 61(1): 13–29.

Black, F., and M. Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *The journal of political economy* 81(3): 637–654.

Bundessortenamt. 2000. *Richtlinie für die Durchführung von Landwirtschaftlichen*

Wertprüfungen und Sortenversuchen. Hannover, Germany.

BVEL [Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft]. 2001. *Statistisches Jahrbuch über Ernährung, Landwirtschaft und Forsten der Bundesrepublik Deutschland 2001*. Münster-Hiltrup, Germany: Landwirtschaftsverlag.

BVEL [Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft]. 2004. *Statistisches Jahrbuch über Ernährung, Landwirtschaft und Forsten der Bundesrepublik Deutschland 2004*. Münster-Hiltrup, Germany: Landwirtschaftsverlag.

Cagan, P. 1956. "The Monetary Dynamics of Hyper-Inflation." In *Studies in the Quantity Theory of Money*, Chicago, USA: University of Chicago Press.

Chambers, R. G. 1988. *Applied Production Analysis. A Dual Approach.* Cambridge, UK: Cambridge University Press.

Chambers, R. G., and J. Quiggin. 1998. "Cost Functions and Duality for Stochastic Technologies." *American Journal of Agricultural Economics* 80(2): 288–295.

Chambers, R. G., and J. Quiggin. 2000. "Uncertainty, Production, Choice, and Agency: The State-Contingent Approach." Cambridge, UK: Cambridge University Press, 1–191.

Chavas, J.-P. 2008. "A Cost Approach to Economic Analysis Under State-Contingent Production Uncertainty." *American Journal of Agricultural Economics* 90(2): 435–446.

Chow, G. C. 2011. *Usefulness of Adaptive and Rational Expectations in Economics*. Center for Economic Policy Studies, Princeton University.

Coelli, T. 2000. "On the Econometric Estimation of the Distance Function Representation of a Production Technology." *CORE Discussion Papers* 2000 (42).

Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.). 2014. *IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland.

Crean, J., K. Parton, J. Mullen, and R. Jones. 2013. "Representing Climatic Uncertainty in Agricultural Models - an Application of State-Contingent Theory." *Australian Journal of Agricultural and Resource Economics* 57(3): 359–378.

Dalhaus, Tobias, and Robert Finger. 2016. "Can Gridded Precipitation Data and Phenological Observations Reduce Basis Risk of Weather Index–Based Insurance?" *Weather, Climate, and Society* 8(4): 409–419.

Deschênes, O., and M. Greenstone. 2007. "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather." *The American Economic Review* 97(1): 354–385.

Deschênes, O., and M. Greenstone. 2011. "Using Panel Data Models to Estimate the Economic Impacts of Climate Change on Agriculture." In *Handbook on Climate Change and Agriculture*, eds. A. Dinar and R. O. Mendelsohn. Cheltenham, UK and Northampton, USA: Edward Elgar Publishing, 112–140.

Deschênes, O., and M. Greenstone. 2012. “The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather: Reply.” *The American Economic Review* 102(7): 3761–3773.

Dictionary.com. 2016a. “Cultivar.” http://www.dictionary.com/browse/cultivar?s=t (accessed on October 27, 2016).

Dictionary.com. 2016b. “Ex Ante.” http://www.dictionary.com/browse/ex-ante?s=t (accessed on October 27, 2016).

Dictionary.com. 2016c. “Ex Post.” http://www.dictionary.com/browse/ex-post?s=t (accessed on October 27, 2016).

Dinar, A., and R. Mendelsohn, eds. 2011. *Handbook on Climate Change and Agriculture*. Cheltenham, UK and Northampton, USA: Edward Elgar Publishing.

Dono, G., and G. Mazzapicchio. 2010. “Uncertain Water Supply in an Irrigated Mediterranean Area: An Analysis of the Possible Economic Impact of Climate Change on the Farm Sector.” *Agricultural Systems* 103(6): 361–370.

Evans, G. W., and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton, US and Woodstock, UK: Princeton University Press.

Färe, R., and D. Primont. 1994. *Multi-Output Production and Duality: Theory and Applications*. Norwell, USA: Kluwer Academic Publishers.

Finger, R., and S. Schmid. 2008. “Modeling Agricultural Production Risk and the Adaptation to Climate Change.” *Agricultural Finance Review* 68(1): 25–41.

Fischer, A. C., W. M. Hanemann, M. J. Roberts, and W. Schlenker. 2012. “The Economic Impacts of Climate Change: Evidence from Agricultural Profits and Random Fluctuations of Weather: Comment.” *The American Economic Review* 102(7): 3749–3760.

Fischer, I. 1911. *Purchasing Power of Money*. New York, USA: Macmillan.

Fischer, I. 1930. *The Theory of Interest, as Determined by Impatience to Spend Income and Opportunity to Invest It*. New York, USA: Macmillan.

Fox, J., and S. Weisberg. 2011. “Nonlinear Regression and Nonlinear Least Squares in R.” In *An R Companion to Applied Regression*, Thousand Oaks, USA: SAGE Publications, 1–20.

Friedman, M. 1957. *Theory of the Consumption Function*. Princeton, USA: Princeton University Press.

Gavin, H. P. 2013. “The Levenberg-Marquardt Method for Nonlinear Least Squares Curve-Fitting Problems.” *Department of Civil and Environmental Engineering, Duke University, Durham, USA*: 1–17. http://people.duke.edu/~hpgavin/ce281/lm.pdf (accessed on January 5, 2016).

Gillis, J. 2013. “Climate Change Seen Posing Risk to Food Supplies.” *The New York Times*. http://www.nytimes.com/2013/11/02/science/earth/science-panel-warns-of-risks-to-food-supply-from-climate-change.html?_r=0 (accessed on November 23, 2016).

Green, S. B. 1991. “How Many Subjects Does It Take To Do A Regression Analysis.”

Multivariate Behavioral Research 26(3): 499–510.

Härdle, W., and L. Simar. 2003. *Applied Multivariate Statistical Analysis*. New York, NY, USA: Springer Science+Business Media.

Hastie, T., R. Tibshirani, and J. Friedman. 2011. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer Science+Business Media.

Heyer, W. 2010. *Bericht zur Erstellung eines aktuellen Rankings der Anwendung von Pflanzenschutzmittelwirkstoffen in der Landwirtschaft Sachsen-Anhalts*. Halle an der Saale, Germany. http://www.lhw.sachsen-anhalt.de/fileadmin/Bibliothek/Politik_und_Verwaltung/Landesbetriebe/LHW/neu_PDF/5.0_GLD/Dokumente_GLD/Berichte_OW_Chemie/Bericht_Wirkstoffranking_2010_final.pdf (accessed on November 23, 2016).

Hicks, J. 1932. *The Theory of Wages*. London, UK: Macmillan.

Iglesias, A., S. Quiroga, and A. Diz. 2011. "Looking into the Future of Agriculture in a Changing Climate." *European Review of Agricultural Economics* 38(3): 427–447.

Iglesias, A., J. Schlickenrieder, D. Pereira, and A. Diz. 2011. "From the Farmer to Global Food Production: Use of Crop Models for Climate Change Assessment." In *Handbook on Climate Change and Agriculture*, eds. A. Dinar and R. O. Mendelsohn. Cheltenham, UK and Northampton, USA: Edward Elgar Publishing, 49–72.

IPCC. 2014. *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. eds. O. Edenhofer et al. Cambridge, United Kingdom and New York, USA: Cambridge University Press.

Kaufman, L., and P. Rousseeuw. 2005. *Finding Groups in Data: An Introduction to Cluster Analysis (Wiley Series in Probability and Statistics)*. Hoboken, USA: John Wiley & Sons.

LUF [Landwirtschaftliche Untersuchungs- und Forschungsanstalt]. 2001. *Versuchsbericht Landessortenversuche Wintergerste 2001*. Halle an der Saale, Germany.

Madsen, K., H. Bruun, and O. Tingleff. 2004. *Methods for Non-Linear Least Squares Problems*. Kongens Lyngby, Denmark: Informatics and Mathematical Modelling, Technical University of Denmark, DTU.

Maechler, M. et al. 2015. "Cluster: Finding Groups in Data: Cluster Analysis Extended. Partitioning Around Medoids." *R package version 2.0.3*. https://stat.ethz.ch/R-manual/R-devel/library/cluster/html/pam.html.

Massetti, E., and R. Mendelsohn. 2011. "The Impact of Climate Change on US Agriculture: A Repeated Cross-Sectional Ricardian Analysis." In *Handbook on Climate Change and Agriculture*, eds. A. Dinar and R. O. Mendelsohn. Cheltenham, UK and Northampton, USA: Edward Elgar Publishing, 141–166.

MELF [Ministerium für Ernährung, Landwirtschaft und Forsten des Landes Sachsen-Anhalt]. 1998. *Bericht zur Lage der Land-, Ernährungs- und Forstwirtschaft des Landes Sachsen-Anhalt 1998*. Magdeburg, Germany.

MELF [Ministerium für Ernährung, Landwirtschaft und Forsten des Landes Sachsen-Anhalt]. 1999. *Land-, Ernährungs- und Forstwirtschaft des Landes Sachsen-Anhalt 1999*. Magdeburg, Germany.

Mendelsohn, R., and A. Dinar. 2003. "Climate, Water, and Agriculture." *Land economics* 79(3): 328–341.

Mendelsohn, R., W. D. Nordhaus, and D. Shaw. 1994. "The Impact of Global Warming on Agriculture: A Ricardian Analysis." *The American Economic Review* 84(4): 753–771.

Mendelsohn, R. O., and A. Dinar. 2009. *Climate Change and Agriculture: An Economic Analysis of Global Impacts, Adaptation and Distributional Effects*. Cheltenham, UK and Northampton, USA: Edward Elgar Publishing.

Ministerium für Raumordnung, Landwirtschaft und Umwelt. 1996. *Agraratlas des Landes Sachsen-Anhalt. Die Agrarwirtschaft des Landes in Karten-Texten-Übersichten.* Magdeburg, Germany.

MLU [Ministerium für Landwirtschaft und Umwelt des Landes Sachsen-Anhalt]. 2002. *Bericht zur Lage der Land-, Ernährungs- und Forstwirtschaft des Landes Sachsen-Anhalt 2002*. Magdeburg, Germany.

MLU [Ministerium für Landwirtschaft und Umwelt des Landes Sachsen-Anhalt]. 2004. *Bericht zur Lage der Land-, Ernährungs- und Forstwirtschaft und Tierschutzbericht des Landes Sachsen-Anhalt 2004*. Magdeburg, Germany.

MLU [Ministerium für Landwirtschaft und Umwelt des Landes Sachsen-Anhalt]. 2005. *Bericht zur Lage der Land-, Ernährungs- und Forstwirtschaft und Tierschutzbericht des Landes Sachsen-Anhalt 2005*. Magdeburg, Germany.

MLU [Ministerium für Landwirtschaft und Umwelt des Landes Sachsen-Anhalt]. 2006. *Bericht zur Lage der Land-, Ernährungs- und Forstwirtschaft und Tierschutzbericht des Landes Sachsen-Anhalt 2006*. Magdeburg, Germany.

MLU [Ministerium für Landwirtschaft und Umwelt des Landes Sachsen-Anhalt]. 2007. *Bericht zur Lage der Land-, Ernährungs- und Forstwirtschaft und Tierschutzbericht des Landes Sachsen-Anhalt 2007*. Magdeburg, Germany.

MLU [Ministerium für Landwirtschaft und Umwelt des Landes Sachsen-Anhalt]. 2008. *Bericht zur Lage der Land-, Ernährungs- und Forstwirtschaft und Tierschutzbericht des Landes Sachsen-Anhalt 2008*. Magdeburg, Germany.

MRLU [Ministerium für Raumordnung, Landwirtschaft und Umwelt des Landes Sachsen-Anhalt]. 1997. *Land- und Forstwirtschaft in Zahlen 1997*. Magdeburg, Germany.

MRLU [Ministerium für Raumordnung, Landwirtschaft und Umwelt des Landes Sachsen-Anhalt]. 2000. *Bericht zur Lage der Land-, Ernährungs- und Forstwirtschaft des Landes Sachsen-Anhalt 2000*. Magdeburg, Germany.

MRLU [Ministerium für Raumordnung, Landwirtschaft und Umwelt des Landes Sachsen-Anhalt]. 2001. *Land-, Ernährungs- und Forstwirtschaft des Landes Sachsen-Anhalt 2001*. Magdeburg, Germany.

MRLU [Ministerium für Raumordnung, Landwirtschaft und Umwelt des Landes

Sachsen-Anhalt]. 2003. *Bericht zur Lage der Land-, Ernährungs- und Forstwirtschaft des Landes Sachsen-Anhalt 2003*. Magdeburg, Germany.

Nauges, C., C. J. O'Donnell, and J. Quiggin. 2011. "Uncertainty and Technical Efficiency in Finnish Agriculture: A State-Contingent Approach." *European Review of Agricultural Economics* 38(4): 449–467.

Nerlove, M. 1958. *The Dynamics of Supply: Estimation of the Farmer's Response to Price*. Balitimore, USA: Johns Hopkins University Press.

Parkin, M. 2008. "Adaptive Expectations." *The New Palgrave Dictionary of Economics, Volume 1*. London, UK: Palgrave Macmillan UK, 34–35.

Passel, S. V., E. Massetti, and R. Mendelsohn. 2014. *A Ricardian Analysis of the Impact of Climate Change on European Agriculture*, CESifo Working Paper, No. 4842.

Peck, D. E., and R. M. Adams. 2011. "Farm-Level Impacts of Climate Change: Alternative Approaches for Modeling Uncertainty." In *Handbook on Climate Change and Agriculture*, eds. A. Dinar and R. Mendelsohn. Cheltenham, UK and Northampton, USA: Edward Elgar Publishing, 89–111.

R Core Team and contributors. 2016a. "Stats: The R Stats Package. Hierarchical Clustering." *R version 3.4.0*. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/hclust.html.

R Core Team and contributors. 2016b. "Stats: The R Stats Package. Non-Linear Least Squares." *R version 3.4.0*. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/nls.html.

Ritz, C., and J. C. Streibig. 2008. *Nonlinear Regression with R*. New York, NY, USA: Springer Science+Business Media.

Scherer, J. 2016. "China Tells Trump Climate Change Is Not a Chinese Hoax." *The Washington Post*. https://www.washingtonpost.com/news/morning-mix/wp/2016/11/17/china-tells-trump-climate-change-is-not-a-chinese-hoax/ (accessed on November 23, 2016).

Schlenker, W., W. M. Hanemann, and A. C. Fisher. 2006. "The Impact of Global Warming on US Agriculture: An Econometric Analysis of Optimal Growing Conditions." *Review of Economics and Statistics* 88(1): 113–125.

Schlenker, W., W. M. Hanemann, and A.C. Fisher. 2007. "Water Availability, Degree Days, and the Potential Impact of Climate Change on Irrigated Agriculture in California." *Climatic Change* 81(1): 19–38.

Schlenker, W., and P. Roberts. 2008. "Nonlinear Effects of Weather on US Crop Yields, Implications for Climate Change, and Why These Effects Matter for Developing Countries." In *Agriculture and Development*, Washington DC, USA: The International Bank for Reconstruction and Development/World Bank, 193–202.

Schlenker, W, W. M. Hanemann, and A. C. Fisher. 2005. "Will US Agriculture Really Benefit from Global Warming? Accounting for Irrigation in the Hedonic Approach." *American Economic Review* 95(1): 395–406.

Shankar, S. 2012. "Production Economics in the Presence of Risk." *Australian Journal* *of Agricultural and Resource Economics* 56: 1–24.

Shephard, R. W. 1953. *Cost and Production Functions*. Princeton, USA: Princeton University Press.

Smith, P. et al. 2014. “Agriculture, Forestry and Other Land Use (AFOLU).” In *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds. O. Edenhofer et al. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 811–922.

United Nations. 2016. “Why Must We Act?” http://cop22.ma/en/#whatscop/post/162 (accessed on November 21, 2016).

United States Department of Agriculture. 2016. “What Are Crop Simulation Models?” https://www.ars.usda.gov/northeast-area/beltsville-md/beltsville-agricultural-research-center/crop-systems-and-global-change-laboratory/docs/what-are-crop-simulation-models/ (accessed on October 27, 2016).

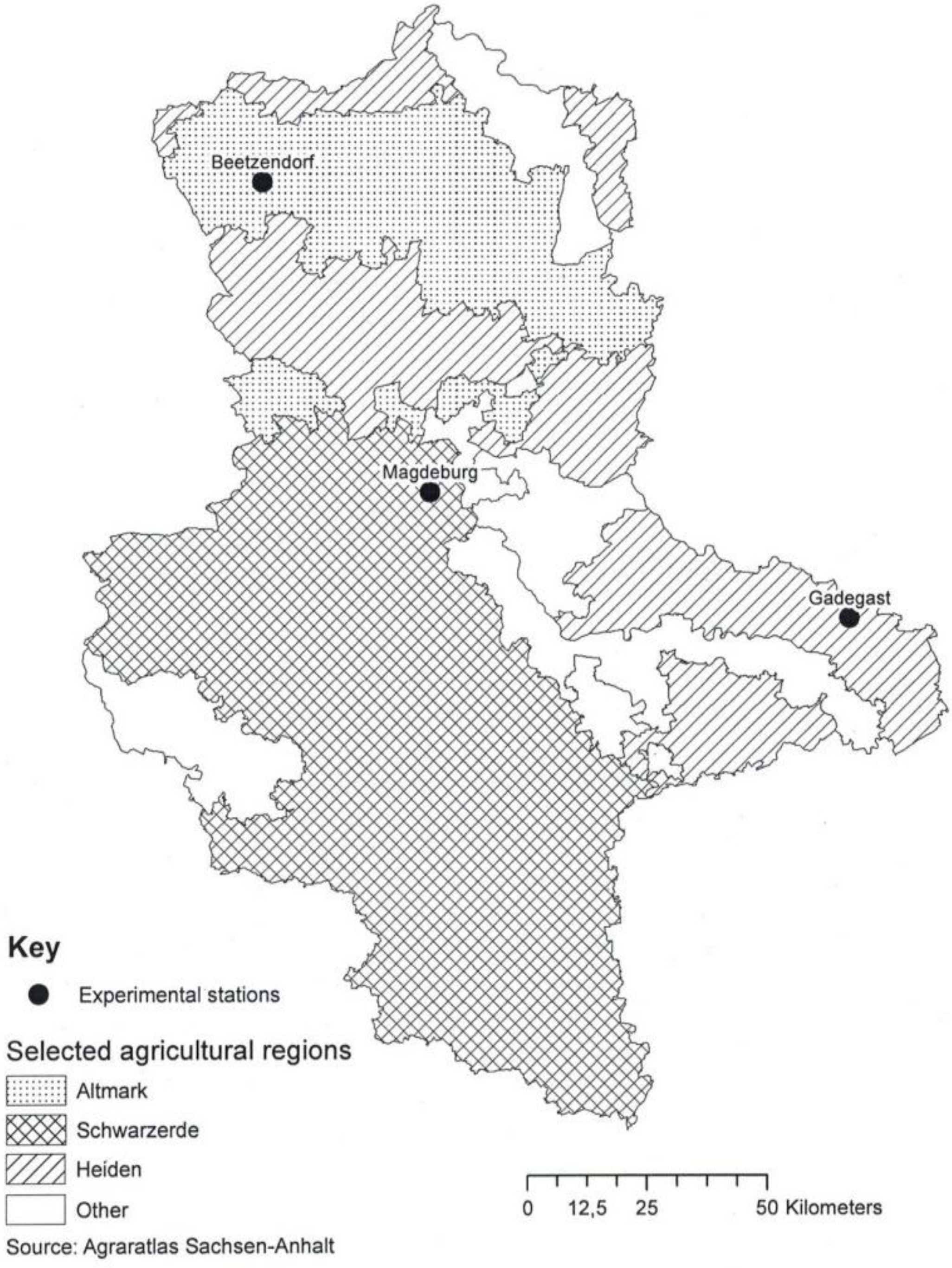
Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with S.* New York, USA: Springer.

Witten, I. H., E. Frank, and M. Hall. 2005. *Data Mining: Practical Machine Learning Tools and Techniques*. Burlington, USA: Morgan Kaufmann.

Wong, E. 2016. “Trump Has Called Climate Change a Chinese Hoax. Beijing Says It Is Anything But.” *The New York Times. http://www.nytimes.com/2016/11/19/world/asia/china-trump-climate-change.html* (accessed on November 23, 2016).

Zellner, A., J. Kmenta, and J. Drèze. 1966. “Specification and Estimation of Cobb-Douglas Production Function Models.” *Econometrica* 34(4): 784–795.

Appendix I: Map of Saxony-Anhalt

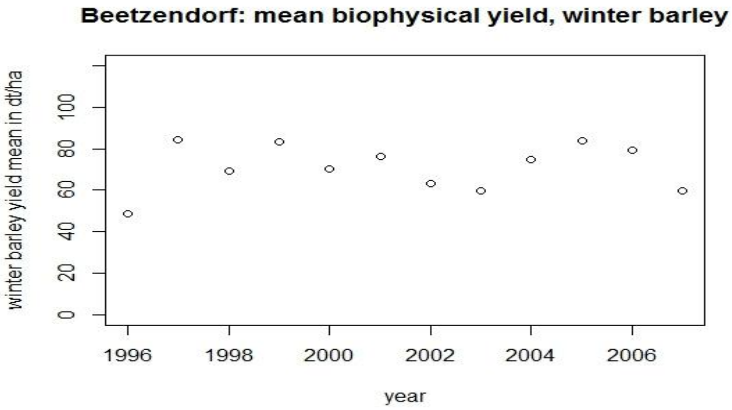


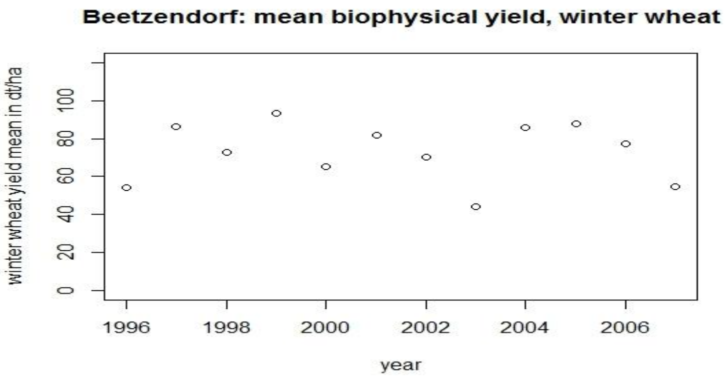
Map 1 Map of Saxony-Anhalt

Note: Location of the three agricultural regions and experimental stations in Saxony-Anhalt.

Source: Ministerium für Raumordnung, Landwirtschaft und Umwelt (1996)

Appendix II: Biophysical yields and ratios





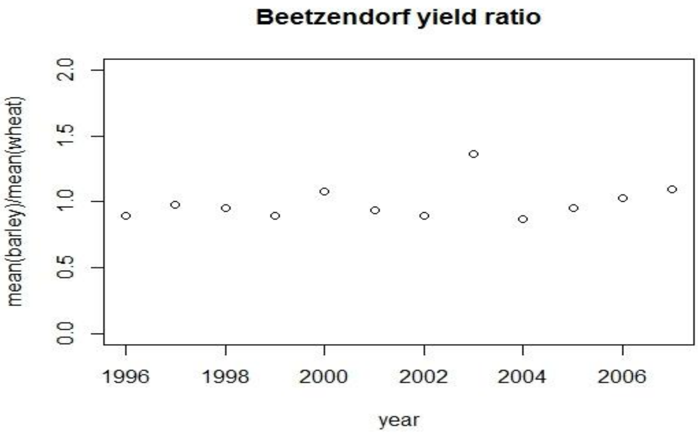


Figure 11 Biophysical crop yields at experimental station Beetzendorf.

Note: Observations on mean yields of winter barley, mean yields of winter wheat and the crop yield ratio between the years 1996 and 2007.

Source: own Illustration.

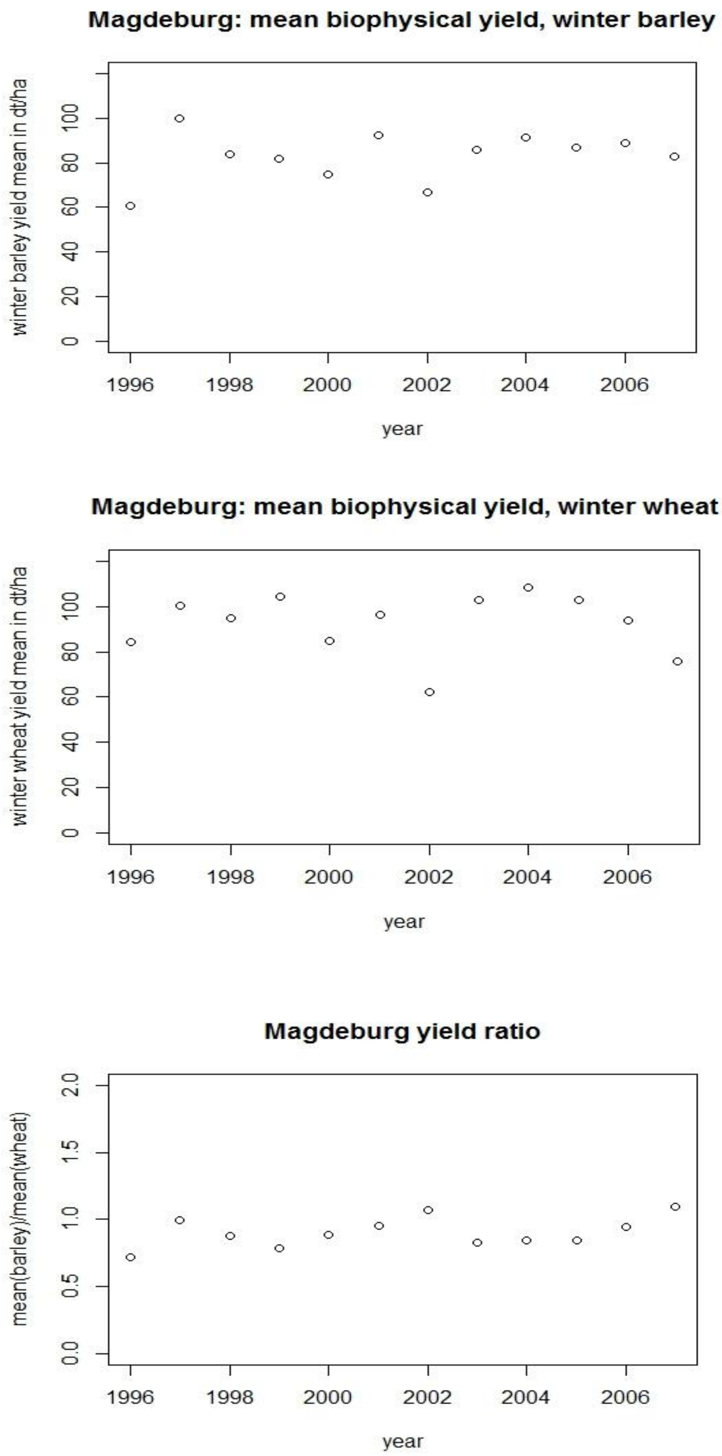


Figure 12 Biophysical crop yields at experimental station Magdeburg.

Note: Observations on mean yields of winter barley, mean yields of winter wheat and the crop yield ratio between the years 1996 and 2007.

Source: own Illustration.

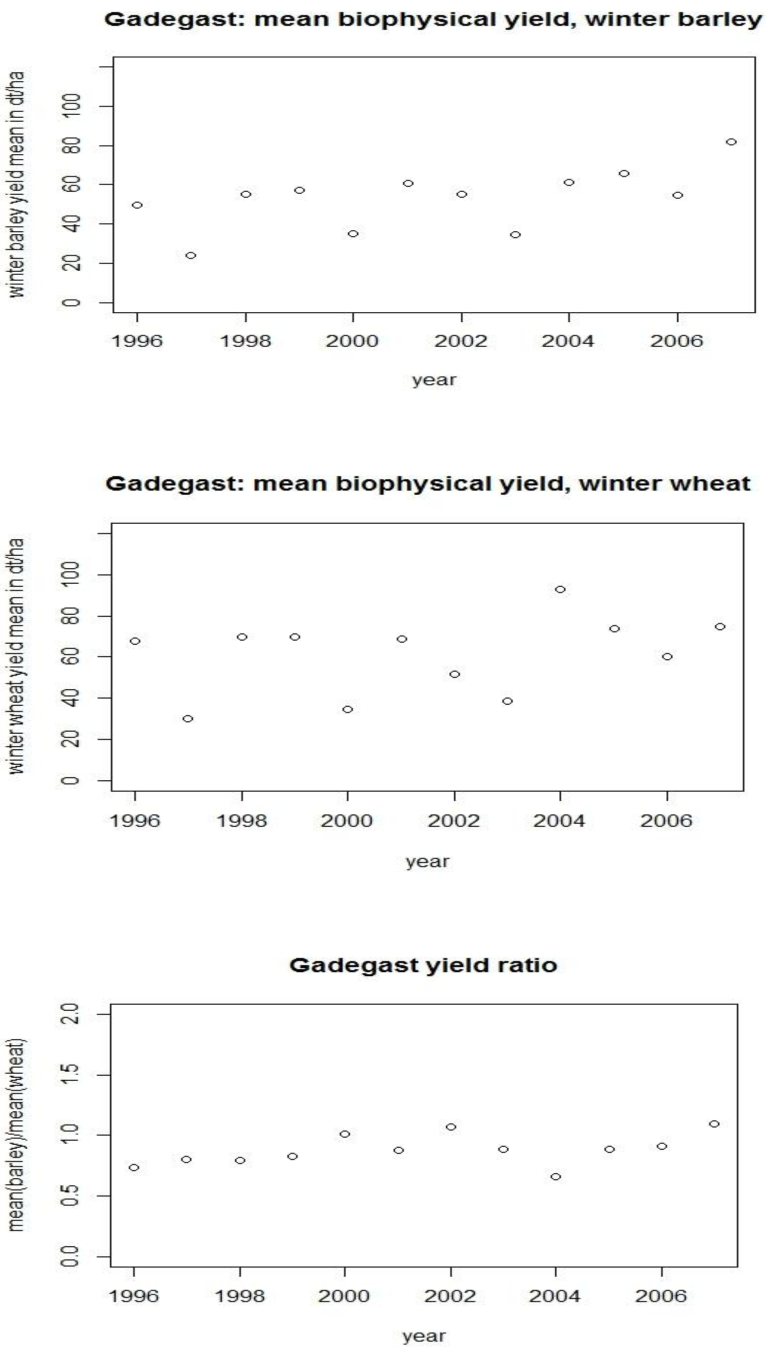
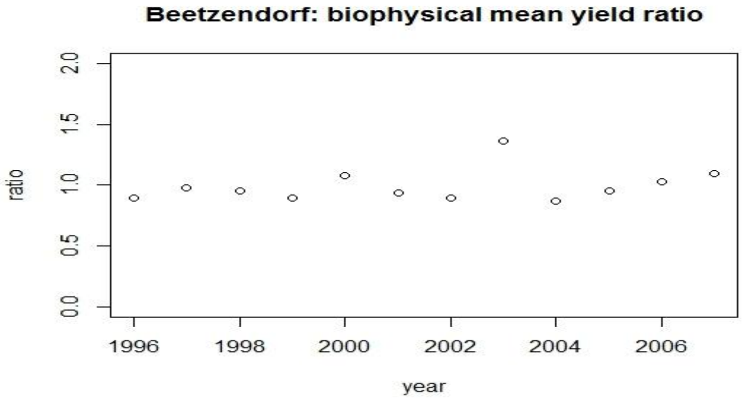


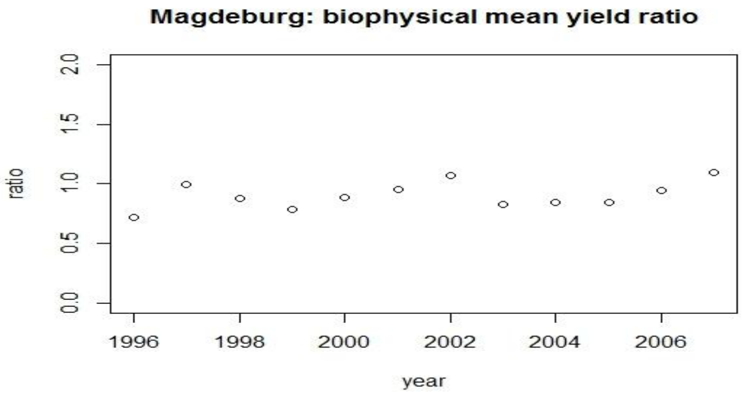
Figure 13 Biophysical crop yields at experimental station Gadegast.

Note: Observations on mean yields of winter barley, mean yields of winter wheat and the crop yield ratio between the years 1996 and 2007.

Source: own Illustration.

Appendix III: Biophysical mean yield ratios





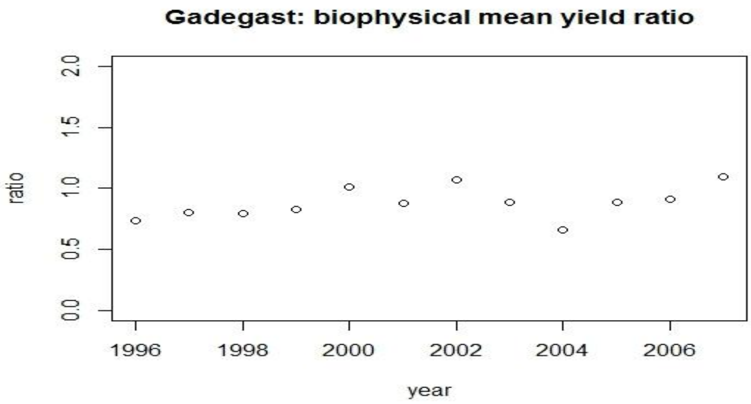
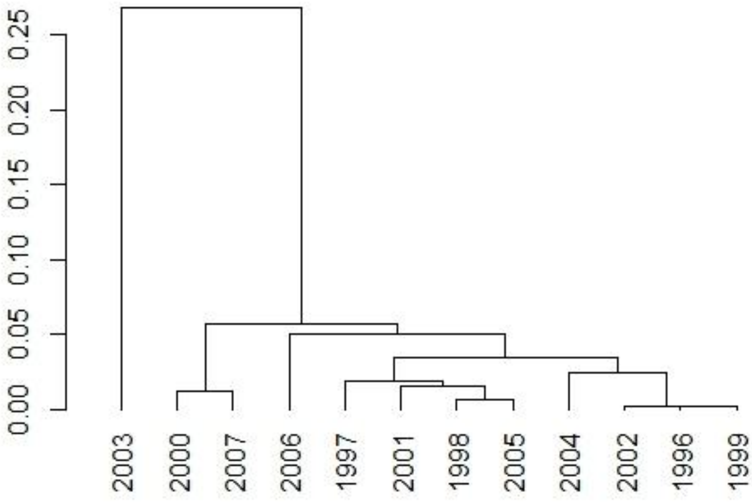


Figure 14 Biophysical mean yield ratios.

Note: The mean yield ratios at Beetzendorf, Magdeburg and Gadegast for the years between 1996 and 2007, displayed in the range of values between 0 and 2.

Source: own Illustration.

Appendix IV: Dendrograms



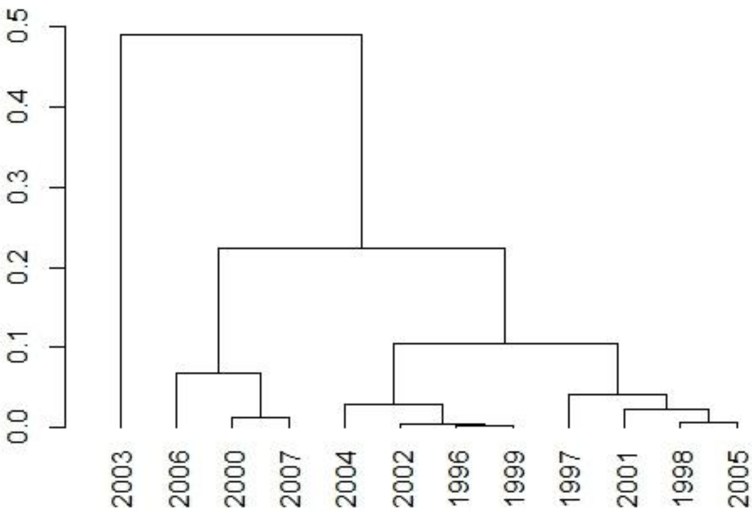
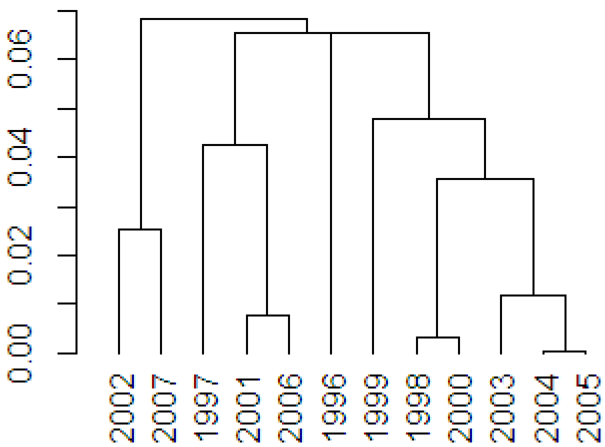


Figure 15 Hierarchical clustering results for Beetzendorf.

Note: Single linkage dendrogram above (consistent with two clusters), complete linkage dendrogram below (consistent with two clusters).

Source: own illustration.



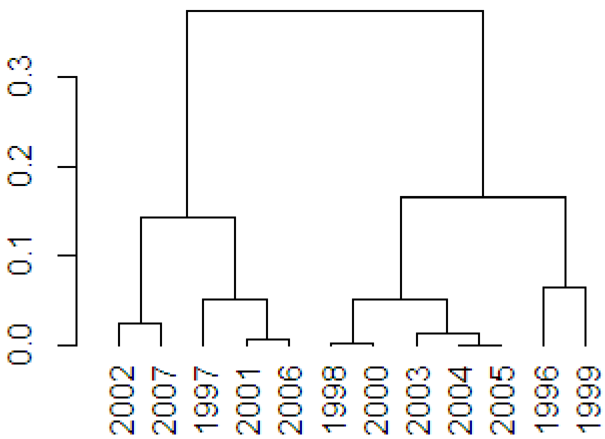
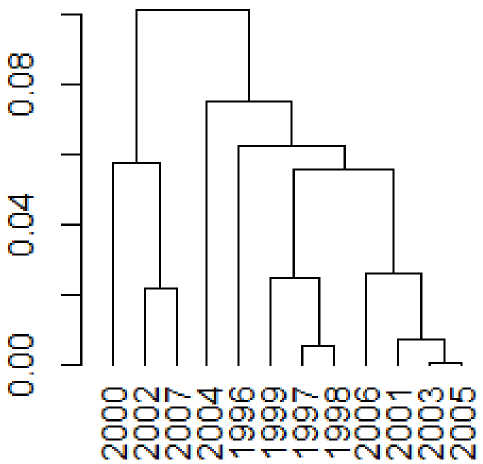


Figure 16 Hierarchical clustering results for Magdeburg.

Note: Single linkage dendrogram above (consistent with three clusters and an outlier observation 1996), complete linkage dendrogram below (consistent with two or four clusters).

Source: own illustration.



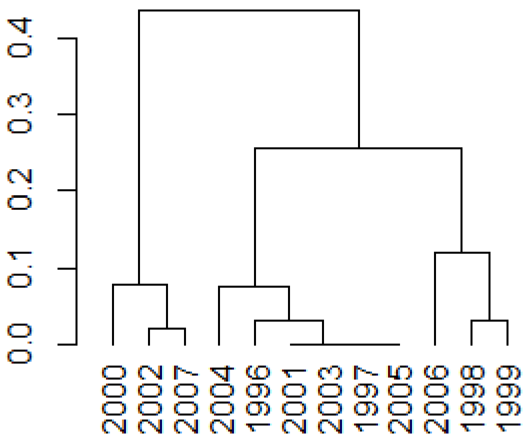


Figure 17 Hierarchical clustering results for Gadegast.

Note: Single linkage dendrogram above (consistent with two or three clusters), complete linkage dendrogram below (consistent with two or three clusters).

Source: own illustration.