AgMIP The Agricultural Model Intercomparison and Improvement Project

Protocols for AgMIP Regional Integrated Assessments

VERSION 6.1

AgMIP gratefully acknowledges the UK Department for International Development's UKaid for major project funding, and the US Department of Agriculture for its sustained support. AgMIP additionally thanks the CGIAR Research Program on Climate Change, Agriculture and Food Security and the United States Agency for International Development (USAID) for project support.

AgMIP additionally acknowledges support from The International Life Sciences Institute/The Center for Integrated Modeling of Sustainable Agriculture & Nutrition Security, Monsanto, University of Florida, The National Oceanic and Atmospheric Administration, iPlant, Bill and Melinda Gates Foundation, Columbia University/Earth Institute, and The National Aeronautics and Space Administration.

AgMIP extends a special thank-you to Columbia University and The Earth Institute for hosting AgMIP Coordination at its Center for Climate System Research. AgMIP appreciates significant in-kind contributions from many organizations and institutions (see also inside back cover).

Suggested Citation:

Rosenzweig, C., J.W. Jones, J. Hatfield, J.M. Antle, A.C. Ruane, K.J. Boote, P. Thorburn, R.O. Valdivia, K. Descheemaeker, C.H. Porter, S. Janssen, W. Bartels, A. Sulivan, and C.Z. Mutter. (2016). *Protocols for AgMIP Regional Integrated Assessments Version 6.1.*

Protocols for AgMIP Regional Integrated Assessments

Version 6.1

AgMIP International Leaders

Cynthia Rosenzweig, Jim Jones, John Antle, and Jerry Hatfield

AgMIP Team Leaders

Climate: **Alex Ruane and Sonali McDermid**

Crops: **Ken Boote and Peter Thorburn**

Livestock: Katrien Descheemaeker

Regional Economics: **John Antle and Roberto Valdivia**

Information Technologies: **Cheryl Porter and Sander Janssen**

Stakeholders: **Wendy-Lin Bartels and Amy Sullivan**

AgMIP Science Coordinator

Alex Ruane (alexander.c.ruane@nasa.gov)

AgMIP International Coordinator

Carolyn Mutter (czm2001@columbia.edu)

November 22nd, 2016

Protocols for AgMIP Regional Integrated Assessments[1]

Table of Contents

[1] Note that previous versions of this document were entitled: "Guide for Regional Integrated Assessments: Handbook of Methods and Procedures"
Cover photos by Shari Lifson and Alex Ruane

AgMIP The Agricultural Model Intercomparison and Improvement Project

i. Introduction

The purpose of this handbook is to describe recommended methods for a trans-disciplinary, systems-based approach for regional-scale (local to national scale) integrated assessment of agricultural systems under future climate, bio-physical and socio-economic conditions. Readers who wish to learn about the overall process should read through the main sections, and others may want to go directly to the numbered sections below that provide a step-by-step description of the procedures. This handbook is written to guide a consistent set of integrated assessments that can be applied to any region globally. A list of the key characteristics of an AgMIP regional integrated assessment (RIA) is provided in the next section. These assessments are the regional manifestation of research activities described by AgMIP in its online protocols document (available at www.agmip.org). AgMIP Protocols were created to guide climate, crop and livestock modeling, economic modeling of farming systems, and information technology components of its projects.

Various regions of the world are now undertaking regional assessments following AgMIP protocols and integrated assessment procedures. This Handbook version also has a number of modifications to the methods and to our description of methods based on what was learned from the use and evaluation of prior versions of the Handbook. However, it is important to recognize that the procedures presented here were designed for the data available to the SSA and SA teams, for implementation of two crop models per integrated assessment region (at least DSSAT and APSIM), and for use of one socio-economic model (TOA-MD) in the integrated impact assessments. Going forward, we recommend the use of multiple crop, livestock, and economic models when available, based in large part to lessons learned in the various crop model intercomparisons (e.g., Rosenzweig et al., 2013; Asseng et al. 2013, 2014; Bassu et al., 2014; Li et al., 2014), global economic model intercomparisions (Nelson et al., 2013; von Lampe et al., 2014), and global gridded crop model intercomparisons (Rosenzweig et al., 2014; Elliott et al., 2015). We envision that specific choices of multiple models may vary among regions, but that a core set of models should be used such that results can be aggregated and compared across all regions.

AgMIP regional integrated assessments require close coordination among economic, climate, crop modeling and IT team members within each regional research team (RRT). Many teams are also now integrating livestock modeling into their assessments, whose linkages are described for the first time in this version. Assessments begin with regional teams working with stakeholders to define what outcomes are to be evaluated and then developing details of the specific production systems that need to be quantified. Each RRT should focus on impacts related to, at minimum, food production, income, and poverty in their regions; emphasizing important food crops and quantifying relevant uncertainties. Where appropriate, livestock components of production systems should be included. Then a plan of work should be developed by teams that will include AgMIP-recommended methods and procedures to accomplish integrated assessments and desired compatibility of outputs across regions.

This handbook was written such that it represents a minimum approach that can be expanded upon in regions where available data and resources allow. The methods and core approach used by all interdisciplinary research teams need to be consistent in order to enable meta-analyses and large-scale studies. Particular care must therefore be taken in introducing new methods and models that could potentially limit the ability of results to be compared beyond the immediate region.

This handbook is a living document that will continue to evolve and be improved through input from the regions as they apply it to their own situations and gain experience in the methods that are aimed at helping to unify methods and outputs.

ii. Key Attributes of an AgMIP Regional Integrated Assessment

- Designed with input from stakeholders, policymakers, and/or other end-users
- Oriented upon production-systems-based approach (rather than specific crops or fields) potentially including multiple crops, livestock, aquaculture, and other sources of income.
- Transdisciplinary in its linking of climate, biophysical, and socio-economic conditions and responses.
- Flexible in that its framework allows for the testing of adaptations and alternative models and methods within a given region.
- Addresses core questions of climate impact on current and future production systems (detailed in the next section)
- Allows evaluation of stakeholder-consulted adaptations in production systems under current as well as future climate.
- Calibrated on current production system using available data with documentation sufficient to enable replication of results.
- Examines the impact of both mean climate changes and potential interactions with climate variability
- Presents results in a probabilistic manner with accounting of major uncertainties.
- Utilizes consistent terminology across disciplines and among various AgMIP assessments and initiatives.
- Uploads results to an online AgMIP database for archival and cross-regional analyses with full attribution of data providers and intellectual contributions.
- Publishes findings in peer-reviewed journals and disseminates information to stakeholders.

iii. Core Climate Impact Questions

AgMIP has identified four core research questions[2] that motivate research activities for regional integrated assessments (**Figure 1**):

1) What is the sensitivity of current agricultural production systems to climate change? This question addresses the isolated impacts of climate changes assuming that the production system does not change from its current state.

[2] Note that previous versions of this handbook and Antle et al., 2015, defined only three core questions. Core question #2, as presented here, was added, resulting in the renumbering of core question #3 (previously #2) and core question #4 (previously #3).

2) What are the benefits of adaptation in current agricultural systems? This question addresses the benefit (e.g., economic and food security resilience) of potential adaptation options to current agricultural systems given current climate. Results also form a basis for comparison with Core Question 4 below, as the proposed adaptations may have a higher benefit when the climate changes.
3) What is the impact of climate change on future agricultural production systems? This question evaluates the isolated role of climate impacts on the future production system, which will differ from the current production system due to development in the agricultural sector not directly motivated by climate changes.
4) What are the benefits of climate change adaptations? This question analyzes the benefit of potential adaptation options in the production system of the future, which may offset or capitalize on climate vulnerabilities identified in Core Question 3 above.

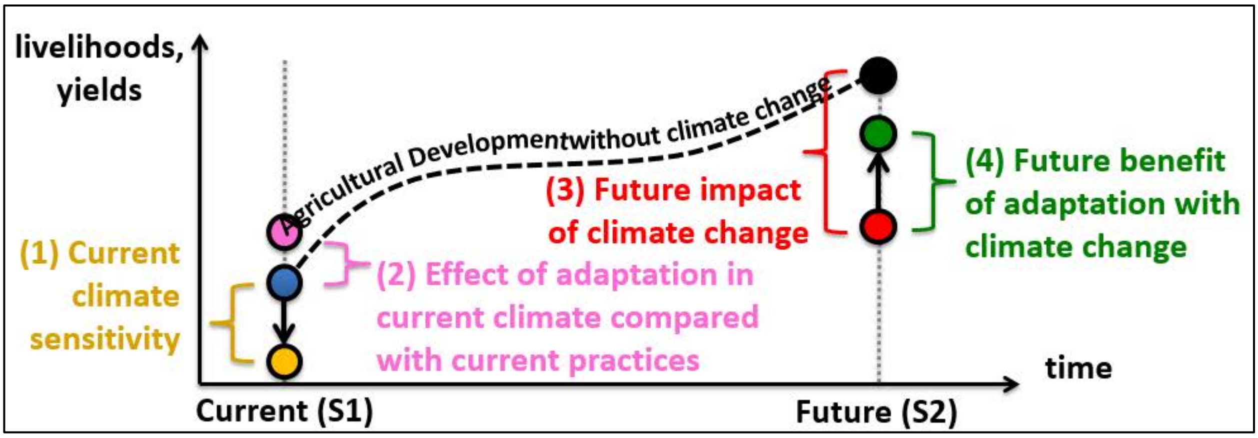


Figure 1. Overview of core climate impact questions and the production system states that will be simulated. The current production system is represented by the blue dot, while the future production of the system is represented in three ways: assuming that there is no climate change (black), assuming that there is climate change and no adaptation (red), and assuming that there is climate change and adaptation (green). The dashed line represents the evolution of the production system in response to development in the agricultural sector that is not directly motivated by climate change. To understand the sensitivity of the current production system to anticipated changes, production in the current period is also estimated using proposed adaptation strategies under present climate (pink) or responding to an instantaneous climate change (orange). Six combinations of simulations, each represented by a colored dot (see Table 1), are needed to address the four core questions (see Table 2).

As each question is designed to allow a comparison between two different production system states, **Table 1** describes the key climate, crop, livestock, and economic modeling components that will describe and compare these states, and **Table 2** describes the comparisons corresponding to each core question.

Table 1. Overview of crop/livestock model simulations needed to represent the systems of interest for the four core questions, along with the climate, agricultural pathway, and adaptation that characterizes each simulation. Note that the agricultural system (colored dot) for each simulation corresponds to the diagram of core questions in figure 1, and for this table the future (climate change) period is assumed to be the RCP8.5 Mid-Century (2040-2069).

System	Crop / Livestock Simulation	Driving Climate (# scenarios passed to Econ)	Crop/Livestock Management configuration (and corresponding RAP)	Major Climate Adaptation
● (blue)	CM1	Current (1)	Current	None
● (yellow)	CM2	CC (5)	Current	None
● (pink)	CM3	Current (1)	Current	Adaptation
● (black)	CM4	Current (1)	Future (RAP1)	None
● (red)	CM5	CC (5)	Future (RAP1)	None
● (green)	CM6	CC (5)	Future (RAP1)	Adaptation

Table 2. Overview of economic model simulations corresponding to the four core questions for AgMIP RIA. Each economic simulation set contrasts two systems (represented by colored dots as in Figure 1 and Table 1) to evaluate the economic impacts of potential changes in the agricultural system.

Core Question	Name	RAP	Climate Adaptation	Yield/Productivity Change Ratio from Ag Model Runs
#1	Climate Impact in Current World	No	No	● (yellow) CM2/CM1 ● (blue)
#2	Climate Adaptation in Current World	No	Yes	● (pink) CM3/CM1 ● (blue)
#3	Climate Impact in Future World	Yes	No	● (red) CM5/CM4 ● (black)
#4	Climate Adaptation in Future World	Yes	Yes	● (green) CM6/CM5 ● (red)

iv. Key Regional Team Outputs

A number of outputs are anticipated from the sum of RRT activities described in this Handbook. This list of anticipated activities is intended to be used for RRT planning, and thus specific outputs and methods are provided in the material that follows. In addition, however, there are several overarching outputs that should be targeted by each RRT. These overarching outputs are summarized below, along with questions that help motivate the construction of these outputs.

a. ***A network of sites where multiple crop and livestock models have been calibrated using locally representative management, soils, cultivars, animal breeds, and climate to simulate food production regions that are important for regional food security, with analysis of calibration uncertainties.*** Key questions include:
   - Which important farm systems, crops, and agricultural sub-regions are to be targeted for simulating regional food security?

- What data are available for calibration of crop and livestock models and to estimate parameters for the economic model?
- How do crops respond to applied levels of fertilizer nitrogen?
- How do livestock respond to variability in the feed composition resulting from climate variability?
- What adaptation measures should be analyzed in the study?

b. ***A set of Representative Agricultural Pathways (RAPs) for each region for use in analyses of regional climate impacts and adaptation.*** Key questions include:

- What RAP narrative(s) best describe the future world that the analyst wants to characterize?
- What output variables from global economic models and analyses are key drivers of agricultural trends in the region (e.g., commodity prices, population growth and GDP growth from Shared Socio-economic Pathways, and global representative agricultural pathways)?
- What key regional variables are likely to be affected by the higher level drivers (policy, socioeconomic, and technology)?
- What quantitative trends in each of the variables (including fertilizer, improved cultivars and breeds, improved management, forage availability, farm size, etc.) are needed to parameterize agricultural models (crop, livestock, and economic) for the regional integrated assessment of future production systems?

c. ***Characterization of historical agro-climate, sensitivity to climate shifts, and climate change scenarios downscaled for use at the regional scale.*** Key questions include:

- How is climate currently changing in the region?
- What are the most important climate factors that impact a given farm or region?
- What types of climate changes are projected to impact the region in the future and how certain are these projections?
- What are the vulnerabilities of crops and livestock to current and future climate variability, and what are the sensitivities of the multiple crop models to climate changes in temperature, CO_2, and rainfall?
- Where are agro-climatic impacts likely to be most acute?

d. Assessment of economic impacts for a subset of agricultural regions under future climate change, adaptation and socio-economic scenarios. Key questions include:

- How will climate change affect the distribution of production, income, and poverty in the farming systems of a given region if adaptations do not occur?
- What are the projected adoption rates of climate-adapted systems? How will various adaptations affect the impacts of climate change? How will alternative future socio-economic scenarios affect the impacts of climate change?
- How do uncertainties in key economic parameters affect the projected climate change impacts?

e. ***Adaptation packages including agronomic, animal husbandry, economic, and policy adaptations that improve outcomes under current and future conditions.***
   Key questions include:
   - What farm-level management adaptations would be beneficial under current and future climate conditions?
   - What changes to the production system would increase resilience under present climate variability and future climate challenges?
   - How can these adaptations be represented consistently in crop, livestock, and economic models?

f. ***Documentation for communication to the scientific community and to stakeholders.*** This includes linkages into the AgMIP Impacts Explorer, web sites, databases, scientific publications, and reports that have been communicated to stakeholders. Key questions include:
   - What are the pressing questions for climate change planning in the region?
   - What agricultural sector decisions are being made that may be affected by climate change?
   - How can researchers best communicate results to support decision-making among a variety of stakeholders?

v. AgMIP Standardized Formats and Tools
To ensure consistency in the archival and translation of data and results from AgMIP integrated assessment regions, several resources, tools, and standardized data formats have been created that will be referenced in the activities below. These standardized formats also ensure compatibility with stand-alone and web-based tools that will facilitate the execution of research activities and the dissemination of integrated assessment results.

- **.AgMIP climate data format** – Standardized format for climate series at a single location, featuring daily climate data and variables needed for crop modeling.
- **Guide for Running AgMIP Climate Scenario Generation Tools with R –** This "AgMIP Climate Scenarios Guidebook" describes how to access the data and suite of scripts required to produce AgMIP climate scenarios using the AgMIP methodologies, using .AgMIP-formatted climate data for both inputs and outputs.
- **C3MP Protocols –** The Coordinated Climate-Crop Modeling Project (C3MP) has established a set of standardized sensitivity tests of crop and livestock models response to carbon dioxide, temperature, and water changes. Protocols may be downloaded at www.agmip.org/c3mp-downloads.
- **AgMIP Crop Experiment (ACE) database** – contains site-based crop experiment (e.g., calibration data) and farm survey data using the AgMIP harmonized format. Crop modeling data can be translated from raw format to ACE and from ACE to crop model-ready formats using the QuadUI desktop utility. These data are archived in the ACE online database through the AgMIP Data Interchange (data.agmip.org).
- **Data Overlay for Multi-model Export (DOME)** refers to field overlays and seasonal strategies. Field overlays contain information related to field conditions which were not measured at the ACE sites, but are needed for crop modeling exercises. These data are estimated based on the best agronomic knowledge of cultural practices in the region.

Seasonal Strategy DOMEs contain baseline and future management and climate inputs which are used to modify existing site data for analysis of hypothetical scenarios. Each DOME dataset will be linked to one or more ACE datasets. These data are archived in the DOME online database through the AgMIP Data Interchange (data.agmip.org).

- **AgMIP Crop Model Output (ACMO)** data are the harmonized outputs from AgMIP ensemble crop model simulations. ACMO data are linked to both ACE and DOME data. These data are archived in the ACMO online database through the AgMIP Data Interchange (data.agmip.org).
- **User's Guide to Crop Model Simulations for Regional Integrated Assessments** – contains complete guidelines and crop modeling advice relative to entering experimental and farm survey yield data into the ACE template, use of DOME files to input standard assumptions, creation of model-ready files, running of the multiple crop models, and storage of output into ACMO files.
- **User's Guide to Livestock Model Simulations for Regional Integrated Assessments** – contains guidelines and advice related to creating livestock model input files, running the livestock model LivSim and consulting and exporting model output for further analysis.
- **Economic model input and output archives** – This repository will store input and output data for the TOA-MD model. Each TOA-MD file will be associated with one or more ACMO datasets via the metadata.
- **DevRAP** – Provides a structure to guide the process to develop Representative Agricultural Pathways (RAPs), to record and document the information systematically, and to translate RAPs into model-specific scenarios. The DevRAP v1.0 provides a structured format for the parameters needed to run the TOA-MD model as well as crop models.
- **AgMIP ftp site** – An ftp site has been established to archive data for review or processing prior to upload to the AgMIP Data Interchange databases. This ftp site can be accessed at ftp://data.agmip.org using the usernames and passwords assigned to each team.
- **Data Journal** – will be used to publish and permanently archive datasets which are complete and form the basis of journal articles, web visualizations, or other references. These published datasets will be assigned a DOI and can be cited with credit given to data authors, as in any other published work.
- **FACE-IT** –An online workflow system is under development which allows the intensive computations required for the RIA system to be performed using chains of applications, deployed on a cloud-server. This system, FACE-IT (Framework to Advance Climate, Economic and Impact Investigations with Information Technology) provides an alternative to using the AgMIP desktop utilities for data translation and allows simulations using DSSAT and APSIM for complex workflows, including multiple climate scenarios, sensitivity analyses, and adaptation scenarios. Procedures for using this system are not covered herein, but interested users are encouraged to learn more at www.learnfaceit.org.

vi. Guidelines for Activities for AgMIP Regional Research Teams

A list of characteristic activities for AgMIP Regional Projects includes 14 categories of activities along with methods that integrate across climate, crop modeling, livestock modeling, economics, and IT teams. These are listed in **Table 3** and presented in the sections below. **Figure 2** shows a schematic of the overall components of the integrated assessment process. Because of the importance of close collaboration among different disciplines (climate, crop, economic and livestock if included in the systems), regional teams may want to define a subset of the overall analysis to make sure that all team members learn how to best interact with other team members to achieve the overall results. Here, we present the overall activities needed to perform the entire integrated assessment. Full documentation of steps and procedures are provided in Appendix 1.

Table 3. Overview of tasks necessary to complete and disseminate regional integrated assessment. The section describing protocols for each task is also identified, as well as the disciplinary team primarily responsible for execution of each task.

Section	Task	Team Responsible
1	Scoping of production systems and developing/refining research work plan for regional integrated assessment	All, led by Stakeholder Assessment Team
2	Develop Representative Agricultural Pathways (RAPs) for use in regional analysis of climate impact and adaptation	All, led by Economic Team
3	Assemble existing data from experiments and calibrate crop models for regionally-relevant cultivars	Crop Modeling Team
4	Assemble existing data and calibrate livestock models for regionally-relevant livestock breeds	Livestock Modeling Team
5	Assemble and quality-control current climate series	Climate Scenarios Team
6	Assemble farm-survey yield data and simulate crop models for analysis of yield variations	Crop Modeling Team
7	Analyze Carbon-Temperature-Water-Nitrogen (CTWN) responses	Crop/Livestock Modeling and Climate Scenarios Teams
8	Assemble farm-survey livestock data and compare with livestock model outputs for analysis of livestock productivity variations	Livestock Modeling Team
9	Assemble economic data for regional economic analysis and develop skills for using the regional economic model	Economic Team
10	Create downscaled climate scenarios	Climate Scenarios Team
11	Conduct multiple crop/livestock model simulations	Crop Modeling Team
12	Analyze regional economic impacts of climate change without and with adaptation using the regional economic model	Economic Team
13	Archive data and analyses results for integrated assessments	All, led by IT
14	Disseminate integrated assessment results	All, led by Stakeholder

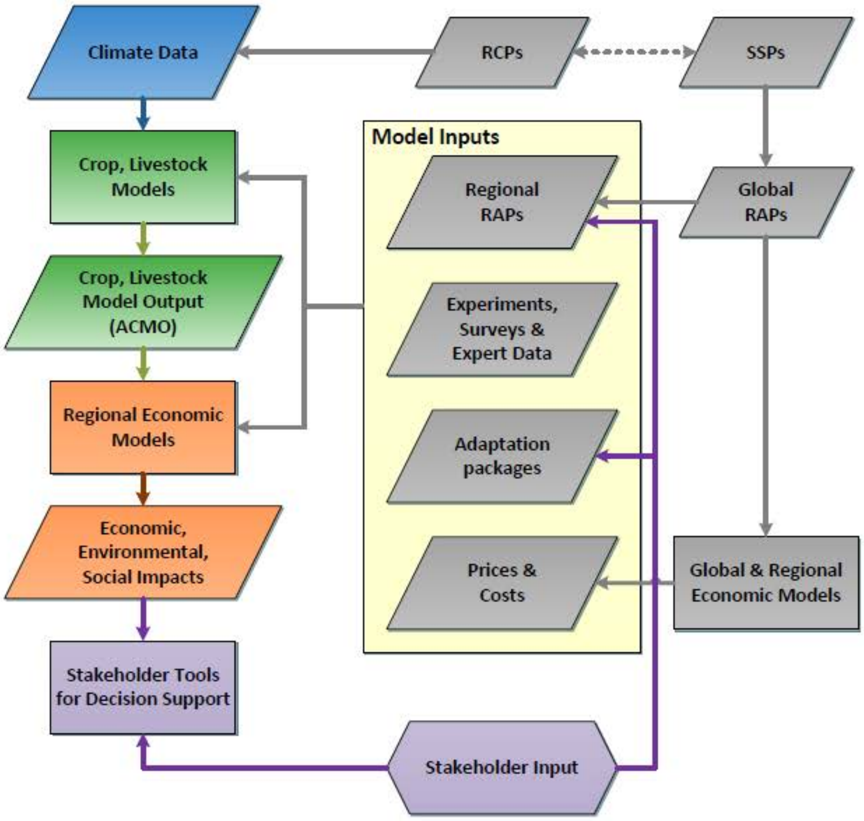


Figure 2. AgMIP Regional IA Framework: Parallel development of system design, data and modeling to couple crop & livestock models with TOA-MD, including input from and outputs to stakeholders.

1. Scoping of production systems and developing/refining research work plan for regional integrated assessment.

The overall outputs from this set of activities is a report describing the region, crops and livestock system components selected for explicit modeling, characteristics of the broader agricultural systems, the availability of data (climate, crop, soil, livestock, and socioeconomic), and stakeholder interactions and inputs. Suggested components of this phase of the projects are as follows.

a. **Review key project objectives, develop or refine research questions**, determine relevant stakeholders and policymakers, and assign team roles.

b. **Define key production systems** to be studied in consultation with stakeholders, identify how they influence food security in the region, and identify current questions and ongoing considerations for long-term planning and investment. Select crops and livestock that will be explicitly modeled in the study, other important components of the production system that must also be represented (e.g., rangelands for livestock grazing), and important sub regions that will be modeled in the study (**Figure 3**).

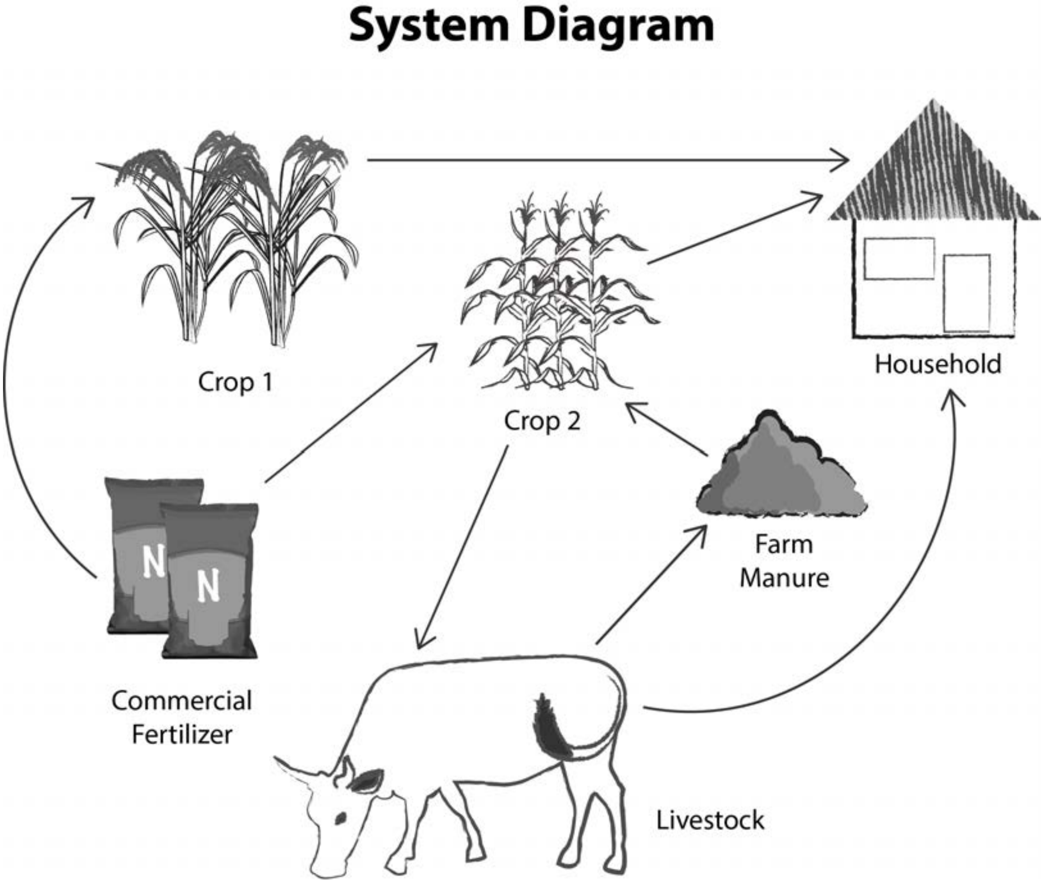


Figure 3. Example diagram describing the major elements and interactions of a production system.

c. **Select (multiple) crop models that will be used**, keeping in mind that the aim is to use at least the DSSAT and APSIM cropping system models across all regions. Assess the level of experience among team members with the selected models and identify additional capacity building needs.

d. **Select (multiple) livestock models that will be used,** with the aim to use at least LivSim across the regions. Potentially a rangeland production model could be included as well (e.g. SAVANNA). Assess the level of experience among team members with the selected models and identify additional capacity building needs

e. **Build capacity in the team of economists to use the Tradeoff Analysis Model for Multi-Dimensional Impact Assessment Model (TOA-MD)**, the economic model that has been used in prior regional efforts, or equivalent regional economic model(s). Identify project team members who will work with the regional economic model. Evaluate regional economic model capacity-building needs and team members in the RRTs who would participate in trainings.

f. **Produce a work plan that includes responsible persons, activities, time lines, and maps** of regions showing administrative boundaries, regions that will be studied, and points showing where climate and crop data are available. The report will include specifics of the information obtained in the above points, including the plan for stakeholder engagement.

2. Develop Representative Agricultural Pathways (RAPs) for use in regional analysis of climate impact and adaptation.

RAPs provide an overall narrative description of a plausible future development pathway, and also contain key variables with qualitative storylines and quantitative trends, consistent with higher-level pathways (e.g. SSPs, global RAPs developed by the AgMIP Global Modeling Group), see **Box 1**, **Box 2**, and **Figure 4**. Prices, policy and productivity trends should be consistent with the higher-level RAPs or scenarios that are available (SSPs, global RAPs, CCAFS regional scenarios). RAPs are translated into one or more scenarios (parameterizations) for the TOA-MD model and crop models. These RAPs represent a set of technology and management changes that will occur over time independent of climate change. These scenarios, developed for specific RAPs, will typically include changes in the types of crops or livestock produced and the way they are managed (e.g., use of fertilizers and improved crop cultivars).

Procedures for RAPs development are based on a step-wise process as shown in **Box 1**, with input from all components (climate, crop, livestock, economic) of the AgMIP Regional Team. Outside experts may need to be consulted if there is an important area of expertise not represented within the team. Stakeholders should also be incorporated into RAPs developed, as described below.

Box 1. Overview of Step-wise Process for RAPs Development

1. A multi-disciplinary team of scientists and other experts is established.
   - Team members need to have knowledge of the agricultural systems and regions to be covered
2. The team reviews general goals and define the time period for analysis and selected higher-level pathways (Shared Socio-economic Pathways, Global RAPs) to follow the nested approach (**Figure 4**)
3. Main drivers from higher level pathways are identified (and quantified if possible, e.g. outputs from global models)
4. Based on drivers and specific agricultural systems, a draft of a title and a short narrative of a RAP is constructed
5. Based on the draft narrative, the team identifies key parameters that will likely be affected by driving forces
6. The team draft storylines for each one of the parameters (see **Figure 5**)
7. The team checks for consistency within the RAP components and with higher level pathways and models' outputs
8. Based on consistency check, agreement and confidence levels among team participants, steps 4 -7 are repeated until an acceptable draft of consistent storylines and levels of agreement and confidence are achieved.
9. The team identifies parameters that will need additional revision (expert opinion, modeled data, etc.) or that will likely be subject to sensitivity analysis.
10. The team elaborate full RAP narrative
11. The RAP narrative is documented and distributed to other experts, scientists and key stakeholders for comments.
    - A workshop is organized to discuss the RAP narratives with key stakeholders and obtain their feedback.
12. The final RAPs are distributed to the modeling teams for parameters quantification (for crop and economic models) and scenario development

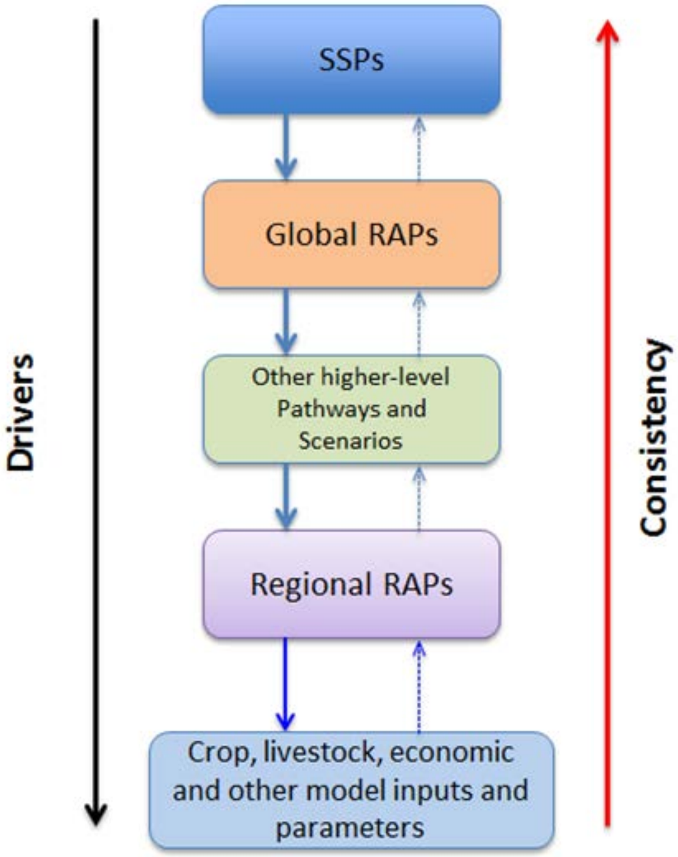


Figure 4. Developing RAPs and Scenarios: Use of a nested approach to assure consistency

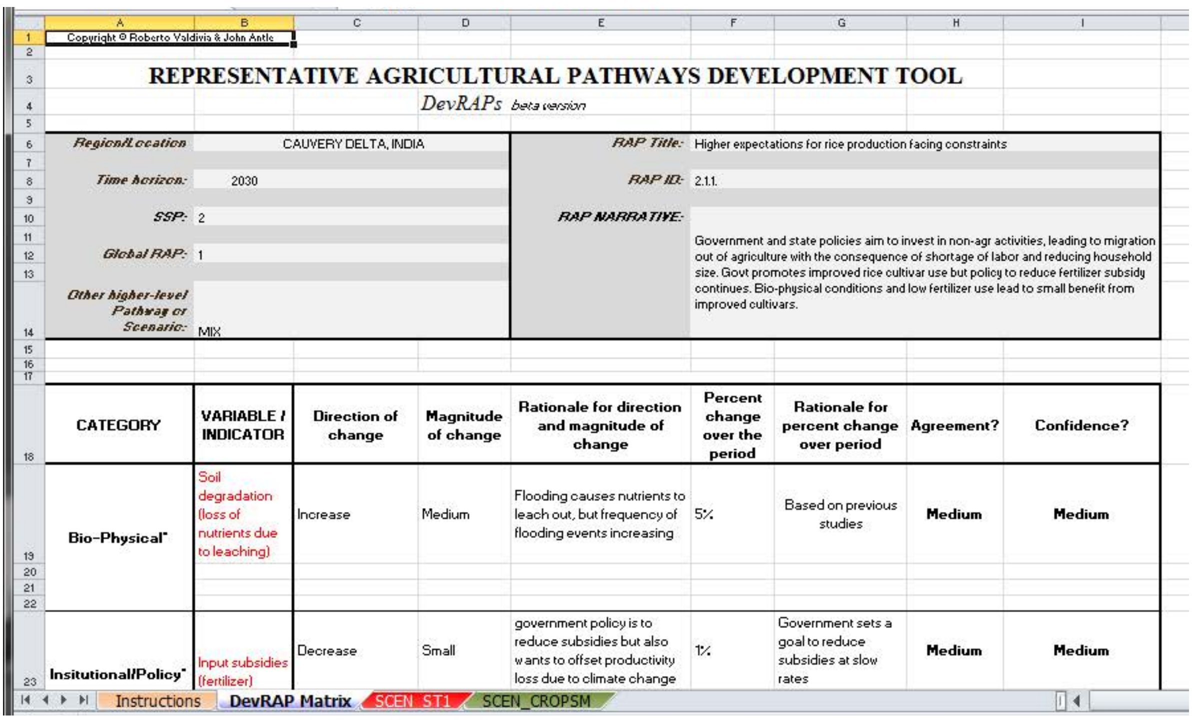


Figure 5. Screenshot of the DevRAP tool v1.0

a. **Building the RAP narratives and quantitative trends.** In this section we outline the steps to build RAPs narratives for AgMIP's regional teams. RRTs should use the DevRAP tool (See **Figure 5**) to develop and document RAPs (Valdivia and Antle 2012).

1) *Identify members of the RAPs development team.* Key members of the research team representing climate, crops & livestock, and economics. Outside members may be solicited if additional expertise is needed.
2) *Define time period for analysis*: AgMIP has designated four "time slices" analysis, current, near-term (2010-2039), mid-century (2040-2069) and end-of-Century (2070-2099). Primary focus is placed on the mid-century period.
3) *Select higher-level pathways:* Following the concept of a nested approach, relevant narratives and quantitative information from selected higher level pathways (e.g. SSPs, Global RAPs) need to be extracted. AgMIP regional teams are recommended to begin using SSP2 (see **Box 2** for a summary description).
4) *RAPs research process:*
    a. First meeting:
        i. Start with a "Business as usual" (BAU) RAP
        ii. Team members identify key parameters that will likely be affected by higher level pathways and draft RAP narrative
        iii. Team members are assigned variables for research
        iv. Team members conduct research –use of templates for reporting and supporting documentation. These templates can be distributed to experts for feedback
    b. Second meeting:
        i. Team members report findings and discuss storylines for each variable
        ii. BAU RAP is finalized using the DevRAP tool and complete the following information:
            1. Complete information for each parameter:
            2. Direction, magnitude & rate of change
            3. Narrative logic for changes
            4. Check for internal consistence and with higher-level pathways and models' variables
            5. Level of agreement among participants
            6. Level of confidence among participants
            7. If level of agreement and/or confidence are low, repeat process until acceptable levels are achieved.
            8. Assess whether one or more parameters need to be revised by other experts or selected for sensitivity analysis.
            9. Document source of information (pathway, model, literature, expert).
        iii. *Additional RAPs are identified*
        iv. *Process similar to BAU is carried out with additional background research*
    c. *Meeting or workshop to present and distribute RAPs to stakeholders and outside experts to obtain their feedback.*
    d. *Meeting(s) to create additional RAPs –Follow similar steps as in a, b and c.*

5) *Modelers develop Scenarios* (see section below)

> **Box 2. Shared Socioeconomic Pathway #2 (SSP2) Summary: Middle of the Road**
>
> In this world, trends typical of recent decades continue, with some progress towards achieving development goals, reductions in resource and energy intensity at historic rates, and slowly decreasing fossil fuel dependency. Development of low-income countries proceeds unevenly, with some countries making relatively good progress while others are left behind. Most economies are politically stable with partially functioning and globally connected markets. A limited number of comparatively weak global institutions exist. Per-capita income levels grow at a medium pace on the global average, with slowly converging income levels between developing and industrialized countries. Intra-regional income distributions improve slightly with increasing national income, but disparities remain high in some regions. Educational investments are not high enough to rapidly slow population growth, particularly in low-income countries. Achievement of the Millennium Development Goals is delayed by several decades, leaving populations without access to safe water, improved sanitation, and medical care. Similarly, there is only intermediate success in addressing air pollution or improving energy access for the poor as well as other factors that reduce vulnerability to climate and other global changes.
>
> Source: O'Neill et al. (2012).

b. **Quantifying Economic Model Parameters.** RAP narratives are next used to construct parameter sets for crop and livestock models and for economic models, including TOA-MD. Here we discuss creating parameters for TOA-MD using the DevRAP tool; research teams can create other parameter sheets for other models they may be using. The sheet SCEN_STi (where i=strata 1,2…) in the DevRAP tool is designed to create and document scenarios for the TOA-MD model. One or more scenarios can be constructed for each RAP as follows:
   a. *Create name and short narrative to describe the scenario*: It is important to document the key characteristics of the scenario, thus the narrative and scenario name must contain elements to understand what the scenario is about.
   b. *Identify model parameters*: The DevRAP tool includes the list of parameters used in the TOA-MD. The team will identify the parameters that will be quantified for the specific scenario.
   c. *Quantify each parameter*: use RAP information to assign a value to each parameter. Data for these parameters can be obtained from the literature, modeled or from expert judgment, and these need to be documented.

c. **Quantifying Management and Technology parameters for crop models**. Similar to steps 6-8 above, the team will use the SCEN_CROPSM sheet in the DevRAP tool to quantify specific crop model parameters (fertilizer level, sowing density, improved cultivars, etc.) based on RAP narratives and scenario details (e.g., RAPs packages).

3. Assemble existing data from experiments and calibrate crop models for regionally-relevant cultivars.

The target outputs from this set of activities are high quality data that are entered into the AgMIP ACE database and used for calibration of multiple crop models for selected sites. The data and model simulations will provide scientific evidence that the models are adapted to the crops and environmental conditions in the region and have cultivar characteristics/parameters that can be used to simulate the crops that are to be studied in the region. This is what is typically done in crop modeling training programs and in research projects. It is likely that the RRTs already have accomplished this for some subset of crops and crop models to be used in the studies. This activity is intended to document those data and past efforts, bring together new data, and ensure that the models to be used have gone through this phase of work. It is anticipated that there will be relatively few site-years with data for any of the selected crops, but those data will be placed in the ACE database and used to calibrate cultivars and improve the adaptation of crop models for the regions. Suggested components of this activity are as follows.

a. **Assemble data from past experiments for calibration of regionally-relevant cultivars for selected crop models** for selected crops. This includes crop, soil, and climate data for site-specific experiments and field trials in the region. This will require input from agronomists, crop modelers, climate, and IT project team members.

b. **Input data into sentinel site ACE database** for use in multiple crop models.

c. **Using the AgMIP IT tools, create input files for each crop model.**

d. **Using methods provided by each crop modeling group (e.g., DSSAT, APSIM, perhaps others), simulate the sentinel site experiments** and estimate cultivar-specific parameters to best simulate the experimental results. These results will help set cultivar characteristics and perhaps soil conditions for regional simulations to be carried out by the teams (see below).

e. **Secondary focus will be estimation of productivity parameters**, relative to initial conditions, crop residue, soil organic matter pools, and soil fertility for the site-specific sentinel data (NOTE: these steps will be repeated to be more appropriate for the household survey and regional simulations where site-specific information is not available).

f. Document model simulations (inputs, management, outputs, soil, climate, cultivar coefficients) by placing them in the ACE database, along with explanatory text and appropriate tables and figures showing the quality of the calibration of cultivar coefficients.

4. Assemble existing data and calibrate livestock models for regionally-relevant livestock breeds.

The target outputs from this set of activities are high quality data can be used for calibration of livestock (and rangeland production) models for selected sites. The data and model simulations will provide scientific evidence that the models have (i) breed parameters that allow simulating the common animal breeds of the region, and (ii) feed quantity and quality input data that characterize the on-farm and off-farm (rangeland) fodder production of the region.

This activity is intended to compile existing data and past efforts, identify gaps and collect the necessary new data, and ensure that the models have been properly calibrated.

The necessary data falls into four broad categories (*with indication of potential data sources in italics)*. Data (including metadata) will be stored in an AgMIP database.

a. Feed trial data, in which body weight, calving rate, milk production and feed input (quantity and quality of feed that was offered to the animals) is recorded

(Experimental data from existing databases, reports, publications).

b. Information on **feeding practices** by farmers and the average feed calendar and feed availability in the area of interest. The following questions should be answered for this aspect

i. In which months do farmers feed crop residues, forages, etc., from which crops, and to which types of animals?
ii. In which months are the herds relying on rangelands (100% or to a certain degree), and does that differ between different animals?

(Data from household surveys, focus group discussions, expert consultations)

c. Information on **rangeland biomass productivity** in relation to climatic variability

(Data from biomass productivity assessments, remote sensing analyses, from databases, reports, publications)

d. Information on the **feed quality** of the different feed sources (forages, crop residues, concentrates, and rangeland) over time (as this varies in the different seasons). Minimum feed quality information requirements include dry matter content, dry matter digestibility, crude protein content, and metabolizable energy.

(Data from laboratory analyses assessments, remote sensing analyses, from databases, reports, publications)

Model calibration will be conducted by estimating the breed specific parameters that result in the closest simulation of important livestock performance indicators such as body weight, calving rate, and milk production. Sensitivity analysis for a number of animal breed and feed input parameters will add confidence that the obtained parameter values are acceptable and result in reasonable model predictions for the region.

Proper documentation of the sensitivity and calibration exercises should include explanatory text, appropriate model performance statistics, and tables and figures showing the quality of the calibration.

5. Assemble and quality-control current climate series.

The key products from this activity will be a high-quality version of in-situ climate observations in .AgMIP format for each location where crop models will be used, a file documenting the changes made to the original raw observations, and summary maps and statistics characterizing the region being analyzed. The following methods are recommended:

a. Assemble and assess quality of station observations.

- Identify weather stations that best represent selected crop modeling regions.
- Obtain as much of the 1980-2010 period as possible (Daily precipitation, maximum and minimum temperatures, solar radiation or sunshine duration, wind speed, dew point temperature, vapor pressure, and relative humidity).

- Convert to .AgMIP units and format with missing data given a value of -99. The AgMIP format is described in the AgMIP protocols available at http://www.agmip.org.
- Name the climate series site with a 4-character code (first 2 characters from internet country code and second 2 characters representing location) following the guidelines in the AgMIP protocols (e.g. "NLHA" for Haarweg, Netherlands).
- Begin a text file to document changes made in the quality assessment and quality control of the raw files (e.g., "NLHA.info").
- Identify outlying (+/- 3 standard deviations probably deserves a closer look) and questionable data that may be corrupted. The best approach remains plotting out the dataset elements as time series to see if anything looks amiss.
- Check to see if data are plausible physically (e.g., questionable value supported by other variables), temporally (e.g., questionable value supported by preceding or following values), or spatially (e.g., questionable value supported by neighboring stations). If values are not plausible, replace with a value of -99.
- If vapor pressure, dewpoint temperature, or relative humidity correspond to a time of day other than mid-afternoon (~maximum temperature), approximate values at the time of day of maximum temperature will be computed, by conserving more robust dewpoint temperature or vapor pressures (which can be calculated using temperature at time of measurement) and then recalculating relative humidity using maximum temperature.

b. Obtain background daily climate time series (1980-2010) from the AgMERRA dataset provided by the AgMIP Climate Team (Ruane et al., 2015). The output of this activity will be a complete set of estimated daily climate data for use in filling in missing data for observation stations. (If the observational dataset is fully complete this step may not be necessary). AgMERRA data are available at http://data.giss.nasa.gov/impacts/agmipcf and are described in Ruane et al., 2014, but an individual location's .AgMIP-formatted time series may be extracted using either FACE-IT workflow tools or via an email to Alex Ruane (alexander.c.ruane@nasa.gov) providing the latitude and longitude, elevation, and site (name and country).

c. Fill in missing/flagged observation data using station observations and the AgMERRA estimated climate series. Note that two overlapping observational sets may be combined in a similar manner. This set of activities will provide a continuous, complete, physically-consistent daily climate series from 1980-2010 in .AgMIP format for use with the crop models. Suggested steps are:

- Go through station observations and fill in all data gaps as follows.
- Use simple interpolations for short data gaps (e.g., if 3 or less days are missing fill in by interpolating from good values on either side). Use caution if strong outlier exists on either side as this may not be an effective approach (e.g., if strong rain event precedes data gap we can't assume that it will have persisted throughout gap.

- For moderate gaps (e.g., 4-10 days) use background dataset to fill in gaps and bias-correct using surrounding good data (adjust mean to ensure approximate continuity with beginning and end points).
- For longer gaps use background datasets to fill in gaps and bias-correct using climatological biases calculated by comparing background dataset to good station observations (e.g., if July Tmax in background dataset is typically 0.6˚C too warm, subtract 0.6˚C from background dataset when filling in a July data gap; if observed rainfall is typically only 90% of background rainfall in October, multiply background dataset by 0.9 to fill in October gaps).
- Ensure that filled in data are physically plausible by checking the following:
    - Relative humidity does not exceed 100%
    - Relative humidity, vapor pressure, and dewpoint temperature are physically consistent at time of day of maximum temperature.
    - Solar radiation is not greater than astronomical maximum (can use historical monthly maximum as proxy) or below zero.
    - Maximum temperature is at least 0.1˚C above minimum temperature.
- Place historical climate data into .AgMIP format using the Excel template provided by Alex Ruane (alexander.c.ruane@nasa.gov).

d. Approximate climate time series in regions for integrated assessments. This set of activities produces a set of climate time series that corresponds to each crop or livestock modeling location in an integrated assessment region and forms the 1980-2009 (current) climate series identified in Table 1. (Note that this procedure is automated in the AgMIP Climate Scenarios Guidebook using the “farmclimate” routine). Working with the crop and economic modeling teams, recommended methods include:

- Obtain desired latitudes and longitudes for each integrated assessment site to be modeled. Name each station with a 4-character code.
- Identify as many weather stations in (or nearby) region as possible. Quality control these datasets following methods above, then assign each of the integrated assessment locations to the most representative weather station (“corresponding station” may not always be selected by geographic distance alone, but may also factor in climatic zones and/or elevation).
- If there are additional precipitation gauges (where other variables are not observed), determine which integrated assessment locations correspond to these and start with this precipitation record.
- Estimate differences in monthly climatologies between integrated assessment locations and corresponding station location using AgMERRA dataset (if distances are greater than ~50km) or WorldClim dataset (if distances are less than ~50km). Adjust corresponding station in a manner similar to the gap-filling bias adjustment to estimate integrated assessment climate series.
- Depending on the number of farms, it may be suitable to categorize each farm into a smaller number of groups that experience nearly the same climate and then create climate series for these groups rather than each individual farm.

e. Create an AgMIP Agro-climatic Atlas for Current Period Climate for eventual publications and integration in AgMIP Impacts Explorer. This atlas will contain maps and plots of important agro-climatic variables for the region. Recommended methods include:

- Generate regional maps of mean temperature and precipitation during historical baseline period from observational data and from GCMs to be used in scenario generation.
- Identify agriculturally important climate metrics. If region is affected by a prominent monsoon, determine which monsoon metrics are important to regional agriculture. Compare climate information with planting rules of thumb from farmers and/or crop model configurations if possible.
- Calculate these metrics and produce maps using observational products during the historical baseline period (in consultation with local experts and stakeholders).
- Identify trends in historical record (utilizing a Mann-Kendall test for statistical significance), most importantly for temperature and precipitation within the growing season.
- Analyze uncertainty among observational products (if available) as reference for future uncertainties.

6. Assemble data and simulate crop models for analysis of yield variations.

Matched case. Ideally, regional projects will use on-farm survey data for which the crop models can be used to simulate each field that was surveyed. This will provide simulated results for the "matched" case where the models use climate, soil, and management for each field to simulate productivity that is then "matched" with observed yields for each field. In order to simulate each field, the teams will need to make assumptions about crop model inputs that are needed but not collected in the farm surveys. These assumed inputs should be developed with advice from agronomists in the region, and they will be documented along with the observed field survey data for each simulated result. Assumed inputs are combined with the survey data by means of a field overlay DOME file (see **Appendix 3**).

Crop modeling team members should analyze these matched results to be sure that they were correctly produced with well-defined and documented inputs and to be sure that results are reasonable. Invariably, there will be biases between simulated and observed survey data, and the modelers should analyze means, variances, biases, and other characteristics of the results prior to confirming that they are ready for use in the economic analyses.

Crop simulation outputs from multiple models will be prepared in the AgMIP harmonized Crop Model output (ACMO) format by the crop modeling team for use by the economists. This file will document all of the inputs and the DOME assumptions used in the model simulations as well as provide a summary of crop productivity outputs (e.g., yield).

Unmatched case. If farm survey data are not available, crop modelers should work with multiple years of historical yield statistics at a district level. In this "unmatched" case, simulated yields cannot be matched one-to-one with observed farm field survey data, and variations in climate, soils, and management inputs across the region will need to be defined and sampled from. This should be done in a representative manner based on available information and expert opinion, particularly about variations in management practices and soils across farms within the district. In this case, comparisons of crop model results will be aggregated to a district level for comparing with district yields and analyzed. Also, a report should be written on methods and results of crop model calibration, aggregation methods, uncertainty associated with seasons, and biases relative to regional aggregated yields.

Recommended steps include:

a. **Matched Case.** Assemble matched yield case data from household farm survey from sub-regions, where crop yield and minimal management (sowing date, fertilization, etc.) are available along with household economics information for 50 to 200 farmers. If it is not possible to simulate each field to produce matched outputs, crop modelers will need to use procedures for unmatched results (see 6.c. below and **Appendix 3**).
   - Download the latest AgMIP Tools (ADA, QuadUI and ACMOUI) from the http://tools.agmip.org/ website.
   - Follow the more detailed instructions in **Appendix 3** to enter yield survey data into spreadsheet templates.
   - Work with regional Agronomists and Soil Scientists to identify the most likely soils for each field in the survey. These data can be added to a separate worksheet in the survey data spreadsheet template.
   - Field Overlay spreadsheets can be used to fill in any information that is missing from the survey, such as initial soil water, initial nitrate and ammonium, soil organic carbon degradation, manure application dates, fertilization dates, prior crop residue, etc.
   - Work with Climate colleagues to identify climate information/sites.
   - Use the ADA and QuadUI applications to convert these spreadsheets into model-ready input files for multiple crop models.
   - Use crop cultivar coefficients that have been calibrated with independent sentinel site data in the region (from Section 3 above).
   - Simulate the matched case survey data, compute means and standard deviation of observed and simulated. Analyze simulated results by computing various statistics and compare with observed statistics, including comparison of yield distributions, means, variances, and characteristics of bias between observed and simulated yields and outliers. Depending on these analyses, crop modelers may decide to accept these inputs as baseline soils and management conditions for further analyses or they may need to make changes in the assumptions in conjunction with agronomists familiar with production in the region. Standard output files (ACMO) are used to provide crop model inputs and outputs for use by economists.

c. <u>Unmatched Survey and Simulation Fields (or Regional Historic Yields).</u> If there are no yield data available from household surveys, it will not be possible to simulate a yield for each farm as in the matched data case. In this case, crop modelers will need to work with economist team members and agronomists in the region to assemble information on variations in management and soils in the region for this “unmatched” case. Assemble soil, typical management, and typical cultivar information for the region along with long-term historical crop statistics data (for district level or higher) for use in evaluating crop model abilities to simulate regional yields and production. Methods for doing this are:

- Yield statistics of crops will be collected for the region over historical time periods of 30 years.
- Cultivar life cycle information will be assumed correct from the site-specific sentinel site data.

- Survey information will be collected with input of agronomists and soil scientists, to represent the distribution of weather stations, soils information, sowing dates, cultivars, residue return, soil organic matter pools, and fertilization that represents the region being predicted.
- Use QuadUI software tools (as above) to create model-ready input files for multiple crop models to simulate historic observed years with and without adaptation as well as future climate with and without adaptation scenarios.
- Similar to the matched case (6.b), crop modelers will create ACMO files for use by economists and prepare reports and publications that describe and interpret biophysical results of the study.
- For purposes of evaluating crop model abilities for simulating regional or district-level yields, crop model teams should aggregate yearly simulated results (over climate sites, soils, sowing dates, cultivars, management) to the district level yield for comparison with historical district yields (e.g., comparing distributions of simulated and observed yields, mean annual bias, etc.).
- Document model simulations (inputs, management, outputs, soil, climate, cultivar coefficients) by placing them in the ACE database, along with explanatory text and appropriate tables and figures showing the yield distributions, analyses of interannual and spatial variations.
- Create maps and summary statistics e.g., spatial distribution of climate, soils, management, and yields illustrated in GIS mapping methods

7. Analyze Carbon-Temperature-Water-Nitrogen (CTWN) responses.

To establish understanding and credibility of the crop and livestock model applications, climate, crop, and livestock experts will undertake analyses of agricultural responses to changes in key climate and nitrogen factors. This analysis will help to identify vulnerabilities and the importance of various uncertainties in the modeling framework.

a. **Select Representative Farmer Field from Phase 1:** RRT teams will come to the workshop with a selected single "representative" farmer field from their farm-survey files from Phase 1, where the yield is relatively median/typical of the farms and where DSSAT and APSIM model predictions are reasonably close and the simulated value for each model is median to that model (necessary to be median to that model if the models differ considerably on average). Each RRT team should come with data for their favored crops, several if they want. However, we hope, at a minimum, to have five maize sites (all African sites and South India) and three rice sites (Pakistan, IGB, and South India) to allow cross-region comparisons of DSSAT and APSIM models for maize and rice. For the selected single field, accept the soil, cultivar, and DOME functions as configured for Phase 1.

b. **Verify simulations** run **for that single farm** and document the soil, initial conditions for soil water, NO3, NH4, root residue, prior crop residue, farmer fertilization with N, and manure application, the soil SOC, the SOC method used, and SOC pools.

c. **CTWN Factor Variation:** for each single farm site, with 30 historical weather years, we will vary one at a time (Table 4): CO2, Tmax/Tmin, rainfall, and fertilizer N over a range for each variable; and we will also conduct the standard 99 sensitivity tests of the Coordinated Climate-Crop Modeling Project (C3MP; Ruane et al., 2014; McDermid et al., 2015). Results will not be passed on to economic analysis, but will be used to interpret different responses of DSSAT and APSIM models to climatic factors and N. *This is not a climate impact assessment exercise, but rather to interpret how and why the models differ.*

Table 4: Description of single factor analyses of CTWN response.

CTWN Single Factor analyses	[CO_2] (ppm) at N=30 kg/ha	[CO_2] (ppm) at N=180 kg/ha	Tmax/Tmin (˚C)	Rainfall (% of current)	Fertilizer (kg/ha)
	360 450 540 630 720	360 450 540 630 720	-2 0 +2 +4 +6 +8	25% 50% 75% 100% 125% 150% 175% 200%	0 30 60 90 120 150 180 210

QUADUI and DOME templates will be available to create linkage and model-ready files to allow simulating the 32 single factor levels and the 99 C3MP levels. Then both the DSSAT and APSIM models will be simulated for that site and the outputs of the models placed in ACMO files. Be sure to check for any model warning messages or log files.

Creating Graphs and Interpreting Differences between Models: R (CTWN) and Matlab (C3MP) programs have been created and distributed that read the ACMO files and create x-y plots of yield responses of two models in same graphs versus the single linear factors of temperature, CO2, rainfall, and N fertilization. Those teams that have capacity and are comfortable with Matlab may obtain, utilize, and modify the Matlab routines (keep C3MP Coordination informed). For C3MP (99 values), the standard C3MP analyses and graphical presentation software will be used.

- **Lineplots of Linear Factor Analysis:** crop response variable versus time (1980-2010). This is mostly for preliminary evaluation prior to doing x-y plots. Don't over-do this. Get to the x-y plots relatively soon.
  - A time-sequence of 30 years of simulated yields will be plotted to evaluate historical patterns in yield,, to identify anomalous years and sources of variability
  - Multiple crop models will be displayed on the same plots for comparison
  - Additional lineplots will be created to show other crop response variables (ET, E, T, and N uptake), in order to identify drivers and reason for yield changes. These plots could lead to ideas on correlating yield to climate and N variables.
  - Cumulative probability of yield exceedance plot will be created with both models on the same plot to illustrate weather risk and crop model variation for that site.
  - If the results suggest, a single year simulation may be evaluated to see how time course of growth, transpiration, soil evaporation, and evapotranspiration varies for APSIM versus DSSAT during the season, especially under a relatively low rainfall year.

- **X-Y Graphs with Boxplots of Linear Factor Analysis: Yield versus C, T, W, N** where the x-axis is the C, T, W, or N variable. Mean yields and box plot (over 30 years) will be computed for each level of the single factors of CO2, temperature, rainfall, and fertilizer N, and plotted against the factor CO2, temperature, rainfall, and N level on the x-axis. The means and box-plots for APSIM and DSSAT will be shown on the same x-y graphs to allow intercomparison of the different models
    - Mean yields and box plot (over 30 years) will be computed for each level of the single factors of CO2, temperature, rainfall, and fertilizer N, and plotted against the factor CO2, temperature, rainfall, and N level on the x-axis. The R-program shows the two crop models for comparison (e.g., side-by-side boxplots at each x-axis level (e.g., +2 degrees Celsius).
    - As appropriate, other variables such as ET, E, T, and N uptake of both models will be plotted against the corresponding CTWN factors.

- **2-Dimensional, Cross-sectional impacts response surfaces (IRS)**
    - Teams will run C3MP sensitivity tests (CO2, temperature, and rainfall) according to C3MP Protocols. N will not be a factor in C3MP unless a strategy is figured out.
    - Teams will use C3MP reporting template (or ACMO file imported directly) to submit results from tests
    - These results will utilized in Matlab scripts to create emulator, and IRS that show contoured percent change in yield for 2D cross-sections of the CTW space for each crop model.
    - IRS can then be created, for contrasting two crop models, and possibly also contrasting across multiple sites (other team sites).

Advantages of both linear factor and C3MP analyses is that it allows the C3MP results to be analyzed for the interactions between the C, T, and W factors , while the linear factor analyses provides a cleaner, simpler evaluation of the crop and crop model response individually. We expect that each of the methods will produce separate, but complementary papers, unique to AgMIP Regional Research Activities.

AgMIP anticipates several strong intercomparisons of DSSAT and APSIM models (one paper on maize, one on rice, and possibly other crops) with emphasis on interpretation of physiological, model structure, soil C methods, transpiration methods, and parameterization reasons for differences among models in response to C, T, W, and N.

In thinking about potential publications resulting from the CTWN/C3MP activities, please consider the following paper structures:

Paper Structure for variation among crop models in response to single factors:

- **Hypothesis/Objective:** The null hypothesis is that the crop models do not differ in their response to temperature, CO2, rainfall, and N fertilization. The objectives are to document any differences in responses and to interpret the reasons for differences in response.
- **Multi-author:** The RRTs may potentially work together to evaluate the sensitivity of one crop, with crop modelers leading the effort, climate scientists involved, with strong connections to original developers of the respective crop models who can help interpret reasons for the responses obtained, based on their knowledge of model structure and method.
- **Figures:** Single factor yield (biomass, ET, transpiration, N uptake, etc.) responses to CO2, Temperature, rainfall, and N fertilizer, always with all models

shown on the same graphs. In some cases, showing time-series simulation of biomass, LAI, N uptake, or ET versus time, is a way to understand why DSSAT and APSIM differ.

Paper Structure for C3MP-type results, following standard C3MP approaches, but adapted to help interpret what the models are doing, possibly with given GCM “delta” temperature and CO2 placed on the plots to hypothesize climate impact for that farm.

- **Multi-author:** is encouraged. Please work with C3MP coordination, to make sure no efforts are duplicated in the wider C3MP community
- **Figures:** IRS figures, and other relevant plots. Remember the goal here is to contrast different crop models on the same figures somehow.

8. Assemble data and simulate livestock models for analysis of livestock productivity at the household level.

In this activity household survey information and outputs from the crop models need to be combined to generate the necessary livestock model input data.

Firstly, the livestock model has to be fed with feed availability information coming from the crop models and, if available, rangeland models. These yield data need to be combined with information from household surveys on field sizes to calculate total farm-level feed production. Secondly, household survey information will also serve to derive the initial herd size and composition for each household, which is needed as input data for the model.

For the grazing component of the livestock data, rangeland models could be used if available and well calibrated. If these are not available, or if confidence in modelling results is not (yet) satisfactory, other options exist to estimate the grazing component. One option is to use a crop model like APSIM or DSSAT to simulate tropical grass productivity in response to climate. Outputs from these crop models should be checked against reported rangeland biomass availability figures from the literature before use. A third option for estimating annual productivity of grazing lands, is to use rainfall use efficiency values from the literature in combination with seasonal rainfall. A final option, which does not allow incorporating annual biomass variability, is to work with reported average values of biomass availability. In all cases, rangeland productivity estimates have to be combined with information on rangeland area and stocking density to derive feed availability per animal.

On-farm crop residue and forage production can be derived from the crop modelling results. Biomass yields have to be multiplied with field sizes (from household surveys) to calculate total farm-level feed production and combined with the actual herd size of a particular year in the simulation, to obtain feed availability per animal, which is the final input used by the livestock model.

Simulated livestock productivity in terms of herd size and dynamics (number of animals born, sold, died) and milk production should be compared with information derived from the household survey. Invariably, there will be biases between simulated and observed survey data, and the modelers should analyze means, variances, biases, and other characteristics of the results prior to confirming that they are ready for use in the economic analyses.

9. Assemble economic data for regional economic analysis and develop skills for using the regional economic model.
Outputs from this set of activities include at least two economist members per project team that are capable of performing economic analyses in their respective regions and data assembled on baseline socioeconomic and agricultural production data in their regions. An output will be crop modelers and economists with experience in interdisciplinary collaboration in co-developing data sets for use by both teams (e.g., historical yields and socioeconomic survey data), with the data input to the AgMIP database. Another output is the TOA-MD model set up to simulate economic outcomes for the region, using baseline socioeconomic data. Specific steps include:

a. **Identify economic data and corresponding study components (see the TOA-MD model and supporting documents for further details).**

b. **Work with the climate and crop model teams** to produce and analyze baseline crop simulations for sites that are jointly selected for the region, based on available data from regional statistics and/or on-farm surveys. This step requires direct cooperation among disciplinary team members and relies on the above steps on collecting climate series and calibration of crop models for regional yields.

c. **Estimate economic model parameters using the available data (see the Appendix 2 and TOA-MD model and supporting documents for details)**.

d. **Prepare a report (**following AgMIP template) describing the existing systems and documenting the data used for regional economic analysis and parameter estimates.

10. Create downscaled climate scenarios
Create downscaled climate scenarios based on AgMIP protocols (Ruane et al., 2015), for use in the assessments of climate change studies, and provide future scenarios for use with crop models in the AgMIP database. Note that these procedures are captured in scripts contained in the AgMIP Guidebook for Climate Scenarios. A key output from this set of activities will be future climate scenarios derived from the latest IPCC climate models and downscaled for use in the target regions. These scenarios will be in the .AgMIP climate data format and ready for multiple crop model simulations of impacts and agricultural adaptation for each region. In addition, a climate atlas will be produced of important climate variables and derived agriculturally-important indices. These atlases will include maps for use in scientific publications and for communication of results to stakeholders.

a. Select subset of GCMs for full analysis and create AgMIP Agroclimatic Atlas showing future climate change scenarios with uncertainties using maps with probabilities. The subset of models is a necessary step considering the limited resources and large number of combinations possible in further combination with crop, livestock, and economics models. We will focus on the RCP8.5 Mid-Century (2040-2069) period. Maps and summary results will be published and also communicated to stakeholders via the Impacts Explorer Tool. Specific methods are:

- Make plot of growing season temperature and precipitation change from full GCM ensemble. Highlight models chosen for representative subset, drawing a relatively hot/dry, hot/wet, cool/dry, and cool/wet GCM as well as a GCM

representing the middle of the ensemble projected changes. Note the weights given to each GCM as these will be used by economic and crop modelers in the final analyses. This can be created using the R CMIP5_TandP script.
- Create monthly box-and-whisker diagram to show current climate and projected range of future climates for mid-century RCP8.5. This can be created using the Matlab 'CMIP5_TandP' script.
- Produce region-wide maps of CMIP5 climate change projections, including median changes in mean quantities, variability, and extremes (along with corresponding uncertainties) for temperatures and precipitation.
- Also produce maps for agriculturally important climate metrics under future climate conditions for comparing with those produced for historical baseline climates.

b. Create CMIP5 mean and variability change scenarios. This activity will produce .AgMIP-formatted climate scenarios including both monthly and sub-monthly changes in temperature and precipitation. These procedures are captured in the "agmipsimple_mandv" scripts in the AgMIP Guidebook for Climate Scenarios. In many regions there are not sufficient resources or available regional climate model (RCM) results to capture important uncertainty in climate projections, however where these are available they are particularly helpful for their representation of sub-seasonal metrics that are often affected by smaller-scale atmospheric dynamics. Suggested methods include:
- Calculate monthly changes in mean maximum temperature, minimum temperature, and precipitation by comparing future 30-year climate periods to the current (1980-2009) climate period from the same GCM/RCM combination (where available).
- Calculate monthly changes in the standard deviation of maximum temperature, the standard deviation of minimum temperature, and the number of rainy days (precipitation>0.1 mm) by comparing future 30-year climate periods to the current climate period from the same GCM/RCM combination (where available). The shape parameter of the gamma distribution for wet events may also be of interest from RCM results, but is generally not of sufficient quality in GCM simulations.
- Impose these monthly changes on baseline climate series for all sites used in the analyses using a stretched distribution approach that adjusts each event by comparing existing and desired values by distributional percentiles.
- Assume that solar radiation, winds, and relative humidity daily variables from the historical daily climate records are unchanged. Ensure that vapor pressure, dewpoint temperatures, and relative humidity are physically consistent at time of maximum daily temperatures.
- Produce mean and variability change scenarios for all CMIP5 GCMs at the best-calibrated site in each region, and then create future scenarios at every farm site using the 5-GCM subset identified above to drive crop and livestock model simulations.

c. Create CMIP5 delta-based climate scenarios (optional). These scenarios will be based on historical baseline daily climate data, with each day's weather variables perturbed using the changes in climate model outputs for future time periods versus those same model outputs for the historical time period. These scenarios are made using the "agmipsimpledelta" routines in the AgMIP Guidebook for Climate Scenarios and may be compared against the more complex mean-and-variability change scripts above. This is a simpler but more straight-forward approach that some teams may want to examine and/or compare against the mean-and-variability approach detailed above. Specific methods include:

- For each of these sites, calculate monthly changes in corresponding mean maximum temperature, minimum temperature, and precipitation by comparing future 30-year climate periods (AgMIP defines three main time periods: "near-term"=2010-2039; "mid-century"=2040-2069; and "end-of-century"=2070-2099) to the baseline climate period (1980-2009; use RCP 4.5 for 2006-2009 period) from the same GCM. The Mid-Century RCP8.5 is the priority period for assessment.
- Impose these monthly changes on baseline climate series for all selected sites by adding temperature changes to the baseline record and multiplying by a precipitation change factor.
- Assume that solar radiation, winds, and relative humidity are fixed at the same values that were in the historical time series. Ensure that vapor pressure, dewpoint temperatures, and relative humidity are physically consistent at time of maximum daily temperatures (warmer temperatures have higher vapor pressures and dewpoint temperature at same relative humidity).
- This will result in a 30-year .AgMIP-formatted climate series for a given future period and GCM.

FACE-IT workflows for RIA. An online workflow system is under development which allows the intensive computations and large data archives required for climate analyses associated with the RIA to be deployed on a cloud-server. This system, FACE-IT (Framework to Advance Climate, Economic and Impact Investigations with Information Technology) provides an alternative to using the R scripts for scenario generation and also offers some convenient forms of analyses. Procedures for using this system are not covered herein, but interested users are encouraged to learn more at www.learnfaceit.org.

Finally, if teams are interested in generating Near-term climate scenarios (2010-2039) as an additional activity (not required for RIA), then AgMIP Climate does employ the methods of Greene et al., (2012a,b). This approach allows the influence of climate variability to be compared against climate change, and is especially useful in identifying extreme events for adaptation planning and risk management.

11. Conduct multiple crop/livestock model simulations

The major outputs of this series of activities include simulations of yields by multiple crop and livestock models for multiple sites within the study region. Table 2 depicts six crop and livestock modeling simulation sets, and Table 2 identifies four associated climate change ratios for resulting economic questions, that are needed to address the **Core Climate Impact Questions** described in the Introduction.

Summary of crop and livestock model simulation data sets needed to address the **Core Climate Impact Questions:**

a. **Calibration.** The calibration dataset is discussed in Section 3.
b. **CM0 (Historical).** A simulation of the conditions under which the farm survey data were collected is typically performed for duration of one to two years for crop models and uses observed weather data for each site. For livestock models, a run time of at least 10-12 years is recommended because the livestock models take a longer time to stabilize and yield a reasonable average value of livestock productivity. The comparison of observed to simulated yields from the historical simulation allows researchers to evaluate the models and input parameters, and to compute biases and probability of exceedance. This is the only simulation for which comparison to observed crop yields and livestock productivity is relevant. These simulations and the associated analyses are discussed in Section 6.
c. **Simulation sets CM1 through CM6** use 30 years of daily weather data based on current or future climatology for the entire set of surveyed farms and crops and livestock.
    a. **CM1 (Current climate, current management).** Simulation of the current climate and current production system.
    b. **CM2 (Climate change, current management).** Simulation of climate change scenarios with the current production system.
    c. **CM3 (Current climate, current management, plus adaptation).** Simulation using current climate, but with crop, livestock, and economic changes specifically designed for climate change adaptation.
    d. **CM4 (Current climate, Future RAP).** Simulation using current climate and future management based on technology trends corresponding to a particular RAP.
    e. **CM5 (Climate change, Future RAP).** Simulation using climate change scenarios and future management based on technology trends corresponding to a particular RAP.
    f. **CM6 (Climate change, Future RAP, plus adaptation).** Simulation using climate change scenarios, future management based on technology trends corresponding to a particular RAP, and with crop, livestock, and economic changes which are specifically designed for climate change adaptation.

The steps to perform Simulation Sets CM1 through CM6 are listed below.

This includes simulation of responses across GCMs, farms, and across years within the 30-year periods. Multiple crop/livestock models will be used to simulate variations in climate, soils, and management, thus obtaining within-region variability of production. These results will be put into the AgMIP ACMO database for use in the economic analyses.

- *CM1: Current climate with current production systems technology*: Simulate current period climate series (identified as planting years 1980-2009 in Table 1) for all farms using the 30-year climate series created in Task 5 above, current production systems and a CO_2 concentration of 360ppm for all years (see **Table 5**). Survey data and field overlay data from the historical simulation (Section 6) and calibrated cultivars and livestock breeds from the calibration simulations (Section 3) are used. Seasonal strategy DOMEs are used to generate the 30-year simulations using the historical survey data.
- *CM2: Climate change scenario(s) with current production technology* (no adaptation or RAPs): Simulate mean-and-variability-based climate change scenarios (beginning with RCP8.5 Mid-Century, identified as planting years 2040-2069 in Table 5) for all farms using the 30-year climate series created in

Activity 10 above, current production systems and a CO_2 concentration corresponding to the central year for all simulations (see Table 5). Teams will utilize five climate change scenarios representative of projected regional changes, in consultation with climate experts.

- *CM3: Crop and livestock model simulations with current climate, using adaptation package(s) created via collaboration between the crop, livestock, and economic modeling teams*. Adaptations could be the same as that used in CM6 (to contrast the value of climate-related adaptations in current climate versus future climate). Examples include heat or drought-tolerant cultivars; added irrigation; subsidies for improved seed, inclusion of heat-tolerant forage crops, economic incentives, etc. requiring major investments. Alternatively, teams could design adaptation/interventions for present climate and present technology.
- *CM4: Crop and livestock models will be simulated with current climate for future production technology* (e.g., improved cultivars and livestock breeds, additional N fertilization, use of feed concentrates, altered management) informed by RAPs and technology trends.
- *CM5: Climate change scenario(s) with future production technology* (improved cultivars and livestock breeds, additional N fertilization, use of feed concentrates, altered management) informed by RAPs and technology trends.
- *CM6: Climate change scenario(s) with future production technology, plus an adaptation package.* Create and document adaptation package(s) via collaboration between the crop, livestock, and economic modeling teams. Adaptations should be connected to climate-related vulnerabilities identified in a comparison between CM4 and CM5 results (also CM1and CM2) such as heat or drought-tolerant cultivars; added irrigation; subsidies for improved seed, inclusion of heat-tolerant forage crops, economic incentives, etc. requiring major investments). Do not attempt improved management options associated with representative agricultural pathway and technology trends that define future production systems.
- Organize simulated crop yields and livestock productivity indicators and enter those results into the AgMIP Crop Model Output database (ACMO), which contains variables needed by the economic team members for the regional economic model analyses using ACMOUI application.
- All outputs should be reviewed by crop and livestock modeling team members working closely with economic and climate team members to ensure the results are plausible, e.g., that there are no unexplained outliers.
- When the team has finalized the crop and livestock model simulations and summarized outputs in the ACMO file, then the outputs from the six cases will be used to compute the four change-ratios (shown in Table 3) to be used by the economists in the TOA-MD economic model. For the TOA analyses, question 1 uses the ratio of CM2/CM1, question 2 uses ratio of CM3/CM1, question 3 uses ratio of CM5/CM4, and question 4 uses ratio of CM6/CM5.
- Summarize crop yield and livestock productivity impacts in tables, graphs, and maps for publication and communication to stakeholders. Included in these tables, graphs, and maps should be:
  - within-region variability in impacts, and

- uncertainties associated with crop, livestock, and climate models
- Interpret reasons for variations among crop, livestock, and climate models as well as between regional households

Table 5: Central year carbon dioxide concentrations for AgMIP climate scenarios and time periods, with the Current and RCP8.5 Mid-Century time periods highlighted as they will be the primary focus of integrated assessment. These are the concentrations to be used for all years in a given scenario experiment.

Scenario and Time Period	Planting Year Coverage	Mid-year	[CO2]
Current	**1980-2009**	**1995**	**360 ppm**
RCP4.5 Near-term	2010-2039	2025	423 ppm
RCP8.5 Near-term	2010-2039	2025	432 ppm
RCP4.5 Mid-Century	2040-2069	2055	499 ppm
RCP8.5 Mid-Century	**2040-2069**	**2055**	**571 ppm**
RCP4.5 End-of-Century	2070-2099	2085	532 ppm
RCP8.5 End-of-Century	2070-2099	2085	801 ppm

d. The following **analysis simulation sets** are performed for a single, best-calibrated and representative site in each integrated assessment region. These results are not used to answer the **Core Climate Impact Questions**, but are used to more fully understand the dynamics of the cropping system, and to interpret causes for differences among crop model responses to climate and management factors.

a. **Full GCM simulations.** Examine the full GCM ensemble for a single farm. The outputs from the single location GCM ensemble simulations will be used by the climate team members to place the subset of GCMs in context.

b. **Sensitivity analyses.** Examine the sensitivity of simulated yield from a single farm to weather and fertilizer nitrogen inputs using C3MP and CTWN protocols (See section 7).

FACE-IT workflows for RIA. An online workflow system is under development which allows the intensive computations required for the RIA system to be performed using chains of applications, deployed on a cloud-server. This system, FACE-IT (Framework to Advance Climate, Economic and Impact Investigations with Information Technology) provides an alternative to using the AgMIP desktop utilities for data translation and allows simulations using DSSAT and APSIM for complex workflows, including analysis of multiple climate change scenarios, sensitivity analyses, and adaptation scenarios. Procedures for using this system are not covered herein, but interested users are encouraged to learn more at www.learnfaceit.org.

12. Analyze regional economic impacts of climate change without and with adaptation using the regional economic model.

Outputs will be impacts of climate change on agricultural production, farm income and poverty, and projected rates of adoption of adapted systems. To the extent possible, teams should use results of these sub-national analyses to draw implications for the national impacts, e.g., by extrapolating impacts to regions with similar production systems. The AgMIP regional integrated assessment framework is summarized in **Figure 1**.

Economist team members will use the TOA-MD model (or similar) following the procedures in Appendix 2 to estimate the economic model parameters. Results from the RIA analyses will be summarized with graphs and reports for scientific publications and for dissemination to stakeholders.

13. Archive data and analyses of results for integrated assessments

An important output of integrated assessments will be databases for the regions that will include climate, soil, management, experiments, surveys, regional economic model parameters, and historical yields that will have been used for the analyses in this set of projects that will be highly valuable for additional future analyses as models improve, research and policy questions change, and adaptation approaches evolve. These archived data will be available for broad use, although it is recognized that some data used in the projects (such as daily climate data in some cases, or confidential survey data) may not be archived due to intellectual property rights and data policies. Additionally, archived results from climate, crop models, livestock models, and economic models will serve as the source for various publications and presentations, including web-based information that will be made available for stakeholders. A well-documented archive of AgMIP experiments, outputs, and analysis tools will facilitate future improvements in capabilities to perform integrated assessments of climate change impacts and adaptation at site and aggregated scales.

Figure 6 presents a data flow diagram for AgMIP Regional Integrated Assessments. Data created using the tools and procedures outlined in this document should be archived in AgMIP databases. Research teams shall contribute data to ACE (AgMIP Crop Experiment), DOME (Data Overlay for Multi-model Export), ACMO (AgMIP Crop Model Output) and Regional Economic databases. The AgMIP IT Team will provide tools and training through the regional workshops and web tutorials so that RRTs can interact with the ACE, DOME, ACMO and regional economic databases directly through the AgMIP Data Interchange (data.agmip.org) which connects to AgMIP data nodes. This will allow for storage of standardized databases of crop experiments and yield trials for the region and outputs of crop model simulations.

Data to be archived includes:

a. Climate data
   - Observed weather data for crop model calibration
   - 1980-2010 quality-controlled daily climate data for use in the AgMIP regional assessment
   - Ensembles of daily future climate scenarios

b. Crop Modeling
   - Harmonized (aceb, dome and alnk) data files associated with detailed calibration data from field experiments or other sources.
   - Calibrated cultivar parameters
   - Soil parameters as used in simulations
   - Harmonized data associated with farm survey sites for regional assessments using baseline and future conditions (aceb, dome and alnk files)
   - Crop model outputs for survey, baseline, sensitivity tests, and various future climate conditions (ACMO files)
   - Text summary of climate impacts on yield, considering crop management in survey fields

c. Livestock Modeling
   - Harmonized data files with information from feeding trials, breed-specific productivity indicators, farmer feeding practices, rangeland biomass availability, feed quality

- Calibrated livestock breed parameters
- Feed input data (on-farm and grazing land) and herd size and composition as used in simulations
- Livestock model outputs (milk production, herd dynamics) for baseline, future climate and adaptation conditions

d. Economic data
- Inputs to regional economic models (including survey metadata)
- DevRAP matrix spreadsheet including output data from global economic models used in the RAPS and productivity trends.
- Regional economic model outputs - Impacts of climate change on agricultural production, farm income and poverty, and percentage of winners and losers and predicted adoption rates of adapted technologies.

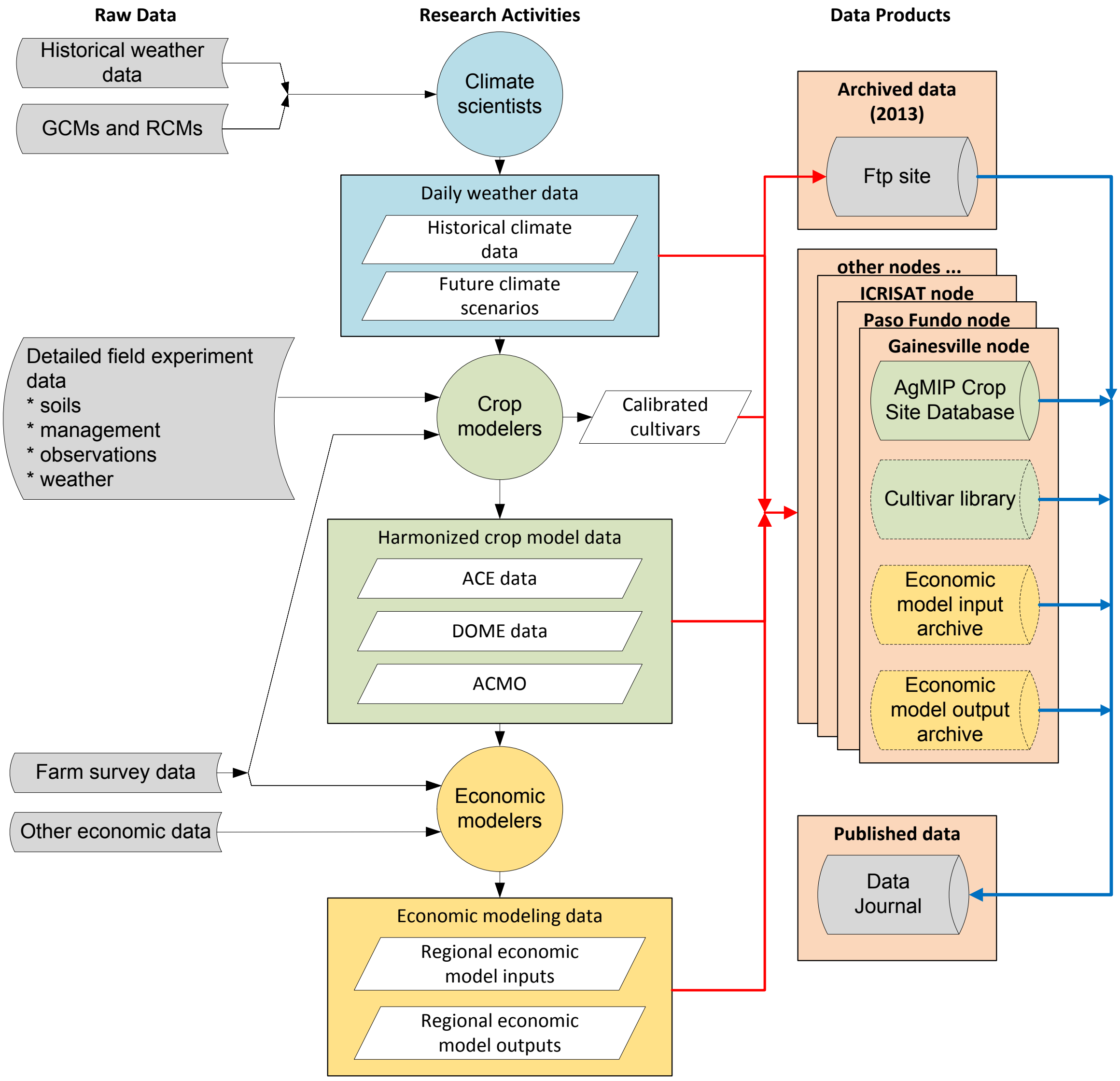


Figure 6. Data flow diagram for AgMIP Regional Integrated Assessments showing AgMIP data products and archive databases

14. Disseminate integrated assessment results.
The key outputs from this set of activities include scientific publications, project reports, results summarized on regional web pages linked to the AgMIP web site, and workshops with stakeholders. Initial and ongoing interaction with stakeholder and policymaking communities are likely to be as valuable as the dissemination of results to these communities, as early and consistent interactions increase buy-in and help develop a more useful and efficient research project

a. Develop RRT-specific web pages for the AgMIP web site. The AgMIP IT Team will provide information on how to create region-specific web pages and will give regional IT team members access to create and maintain that web information. Each region will have its project goals and methods on the site as well as pictures of project activities, output tables, maps, and graphs, as well as news items, for example.

b. Conduct project workshop with stakeholders.

- Invite stakeholders to SSA and SA workshops
- Organize stakeholder sessions at a region-specific workshops to keep them informed and learn from them what information they need for their planning and policy-making responsibilities

c. Prepare scientific publications. AgMIP research is designed to provide results that are well-suited for peer-reviewed journal publications and informing national and regional publications related to climate vulnerabilities, economic development, and adaptation/mitigation planning relative to food production and food security.

Appendix 1

Phase 2 End-to-End Connections and Priorities for Decision Support

AgMIP Regional Integrated Assessments are motivated by the need for cutting-edge scientific information that will aid stakeholders considering various options for policy change or investment. Figure A1.1 demonstrates how this decision support requires a modeling framework connecting economics, crop/livestock, and climate model inputs and outputs, but is also built upon a foundation of credibility established through key validation and analyses (**Figure A1.1**). The protocols and activities described in this document provide credible information and context in support of a range of stakeholders around the world according to this model.

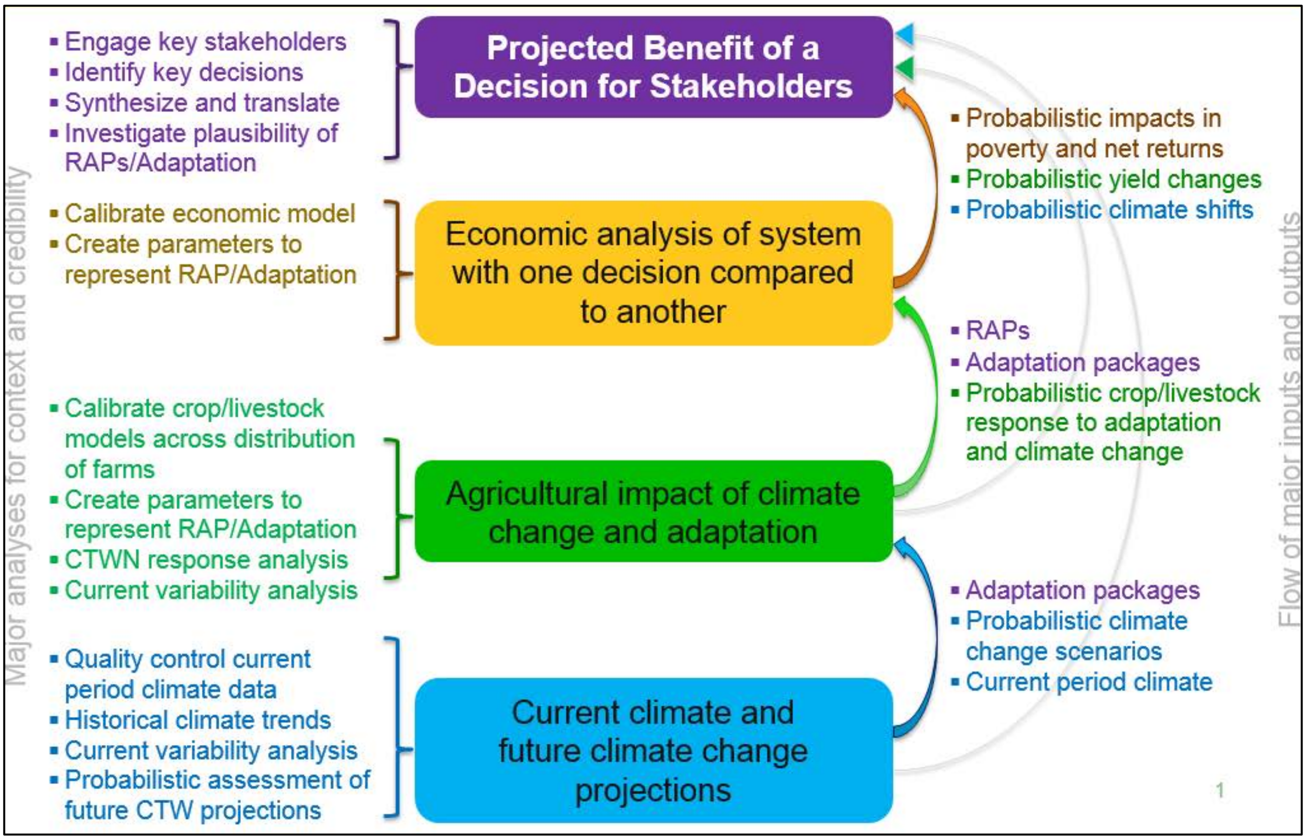


Figure A1.1: Overview of Phase 2 end-to-end project design. *Center:* To inform stakeholder decisions we need economic simulations driven by crop/livestock models driven by climate information. *Right:* Flow of the major inputs and outputs to enable the end-to-end regional integrated assessment. *Left:* Major analyses that are needed to give context and credibility to the outputs of each disciplinary component of the regional integrated assessment. Colors indicate the RRT project teams responsible for each activity (purple=stakeholder unit; gold=economics; green=crop/livestock; blue=climate), and all arrows will be facilitated by IT infrastructure and project communications.

Appendix 2

Calculating Statistics for Climate Impact Assessments Using Crop/Livestock Model Simulations, RAPs and the TOA-MD Model

John Antle and Roberto Valdivia

November 2015

Introduction

This document describes how crop and livestock model simulations and Representative Agricultural Pathways can be used with TOA-MD to implement assessments of climate change impact and adaptation using "matched" and "unmatched" data from crop or livestock simulation models. We use the case of a population of heterogeneous farms with a single stratum and one production activity to illustrate the methods but this can be generalized to multiple activities and multiple strata. This appendix presents methods for the use of data from crop or livestock models to simulate climate impacts by averaging data over time within the "current period" and within the "future period" defined for the analysis.

It is important to recognize that the methods presented here are not designed to represent temporal variability within the current period or within the future period. We focus on the time averaged case because of key limitations of the data that are usually available. In most cases, we *do not* observe yields or management over enough years to measure variation over time for individual farms. Thus, our methodology is designed to use cross-sectional survey data to estimate spatial heterogeneity reflecting bio-physical differences and management differences across farms.

The first section presents concepts and definitions. The second section describes the calculations used to estimate the parameters of the TOA-MD model.

A2.1. Concepts, Definitions and Assumptions

The Four Core Questions

The methods described here can be used to answer the "core questions" described in the first part of this Handbook. Note that Questions 1 and 3 involve assessing climate impacts, and so the TOA-MD model is used as an impact assessment tool. Questions 2 and 4 involve adaptation, in either a current or future period. Analysis of adaptation involves procedures similar to a standard technology adoption analysis as discussed in the TOA-MD documentation.

It is also important to recognize that these Core Questions are not the only logically possible or useful questions that can be investigated with the methods described here. For example, Core Question 2 can be modified to use a changed climate rather than the current or historical climate; also, Core Question 3 can be modified so that the technology specified for System 2 in the economic analysis is adapted to the future climate rather than a technology adapted to the current or historical climate.

Incorporating Spatial and Temporal Variability

We know yields and related outcomes (economic returns) vary over space and time, and this variation is important to understand vulnerability of farms to climate change. Therefore we need to project these distributions into the future for climate impact assessment.

We can describe a variable such as a yield for a production system h used at location j at time t as y_{jt}. Let μ_j be the mean for farm j obtained by averaging its values of y_{jt} over time and let μ_t be the mean for year t obtained by averaging y_{jt} over all farms in that year. We will say that μ_j is the *time-averaged* mean for farm j and μ_t is the *spatially-averaged* mean for year t. Similarly, we can decompose the variance of y_{jt} into spatial and temporal components. To obtain meaningful approximations to the distribution of outcome variables for the TOA-MD model, we often need to stratify populations of farms that come from different sub-populations or different time periods. For example, we may need to stratify farms geographically or by socio-economic characteristics such as size or ownership of livestock.

Our goal is to use the available data to estimate distributions of realized or expected returns to a farming system using the available data. The data needed are:

- Farm survey data that provide observations of current yields, management and other variables.
- Secondary data on average yields for the study region.
- Projected yield growth rates from global agricultural economic models or RAPs.
- Current and future simulated yields of crops and livestock obtained from the crop and livestock teams.

A key limitation of the data is that, in most cases, we *do not* observe yields or management over enough years to measure variation over time for individual farms. Thus, our methodology is designed to use cross-sectional survey data to estimate spatial heterogeneity reflecting bio-physical differences and management differences across farms.

Defining the Study Region and Time Periods

The presentation here is for the analysis of a farm population in an “integrated assessment region,” i.e., a study area defined geographically and possibly in terms of other socio-economic characteristics. Our convention for time t is that it represents a calendar year within a time “period.” The current period H covers a specified number of years, and the future period F is some number of years ahead.

In most cases available farm survey data will come from a year (or years) near the end of the current period used for climate data and bio-physical model simulations, and management data used in these simulations will come from these survey data. For example, AgMIP’s Regional Research Teams are using 1980-2009 as the current period for climate data and crop and livestock simulations. However, most survey data being used are from 2005 or later. For the economic analysis, using a 30-year period as “current” is not practical due to data limitations, the challenges of dealing with real and nominal trends, etc. Therefore, for the economic analysis, we are using the most recent 5-year period centered as closely as possible on the year(s) of the economic survey data for purposes of defining the current period for economic data.

Interpretation of TOA-MD Systems

Following the TOA-MD terminology, every simulation experiment involves two systems, denoted in TOA-MD as System 1 and System 2. Note that the interpretation of system 1 and system 2 depends on the type of analysis being done. For example in core question 1, to assess the effects of climate change on productivity, we interpret system 1 as the current production system in the current period and system 2 as the same system if it were observed in use with the future climate. However, for analysis of the four questions of Table 1, system 1 and system 2 are constructed to represent various combinations of climate change effects, socio-economic conditions and technologies.

To further simplify this presentation, we consider the case of a production system that has a single activity (say, a crop). More generally, the same types of calculations would be applied to each activity in each sub-system (i.e., to all crops, all livestock, all aquaculture activities).

Definition of Climate and Technology

We define a climate as a distribution of weather outcomes, and denote it with the parameter γ_t, where t = H or F. Note that in the Core Questions 1 and 2, a “future” climate is used under current-world socio-economic conditions. This is done for two purposes. First, it can be useful for evaluating how a change in climate could affect current systems; second, the “future” climate can be defined as a climate different from the historically observed climate, e.g., with an increase in extreme events, that could be occurring now under current socio-economic and technological conditions.

Production system technology is defined here in two dimensions: the period when it is used, and the climate it was developed for and presumed to be adapted to. This means that a given technology, e.g., a specific seed variety, performs best in the climate it is adapted to. However, this does not mean that there cannot be a better-performing technology in that climate, even one adapted to a different climate. Technology is represented as T_{ti} where t = H or F represents the time period the technology is used and i = H or F denotes the climate it is adapted to. Note that in the experimental design of the simulations, the technology T_{HF} is used in analysis of Core Question 2 with current climate, so we interpret this technology to be better adapted to a future climate, but could be better performing in the current climate than the current technology.

Technology and Climate Combinations Used in the Four Core Questions

According to this definition, there are four possible combinations of time period and technology adaptation that are used to parameterize the crop, livestock and economic models. These are combined with climates according to Table to A2.1 to construct the simulations for analysis of the Four Core Questions.

Table A2.1. Technology and Climate Combinations Used in Analysis of the Four Core Questions

Core Question	System 1	System 2
1	τ_{HH}, γ_H	τ_{HH}, γ_F
2	τ_{HH}, γ_H	τ_{HF}, γ_H
3	τ_{FH}, γ_H	τ_{FH}, γ_F
4	τ_{FH}, γ_F	τ_{FF}, γ_F

Variable Definitions

t = individual year or time period

H = current time period

F = future time period

j = farm index, j = 1,…,J farms in data sample representing the integrated assessment region study area

t = 0 = base year(s) for the analysis, typically the year(s) when survey data were collected

τ_{ti} = technology and management used in period t = H or F, adapted to climate i = H or F

γ_t = climate in period t = H or F

p_t = representative output price (currency units/kg), t = H or F

y_{jt} = crop yield in year t (kg/ha)

$\mu_j(\tau_{ti}, \gamma_t)$ = time-averaged mean of yields for farm j using technology τ_{ti} with climate γ_t

Y_0 = mean of observed yields in the survey data for base year t = 0

Y_H = mean of yields averaged over all farms and years in the current period, obtained from secondary data in the study area

$\beta_{y0} = Y_H/Y_0$ = normalization factor used to scale survey data yields to the current period mean

$s_j(\tau_{ti}, \gamma_t)$ = simulated crop yield for farm j using technology τ_{ti} with climate γ_t

r_{jk} = relative yield for farm j used for Core Question k.

$r_{j1} = s_j(\tau_{HH}, \gamma_F) / s_j(\tau_{HH}, \gamma_H)$ = relative yield for analysis of Core Question 1

$r_{j2} = s_j(\tau_{HF}, \gamma_H) / s_j(\tau_{HH}, \gamma_H)$ = relative yield for analysis of Core Question 2

$r_{j3} = s_j(\tau_{FH}, \gamma_F) / s_j(\tau_{FH}, \gamma_H)$ = relative yield for analysis of Core Question 3

$r_{j4} = s_j(\tau_{FF}, \gamma_F) / s_j(\tau_{FH}, \gamma_F)$ = relative yield for analysis of Core Question 4

a_{jt} = total crop area on the farm in period t (ha)

R_{jt} = revenue = $p_t \cdot y_{jt} \cdot a_{jt}$ (currency units/farm/time)

R_{jqs} = time-averaged revenue for question q and system s (currency units/farm)

C_{jt} = production cost for period t (currency units/farm/time)

C_{jqs} = time-averaged production cost for question q and system s (currency units/farm)

C_t = mean of production cost averaged over all years in the current period (t = H), or the mean production cost for the base year (C_0) obtained from secondary data in the study area (if available)

$\beta_{c0} = C_H/C_0$ = normalization factor used to scale production cost survey data to the current period mean (note, If β_c can't be estimated, then use $\beta_{c0} = \beta_{y0}$ to assume that production costs from survey data deviates from what is representative for the current period and costs are normalized by the same factor as yields; or use $\beta_{c0} = 1$ when cost data is representative for the current period).

$G_{jt} = C_{jt}/R_{jt}$ = production cost relative to revenue (unit-free)

$G_{jqs} = C_{jqs}/R_{jqs}$ = time-averaged production cost relative to time-averaged revenue for question q and system s

$V_{jt} = R_{jt} – C_{jt}$ = crop net returns for the farm (currency units/time)

V_{jqs} = time-averaged net returns for question q and system s (currency units)

Bias12 = factor used to adjust RHO12 for bias (see discussion below).

The Relative Yield Model

We use both survey data and simulated data to represent the effects of climate change on productivity using the *relative yield model.* The idea behind this model is as follows: suppose we interpret system 2 as the current system being used under conditions of a future climate, and we interpret system 1 as the current system being used under conditions of the current climate. The average yield under climate change can then be related to the mean of the current system as $\mu_j(\tau_{HH}, \gamma_F)/ \mu_j(\tau_{HH}, \gamma_H) \equiv r_{j1}$ (this is the comparison used in Core Question 1). We define r_{j1} as the *relative yield under climate change*. We assume that we can approximate a yield impacted by climate change by estimating r_{j1} with crop model simulations as $r_{j1} = s_j(\tau_{HH}, \gamma_F) / s_j(\tau_{HH}, \gamma_H)$ where $s_j(\tau_{HH}, \gamma_F)$ is the time-averaged simulated yield for farm j under climate change, and $s_j(\tau_{HH}, \gamma_H)$ is the time-averaged simulated yield for farm j in the current period climate and technology. Then we project the yield with climate change and technology τ_{HH} as $\mu_j(\tau_{HH}, \gamma_F) = r_{j1} \cdot \mu_j(\tau_{HH}, \gamma_H)$ where $\mu_j(\tau_{HH}, \gamma_H)$ is the time-averaged yield for the current period. Since $\mu_j(\tau_{HH}, \gamma_H)$ is not observable in most cases, we approximate it with the observed yield from a farm survey in the current period for farm j, and scale the observed yields if necessary so that they represent the current period population mean.

Calculating the Between-System Correlation in the TOA-MD Model (RHO12)

The TOA-MD model requires an estimate of the correlation between the returns to each system (parameter RHO12 in the TOA-MD data sheet RHO). As noted above for Core Question 1, we estimate system 2 yields by assuming that $\mu_j(\tau_{HH}, \gamma_F) = r_{j1} \cdot \mu_j(\tau_{HH}, \gamma_H)$ where $\mu_j(\tau_{HH}, \gamma_H)$ is the mean observed yield from a farm survey in the current period for farm j. Note that we typically estimate $\mu_j(\tau_{HH}, \gamma_H)$ with the observed base year yield y_{j0} (adjusted by β_{y0} if necessary). We can write base year yield as $y_{j0} = \mu_j(\tau_{HH}, \gamma_H) + e_{j0}$, where $\mu_j(\tau_{HH}, \gamma_H)$ is the mean yield and e is a random component. The problem with the relative yield procedure for the calculation of RHO12 is that by correlating $\mu_j(\tau_{HH}, \gamma_F) = r_{j1}\, y_{j0}$ with y_{j0} we overestimate the correlation (note, the true RHO12 is the correlation between $\mu_j(\tau_{HH}, \gamma_F)$ and $\mu_j(\tau_{HH}, \gamma_H)$, but our procedure gives RHO12 equal to this correlation plus r_{j1} times the variance of e_{j0}). We can show that the bias that results is equal to Bias12 = $\mathrm{var}[\mu_j(\tau_{HH}, \gamma_H)] / \mathrm{var}(y_{j0})$. These variance components can be estimated with panel data using a fixed effects model. If panel data are not available, we suggest using Bias12 = 0.85 which is the approximately the value that has been obtained from several panel datasets.

Matched and Un-Matched Data

Two situations may be encountered with analysis using this type of farm survey data:

Matched Data*:* a crop yield can be simulated for each survey farm, for each crop in the system for which a crop model is available. This is true when weather and soil data can be associated with each survey farm, and some crop management data are included in the survey.

Data matching is possible in most cases where farm survey data are available and some kind of information is included in the survey to identify the survey farms' locations. Ideally, the spatial identifier is the farm's spatial coordinates (or even better, the centroids of individual fields). Note that when spatial coordinates are not included in a survey, they can be approximated with other location identifiers. For example, a legal address or village name may be available, and this may be used to approximate the spatial coordinates of the farm.

It is important to note that the matching of weather and soil data to survey farms will typically require using the *best approximation possible given available data*, because farm-specific weather and soils data are almost never available. Nevertheless, **as long as weather and soil data can be assigned to each survey farm through some reasonable procedure, the term "matched data" is used**, because with the farm specific management data, it is possible to simulate yields for each farm.

Un-Matched Data: a distinct crop yield ***cannot*** be simulated for each survey farm; however, spatially varying weather and soil data are available to run crop model simulations with representative management for the region.

Note that in the un-matched case, it is possible to estimate a simulated yield *distribution* that corresponds to the population of farms represented by the survey; however, it is not possible to match simulated yields to the survey farms.

Accounting for Future World Conditions: RAPs and Future Scenario Data from Global Economic Models

RAPs are used to represent future conditions, including productivity trends and effects of future economic conditions on output prices and costs of production. Regional RAPs must incorporate trends (e.g. yield trends from global econ models) following the methodology presented below, to translate current production systems into the future conditions defined by a RAP. If the analysis is linked to a global pathway and economic model scenario, data from that scenario (e.g., prices, productivity trends) should be linked to the regional RAP and scenario assumptions.

To parameterize the TOA-MD model to analyze the Core Questions, the analyst must construct parameters to reflect the effects of climate and adaptations on yields and costs, and also must adjust all other economic parameters to match the question conditions of current world (Questions 1 and 2) or a future world defined by the RAP (Questions 3 and 4). Note that for Question 1, only yields are adjusted for System 2 to quantify climate impacts under current world conditions. For Question 2, the analysis is implemented as a technology adoption analysis under current world conditions. Questions 3 and 4 are the same logical structure as Questions 1 and 2, but are implemented with economic data projected into the future world conditions.

The following parameters are used to project from current to future world conditions. They can be derived from model projections or RAPs as appropriate.

Γ = compounded yield growth factor between current and future periods. Used to estimate trended parameters of system 1 for Core Question 3 (e.g. use AgMIP Reference scenario data from IMPACT global model).

$\boldsymbol{\phi_t}$ = compounded price growth factor between current and future periods. Used to estimate trended output price parameters.

- $\boldsymbol{\phi_H}$ is the price growth factor **without climate change** and it is used to estimate parameters for system 1 for **Core Question 3** (e.g. use AgMIP Reference scenario data from IMPACT global model).

- **Φ_F** is the price growth factor **with climate change** and it is used to estimate parameters for **system 2 for Core Question 3** and for **system 1 and 2 parameters for Core Question 4.**

Ψ = compounded variable production cost growth factor between current and future periods. Used to estimate trended parameters of **system 1 for Core Question 3** and for **system 1 and 2 parameters for Core Question 4.** This factor should be defined as part of the RAPs.

Key Assumptions

A1: *The distribution of $\mu_j(\tau_{HH}, \gamma_H)$ (the true time-averaged mean of farm j in the current period) is approximated by the distribution of y_{jt} in the current year t in which the spatial yield distribution is observed.* This assumption allows us to use the observed yield in year t, scaled to the mean of the current period, as a proxy for $\mu_j(\tau_{HH}, \gamma_H)$. However, since we know that the observed yields for each farm will vary from the average in the current period, we know that the projected future yields include this variation. Thus, we need to take care in using data from the current period. The more years of data that can be used, the more we can average out the individual-year variation from the current period data, and doing so should result in better estimates of $\mu_j(\tau_{HH}, \gamma_H)$ and thus better projections of future yields.

A2: *For each Core Question, crop simulation biases are equal for each System.* For each technology and climate combination, we can define the bias in the crop model, e.g., let $b_{jH} = s_j(\tau_{HH}, \gamma_H)/\mu_j(\tau_{HH}, \gamma_H)$ for current period technology and climate. Now also define $b_{jF} = s_j(\tau_{HH}, \gamma_F)/\mu_j(\tau_{HH}, \gamma_F)$. If $b_{jH} = b_{jF}$, then it follows that

$$r_{j1} = s_j(\tau_{HH}, \gamma_F) / s_j(\tau_{HH}, \gamma_H) = b_{jF}\, \mu_j(\tau_{HH}, \gamma_F)/ b_{jH}\, \mu_j(\tau_{HH}, \gamma_H) = \mu_j(\tau_{HH}, \gamma_F)/\mu_j(\tau_{HH}, \gamma_H),$$

and thus $\mu_j(\tau_{HH}, \gamma_F) = \mu_j(\tau_{HH}, \gamma_H)\, r_{j1}$, proving that the relative yield provides an unbiased prediction of the System 2 mean yield.

A3: *$G_{jq1} = G_{jq2}$. The ratio of cost/revenue is the same for both systems in the analysis.* This assumption means that the profit margin is the same for the two systems being compared. This assumption provides a standardized way to project future cost based on current costs, or to project cost for an alternative system based on an observed system, but note that this assumption can be modified to fit a future situation where costs are expected to deviate from this relationship.

A4: *Yields in the integrated assessment region grow at compound rate Γ, and crop model simulations for the future period do not incorporate factors accounting for this growth between the current and future periods.* In the approach presented here, we assume that there is an independent yield growth factor associated with technological change that is not accounted for in crop model simulations.

A5: *Total land (Area in the TOA-MD model) allocated to the farming system in the population being modeled is constant within the current and within the future time period (but not necessarily the same between the two periods).* This assumption is based on the premise that data on area variation over time are not available within the current period, and are not modeled for the future period; alternatively, the analyst can use year-specific data if such information is available.

A2.2. Calculating TOA-MD Model Parameters

For Core Questions 1 and 2, the analysis is done assuming that the survey and other observational data represent the current world conditions of the analysis so set $\mathbf{\Gamma = 1, \phi_H = 1, \phi_F = 1, \Psi = 1.}$

Matched Data

Question 1

Step MA11: Calculate the relative yields r_{j1} for each farm j = 1,…,J in the survey.

Step MA12: Survey data observations of y_{j0} (base year) provide information to calculate the parameters for the historical period and historical technology (System 1):

$$\mu_j(\tau_{HH}, \gamma_H) = \beta_{y0} \cdot y_{j0}$$

$$R_{j11} = p_H \cdot a_{jH} \cdot \mu_j(\tau_{HH}, \gamma_H)$$

$$C_{j11} = \beta_{c0} \cdot C_{jH}$$

$$V_{j11} = R_{j11} - C_{j11}$$

Note: recall that p_H is a representative price, adjusted to the historical period average as necessary. β_{y0} *is the normalization factor used to adjust observed yields in the data to the historical period population average, and* β_{c0} is used to adjust observed costs to the historical average. *The historical period is defined as the five-year period centered as closely as possible on the year(s) of the economic survey data.*

Step MA13: calculate parameters with climate change for each farm in the survey data as follows:

$$\mu_j(\tau_{HH}, \gamma_F) = r_{j1} \cdot \mu_j(\tau_{HH}, \gamma_H)$$

$$R_{j12} = p_H \cdot a_{jH} \cdot \mu_j(\tau_{HH}, \gamma_F)$$

$$G_{j12} = C_{j11}/R_{j11}$$

$$C_{j12} = G_{j12} \cdot R_{j12}$$

$$V_{j12} = R_{j12} - C_{j12}$$

Step MA14: Using the data from MA12 and MA13, calculate the means for R_{j11}, C_{j11}, R_{j12} and C_{j12}, and the standard deviations of V_{j11} and V_{j12}.

Step MA15: Calculate RHO12 as the correlation between V_{j11} and V_{j12} times the bias factor Bias12. If this bias factor cannot be estimated, set it equal to 0.85.

Question 2

Step MA21: Calculate the relative yields r_{j2} for each farm j = 1,…,J in the survey.

Step MA22: Survey data observations of y_{j0} (base year) provide information to calculate the parameters for the historical period and historical technology (System 1):

$$\mu_j(\tau_{HH}, \gamma_H) = \beta_{y0} \cdot y_{j0}$$

$$R_{j21} = p_H \cdot a_{jH} \cdot \mu_j(\tau_{HH}, \gamma_H)$$

$$C_{j21} = \beta_{c0} \cdot C_{jH}$$

$$V_{j21} = R_{j21} - C_{j21}$$

Note: these are the same calculations as step MA12.

Step MA23: calculate parameters with adaptation for each farm in the survey data as follows:

$$\mu_j(\tau_{HF}, \gamma_H) = r_{j2} \cdot \mu_j(\tau_{HH}, \gamma_H)$$

$$R_{j22} = p_H \cdot a_{jH} \cdot \mu_j(\tau_{HF}, \gamma_H)$$

$$G_{j22} = C_{j21}/R_{j21}$$

$$C_{j22} = G_{j22} \cdot R_{j22}$$

$$V_{j22} = R_{j22} - C_{j22}$$

Step MA24: Using the data from MA22 and MA23, calculate the means for R_{j21}, C_{j21}, R_{j22} and C_{j22}, and the standard deviations of V_{j21} and V_{j22}.

Step MA25: Calculate RHO12 as the correlation between V_{j21} and V_{j22} times the bias factor Bias12. If this bias factor cannot be estimated, set it equal to 0.85.

Question 3

Step MA31: Calculate the relative yields r_{j3} for each farm $j = 1,\ldots,J$ in the survey.

Step MA32: Survey data observations of y_{j0} (base year), RAPs, and global economic models provide information to calculate the parameters for the future period without climate change.

$$\mu_j(\tau_{FH}, \gamma_H) = \Gamma \cdot \mu_j(\tau_{HH}, \gamma_H) = \Gamma \cdot \beta_{y0} \cdot y_{j0}$$

$$R_{j31} = \phi_H \cdot p_H \cdot a_{jF} \cdot \mu_j(\tau_{FH}, \gamma_H)$$

$$C_{j31} = \Psi \cdot \beta_{c0} \cdot C_{jH}$$

$$G_{j31} = C_{j31}/R_{j31},$$

$$V_{j31} = R_{j31} - C_{j31}$$

Step MA33: calculate parameters with climate change for each farm in the survey data as follows:

$$\mu_j(\tau_{FH}, \gamma_F) = r_{j3} \cdot \mu_j(\tau_{FH}, \gamma_H)$$

$R_{j32} = \phi_F \cdot p_H \cdot a_{jF} \cdot \mu_j(\tau_{FH}, \gamma_F)$

$G_{j32} = G_{j31}$

$C_{j32} = G_{j32} \cdot R_{j32}$

$V_{j32} = R_{j32} - C_{j32}$

Step MA34: Using the data from MA32 and MA33, calculate the means for R_{j31}, C_{j31}, R_{j32} and C_{j32}, and the standard deviations of V_{j31} and V_{j32}.

Step MA35: Calculate RHO12 as the correlation between V_{j31} and V_{j32} times the bias factor Bias12. If this bias factor cannot be estimated, set it equal to 0.85.

Question 4

Step MA41: Calculate the relative yields r_{j4} for each farm j = 1,…,J in the survey.

Step MA42: Survey data observations of y_{j0} (base year), RAPs, and global economic models provide information to calculate the parameters for the future period without climate change.

$\mu_j(\tau_{FH}, \gamma_F) = r_{j3} \cdot \mu_j(\tau_{FH}, \gamma_H)$

$R_{j41} = \phi_F \cdot p_H \cdot a_{jF} \cdot \mu_j(\tau_{FH}, \gamma_F)$

$G_{j41} = G_{j31}$

$C_{j41} = G_{j41} \cdot R_{j41}$

$V_{j41} = R_{j41} - C_{j41}$

Note: these are the same calculations as used for Question 3, System 2.

Step MA43: calculate parameters with climate change for each farm in the survey data as follows:

$\mu_j(\tau_{FF}, \gamma_F) = r_{j4} \cdot \mu_j(\tau_{FH}, \gamma_F)$

$R_{j42} = \phi_F \cdot p_H \cdot a_{jF} \cdot \mu_j(\tau_{FF}, \gamma_F)$

$G_{j42} = G_{j41}$

$C_{j42} = G_{j42} \cdot R_{j42}$

$V_{j42} = R_{j42} - C_{j42}$

Step MA44: Using the data from MA42 and MA43, calculate the means for R_{j41}, C_{j41}, R_{j42} and C_{j42}, and the standard deviations of V_{j41} and V_{j42}.

Step MA45: Calculate RHO12 as the correlation between V_{j41} and V_{j42} times the bias factor Bias12. If this bias factor cannot be estimated, set it equal to 0.85.

Multiple Activities

For systems with multiple activities, we apply the above calculations to each system. In addition, we need to estimate the within-system correlations between the returns to the activities. With matched data we can calculate the within-system correlations for system 2 the same way as for system 1 (i.e., by using the survey data to estimate the within-system average correlation between activities). For unmatched data, we typically assume that within-system correlations are the same for systems 1 and 2.

For trend calculations, yield trends for major crops from global models are used as the starting point, with adjustments to regional conditions as appropriate. Minor crop trends should be defined by the team based on the major crop trends. Livestock trends should be based on global model trends for milk and meat as appropriate, adjusted to regional conditions.

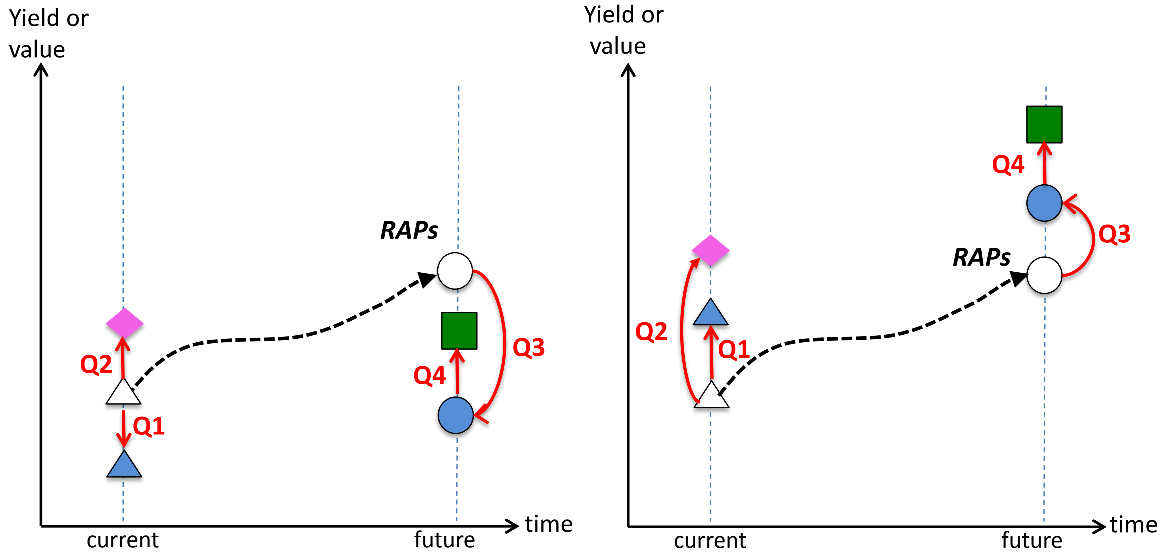


Figure A2.1: Overview of core climate impact questions and the production system states that will be simulated, as in Figure 1, but contrasting situations where climate change has a detrimental impact (left) with those in which climate change has a beneficial impact (right).

	System 1	System 2	Key Outputs
Question #1	Production system in Current Period with Current climate	Production system in Current Period with Future Climate	***% gainers & losers and net impacts*** % change in mean farm income % change in per capita income % change in Poverty rate
Question #2	Production system in Current Period with Current climate	Adapted Production system in Current Period	***Adoption rate (%)*** % change in mean farm income % change in per capita income % change in Poverty rate
Question #3	Production system in Future Period with Current climate **Productivity and price trends with no climate Change and RAPs**	Production system in Future Period with Future Climate **Price trends with climate Change and RAPs**	***% gainers & losers and net impacts*** % change in mean farm income % change in per capita income % change in Poverty rate
Question #4	Production system in Future Period with Future Climate **Price trends with climate Change and RAPs**	Adapted Production system in Future Period with Future Climate **Price trends with climate Change, RAPs and Adaptation Package**	***Adoption rate (%)*** % change in mean farm income % change in per capita income % change in Poverty rate

Figure A2.2: Overview of core climate impact questions and the production system states that will be simulated and key economic components and output indicators for TOA-MD simulation runs.

Appendix 3

User's Guide to Crop Model Simulations for Regional Integrated Assessments

K. J. Boote, C. Porter, C. Villalobos, J. Hargreaves, J. Antle, R. Valdivia, and J. W. Jones

December 3, 2012
revised June 11, 2015

Managing and Documenting Crop Model Inputs

The crop model simulation sets required to answer the 4 research questions for the Integrated Assessment are listed below and shown graphically in Figure A3.1.

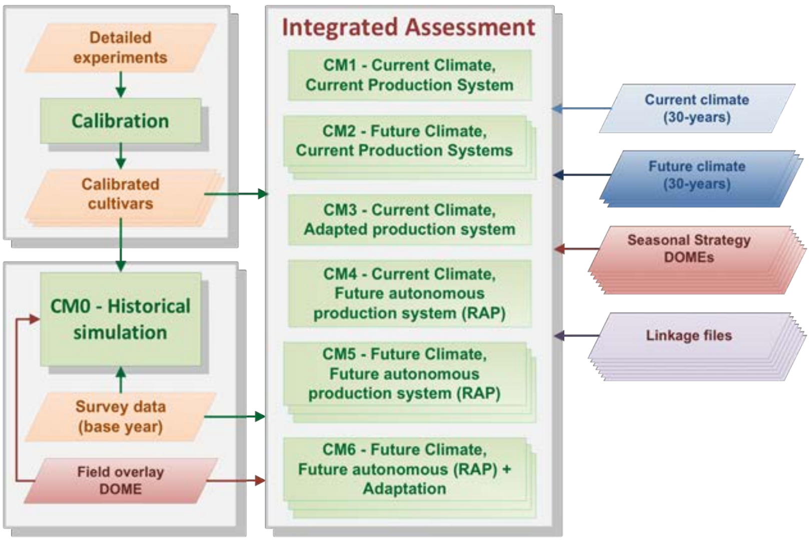


Figure A3.1. Crop model simulation sets required for Regional Integrated Assessments

- **Calibration.**
- **CM0 – Historical.** A simulation of the conditions under which the farm survey data were collected is typically performed for duration of one to two years and uses observed weather data for each site. The comparison of observed to simulated yields from the historical simulation allows researchers to validate the models and input parameters, and to compute biases and probability of exceedance. This is the only simulation for which comparison to observed yields is relevant.
- **CM1 – Current.** Simulation of the current climate and current production system uses 30 years of weather data based on current climatology.
- **CM2 – Future.** Simulation of future climate scenarios with the current production system.
- **CM3 – Current, adapted.** Simulation using current climate, but with a management system which is specifically designed for climate adaptation.
- **CM4 – Current, RAP.** Simulation using current climate, but with a management trend which includes production technology change corresponding to a particular RAP.
- **CM5 – Future, RAP.** Simulation using future climate scenarios, but with a management trend which includes production technology change corresponding to a particular RAP.
- **CM6 – Future, RAP, adapted.** Simulation using future climate scenarios, a management trend corresponding to a RAP, and with management changes which are specifically designed for climate adaptation.
- **Full GCM simulations.** Examine the full GCM ensemble for a single, best-calibrated and representative site in each integrated assessment region (these latter results will not be passed on to economic analysis) (not shown in Figure A3.1).

- **CTWN sensitivity test simulations** – single farm, 30 years; one at a time vary CO_2, Tmax/Tmin, rainfall, fertilizer N over a range for each variable.
    - **CO_2 –** 360, 450, 540, 630, 720ppm (run for high and low N) - 10 simulations
    - **Tmax/Tmin -** -2, 0, +2, +4, +6, +8 oC – 6 simulations
    - **Rainfall –** 25%, 50%, 75%, 100%, 125%, 150%, 175%, 200% - 8 simulations
    - **Fertilizer N –** 0, 30, 60, 90, 120, 150, 180, 210 kg/ha – 8 simulations
    - **Standard C3MP** at farm level N – 99 simulations

Each simulation is carried out through some combination of survey data, soil data, current and future weather data, assumed model inputs, and hypothetical management regimens. Data types used for these analyses are listed in Table A3.1 and described in the paragraphs below. All data are input to QuadUI, a data translation utility which provides the following functions:

(1) Translates the data to harmonized format, which can then be archived on the AgMIP Crop Site Database (https://data.agmip.org/).
(2) Translates the data to model-ready formats for multiple crop models
(3) Generates metadata which fully describe the simulation and data used to generate model input files

Table A3.1. Description of data files used by AgMIP IT tools to create multiple crop model input files.

Data type	File type	Description	File Formats
Raw data	Survey_Data	Observed field survey data for use in creating multiple model inputs. Survey data include experimental and management data in one file and soils data in a separate file.	Excel Spreadsheet, one line per field, which is exported to a zip archive (*.zip) containing comma-delimited (*.csv) files for import and translation
Raw data	Weather	Daily weather data for historical, current or future climate scenarios	Various formats including .AgMIP, csv, and DSSAT WTH files, compressed into a zip archive (*.zip) file
Raw data	Cultivar	Model-specific cultivar parameter files are passed by the translation utility to the model simulation directory.	Model-specific formats, in zip archive (*.zip)
DOME	Field_Overlay	Data and parameters needed by crop models, but which were not recorded in the field survey data	Excel Spreadsheet, which is exported to a zip archive (*.zip) containing comma-delimited (*.csv) files
DOME	Seasonal_Strategy	Used to set conditions for multi-year model simulation of current or alternative management practices for current or future weather scenarios.	Excel Spreadsheet, which is exported to a zip archive (*.zip) containing comma-delimited (*.csv) files
DOME	Rotation_Strategy	Used to set conditions for multi-year model simulation of crop rotations, having just one set of initial conditions at year 1 (under development)	Excel Spreadsheet, which is exported to a zip archive (*.zip) containing comma-delimited (*.csv) files
Linkage	Linkage	Used to assign one or more DOMEs to each entry in the farm survey data.	Comma delimited (csv)
ACMO	AgMIP Crop Model Output file	Summary of crop model simulation metadata and simulated results.	Comma delimited (csv)

Raw data include survey data, soil data, weather data and cultivar parameters. The survey data are measured at individual sites and stored in a Survey_data file, typically one line per site / season observation. Data include metadata regarding the site location; management data including planting, irrigation, fertilization and harvesting; and observations of crop growth and development, including harvested yield and dates of anthesis and harvest.

Microsoft Excel files are generally used to collate and organize the survey data and to convert units to conform to AgMIP standards. Table A3.2 lists the data that are typically provided in this file. Generally, household survey information includes crop yield, some management information, and economic data on a per farm-field basis.

Site-specific soil profile information is bundled with the farm survey data, in a separate worksheet, as shown in the survey data templates.

Weather data are stored separately to facilitate re-use of the survey data for multiple climate scenarios, including current climate conditions. These data can be entered in a spreadsheet using the ICASA notations, or supplied in .AgMIP format or DSSAT WTH files.

Model-specific cultivar parameters, from the calibration step, should be supplied with the raw data. These are not converted to harmonized format, but are passed through to the crop model simulation data directory in the formats required for each model.

Table A3.2. List of variables in the household survey data needed to run crop models. These data are in the Survey_Data.xlsx EXCEL file. Note that some of these data may not have been collected in the survey but are provided later by those who are preparing the data for translation by AgMIP software tools into model-specific input files needed to run each crop model.

Survey data variable	Units	ICASA Variable Name
Field/Farm name		EXNAME
Field overlay name(s)		FIELD_OVERLAY
Seasonal strategy name(s)		SEASONAL_STRATEGY
Latitude	dec. degrees	FL_LAT
Longitude	dec. degrees	FL_LONG
Weather station identifier to link to site information		WST_ID
Soil profile identifier		SOIL_ID
Planting date	yyyy-mm-dd	PDATE
Crop ID (see list of codes above)	code	CRID
Total seasonal N applied	kg[N]/ha	FEN_TOT
Manure/Organic matter applied	kg[DM]/ha	OMAMT
Harvest date	yyyy-mm-dd	HDATE
Harvest yield (dry wt)	kg[dry]/ha	HWAH
By-product removed at harvest as dry wt	kg[dry]/ha	BWAH
Indicates whether the field has been irrigated	Y or N	IRRIG
Notes (as desired, optional)		TR_NOTES
Survey data variable	Units	ICASA Variable Name

DOME data. Invariably, some required crop model inputs are not measured and must be assumed. Some crop models have internal assumptions that provide missing inputs but these are “hidden” from users, they vary across models, and they are not likely to be relevant for all regions where the models will be applied. In addition, some computer simulations make use of observed management, soil, and climate, but modify some of these factors to evaluate climate variability effects at a location, to assess impacts of future climate, and to evaluate hypothetical management options. The “Data Overlay for Multimodel Export”, or DOME, is a file type that is used by AgMIP tools to provide additional data used by each crop model to simulate crop growth and yield. Table A3.1 describes different types of DOME files currently implemented by AgMIP IT tools. All DOME functions are documented on the AgMIP research site at http://research.agmip.org/display/itwiki/The+DOME.

The Field_Overlay DOME is used to supply the needed inputs that are missing so that all of the models make use of the same regional or site-specific assumptions. For example, data collected in regional surveys usually do not have all inputs needed for crop model simulations. This concept ensures the integrity of observed values, clearly documenting assumptions made for simulation analyses, and it ensures consistency across crop models for multi-model applications.

A second type of DOME, the Seasonal_Strategy file, is used to provide information needed to create synthetic simulation experiments for multiple seasons of weather data. These files provide information for controlling simulations for multiple years, for example, and/or management practices, cultivars, and other inputs that define a particular hypothetical scenario that is to be simulated. These files can provide information to set up baseline management and climate simulations over multiple years, or they can be used to set up management associated with Representative Agricultural Pathways (RAPs) or climate change adaptation analyses. In these cases, the soil, climate, and management regimens in Seasonal_Analysis DOME files would override existing recorded management and replace those data with the prescribed regimen.

Linkage files are used to associate each entry in the survey data (farm site and season) with one or more DOMEs. The QuadUI utility reads the The Field_Overlay and Seasonal_Strategy DOME files are combined with archived survey data (Survey_Data files) and used by the data translators to produce model-ready crop model input files for multiple crop model.

ACMO files contain a select set of outputs from crop simulations, with metadata describing the simulation. The ACMOUI application is used to generate ACMO files.

Data templates for survey and DOME inputs are available for download from the AgMIP GitHub site (https://github.com/agmip/json-translation-samples). These templates contain headers which correspond to variables in the ICASA Master Variable list for which precise definitions and units are listed. Definitions and units are replicated in the templates as comments to help guide the user to the correct form of the input data. Templates can be extended to include additional survey data by consulting the complete list of ICASA variables at www.tinyurl.com/icasa-mvl . Dome functions can be added to the templates as needed. These functions are documented fully on the AgMIP research site at http://research.agmip.org/display/itwiki/The+DOME .

Software for AgMIP RIAs

QuadUI is a desktop utility that reads survey, cultivar, weather, DOME and linkage files and translates to model-ready formats. In addition to model input data, the utility produces aceb, dome and alnk files, ready for archiving in the AgMIP Crop Site Database. An ACMO metadata file is produced, which is used by ACMOUI to produce ACMO files.

ACMOUI combines the output files produced by crop models with the ACMO metadata created during the data translation phase by QuadUI and produces an ACMO.csv file. These files can then be archived on the Crop Site Database and are permanently linked to the survey data, DOMEs, weather data and cultivar files used to produce the outputs.

ADA is a Windows desktop utility which converts Microsoft Excel files into comma-delimited files (one per worksheet), zipped and ready for input to QuadUI.

Directory structure

The following list shows the recommended directory structure for a single crop in a single region. This pattern should be followed for each crop within the region and for each region analyzed. For each crop, data should be organized by the seven crop simulation data sets required in the Regional Integrated Assessment (labelled CM0 through CM6 below).

CM0-Historical

a) **Survey data** contains survey data plus soils data. Weather data are provided separately. There should be only one set of survey data which are used without modification for all analyses including future scenarios.
b) **Field overlay.** The data should be sufficient to allow simulation of historical conditions for multiple models. Typically, the field overlay DOME for historical conditions will be re-used without modification for all simulations. Additional field overlay DOMEs may be added for hypothetical management inputs for RAPs and adaptation packages.
c) **Linkage**

CM1-Current –This data set uses the survey data and field overlay of the Historical simulation.

a) **Seasonal Strategy**
b) **Linkage**

CM2-Future –This data set uses the survey data and field overlay of the Historical simulation. Sub-directories may be used for each climate scenario.

a) **Seasonal Strategy.** The Seasonal Strategy DOMEs used to simulate future climate conditions and current management should be the same as for simulation set CM1, except that the climate ID and the atmospheric CO_2 levels are specified for each climate scenario modeled. There will be one seasonal strategy file for each climate scenario
b) **Linkage.** A separate linkage file is needed for each climate scenario to connect survey data to the appropriate DOMEs.

CM3-Current, adapted –There should be one directory for each climate adapted management package (e.g., CM3-A1, CM3-A2, etc.). Adaptation packages for current climate conditions may differ from those for future climate scenarios. Modifications to the survey data for climate adapted management should be done through DOMEs.

a) **Field overlay (optional)** DOMEs may be needed to modify data originally provided in the survey data to impose management elements of the adaptation package. These could be used to indicate changes to soil properties or to use different cultivars. Separate soil data may need to be provided, but these should be given unique SOIL_IDs, separate from the original data. (For example, drought resistant cultivar traits have been simulated by using modified soil traits. In this case, the soil ID should be different than the original soil data.) Modified cultivar data should be included in the separate model-specific cultivar data directory with unique names.
b) **Seasonal Strategy (optional)** It may be possible to re-use the CM1 Seasonal strategy files, depending on the adaptation package modeled.
c) **Linkage**

CM4-Current, RAP –Multiple RAPs should be handled in separate directories (e.g., CM4-RAP1, CM4-RAP2, etc.).

a) **Field overlay (optional) -**
b) **Seasonal Strategy (needed)** It may be possible to re-use (modify) the CM1 Seasonal strategy files, updating for management depending on the RAP package modeled. Current climate.
c) **Linkage**

CM5-Future, RAP –Data relevant to each RAP scenario should be maintained in separate directories (e.g., CM5-RAP1, CM5-RAP2, etc.). Under each RAP directory, multiple climate scenarios may be stored in separate folders.

a) **Field overlay (optional)**
b) **Seasonal Strategy (needed)** Must use the same as the CM4 Seasonal strategy file, which specifies management depending on the RAP package modeled. But using future climate.

c) **Linkage**

CM6-Future, RAP, adapted –Data relevant to each RAP / Adaptation scenario should be maintained in a separate directory (e.g., CM6-RAP1-A1, CM5-RAP2-A2, etc.). Under each RAP directory, multiple climate scenarios may be stored in separate folders.

a) **Field overlay (optional)**

b) **Seasonal Strategy (needed)** Start with the CM5 Seasonal strategy file, which specifies management depending on the RAP package modeled, but modified to a climate-adaptation. Uses future climate.

c) **Linkage**

Weather – All weather data should be put in a separate weather directory. Simulation data sets CM0, CM1, CM3 and CM4 share the current climate conditions weather data. (The exception to this rule is when the surveyed data year falls outside the 1980 – 2010 range of the current climate weather data and the historical simulation data set will have a separate weather file.) Each weather data file should contain the climate ID. Sub-directories may be used to separate climate scenario data if many weather stations are used. Note that QuadUI accepts climate data in comma delimited format (csv), .agmip format and DSSAT WTH format; data must be in zip archive regardless of format provided.

Cultivar – Model-specific cultivar data files should be put in a cultivars.zip file with an internal directory structure which reflects each appropriate model, as shown in the WinZip example in Figure A3.2. DSSAT cultivars must be put in a folder “dssat_specific” and APSIM cultivars must be put in a folder “apsim_specific”.

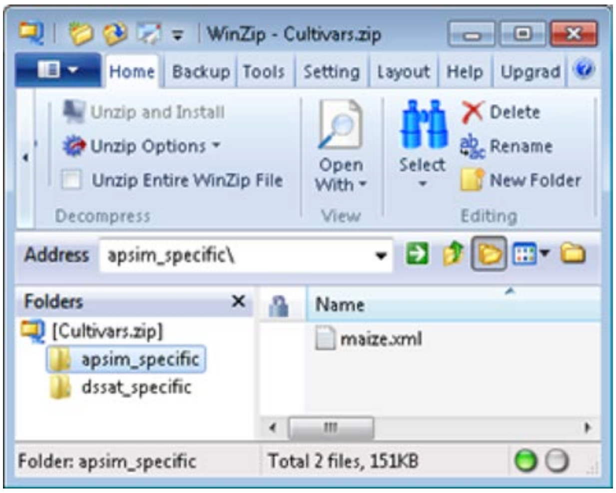


Figure A3.2. Organization of model-specific cultivar data in a zip file.

File naming conventions

In order to keep track of the many different management and climate scenarios modeled, the following file naming conventions should be used so that each data file fully describes its contents and the correct file can be chosen for each translation and simulation.

<u>Crop modeling data:</u>

File_type-Region-Crop-ClimID-RAP_ID-MgmtID.ext

<u>Examples:</u>

Survey_data-REG1-MAZ.xlsx (measured field conditions do not include climate or management id)

Field_Overlay-REG1-MAZ.xlsx

Seasonal_strategy-REG1-MAZ-4IFA-0-0.xlsx (current management)

Seasonal_strategy-REG1-MAZ-4IFA-0-A2.zip (adaptation scenario)

Seasonal_strategy-REG1-MAZ-4IFA-R1-A1.zip (RAP and adaptation scenario)

<u>ACMO File Naming Convention:</u>

ACMO-*Region-Crop-ClimID-RAP_ID-MgmtID-CropModel*.csv

<u>Example:</u>

ACMO-REG1-PNT-4IFA-R2-A2-DSSAT.csv

These files are automatically named by ACMOUI.

<u>TOA-MD File Naming Convention:</u>

TOAin-*Region-ResearchQuestion(1,2,3 or4)-ClimID-RAP_ID-MgmtID-CropModel*.xlsx

<u>Examples:</u>

TOAin-REG1-1-4IFA-0-A1-DSSAT.xlsx

TOAout-REG1-1-4IFA-R1-A1-DSSAT.xlsx

These must be named manually, and should match the ACMO name.

Metadata

The final product of the crop simulations are the ACMO files. These files will be archived in the Crop Site Database and made available for download or for use in analysis and visualization in the AgMIP Impacts Explorer. Complete metadata to describe each simulation must be included in the ACMO files and these metadata are passed through from DOME files. These metadata are particularly important to identify the climate ID for all climate scenarios and the management ID for the adaptation packages. The Climate ID will be assigned in accordance with the Climate Team protocols and should match the names of the daily weather files generated by the Climate Team. The Man_ID metadata variable must be used to distinguish between current management and adaptation packages. The Region, Man_ID and RAP_ID values must be co-developed with the Economic modeling team such that crop modeling metadata and filenames are associated with the corresponding TOA-MD metadata and filenames. Table A3.3 lists metadata associated with each DOME file.

Table A3.3. Metadata included in DOME "INFO" section:

Metadata	Sample value	Definition
REG_ID	REG1	Region name
STRATUM	2	Assigned by econ modeling teams
RAP_ID	R1	Code for RAP being modeled (leave blank if no RAP)
MAN_ID		Code for climate adaptation package being modeled (leave blank if no adaptation package)
RAP_VER		Version code for RAP ID (leave blank if no version)
CLIM_ID	IKFA	Climate ID for scenario being modeled
DESCRIPTION	P1	Short descriptive text for this DOME file (important if there are multiple DOMEs for this scenario)

The DOME name is derived from the values of metadata provided. In this case, the DOME name used in the linkage file would be "REG1-2-R1---IKFA-P1", which is the concatenation of all metadata fields, separated by hyphens. Because of this DOME naming convention, it is important that hyphens are not used in the metadata values (i.e., "P1", not "P-1").

Procedures for Creating Crop Model-Ready Input Files for Survey Fields

Start with generating data for the historical simulation (CM0) which is the simplest case and uses the survey data and a field overlay, but no seasonal strategy DOME. An iterative procedure is usually required to get the correct format and units for the survey data and sufficient field overlay information to produce reliable simulations for multiple crop models.

A crop model simulation "roadmap" can help track which files are used for each simulation set. An example is provided in Table A3.4. In this case the base survey data and field overlay DOME are used for every simulation, without modification. Weather data are supplied based on the climate scenario being modeled. Each simulation, except the historical simulation, requires a seasonal strategy DOME to generate multi-year simulations. Each simulation requires a linkage file to link the survey data to the appropriate DOMEs. The table also lists the associated folder in which the file resides, so that the crop modeler can easily find the file when running QuadUI for data translation.

Additional field overlay DOMEs can be used to describe management imposed by a RAP or an adaptation package. In this example, additional field overlay DOMEs were used for the CM3 adapted management for current climate conditions, the future technology management associated with a RAP (CM4, CM5 and CM6), and with the future climate adaptations (CM6).

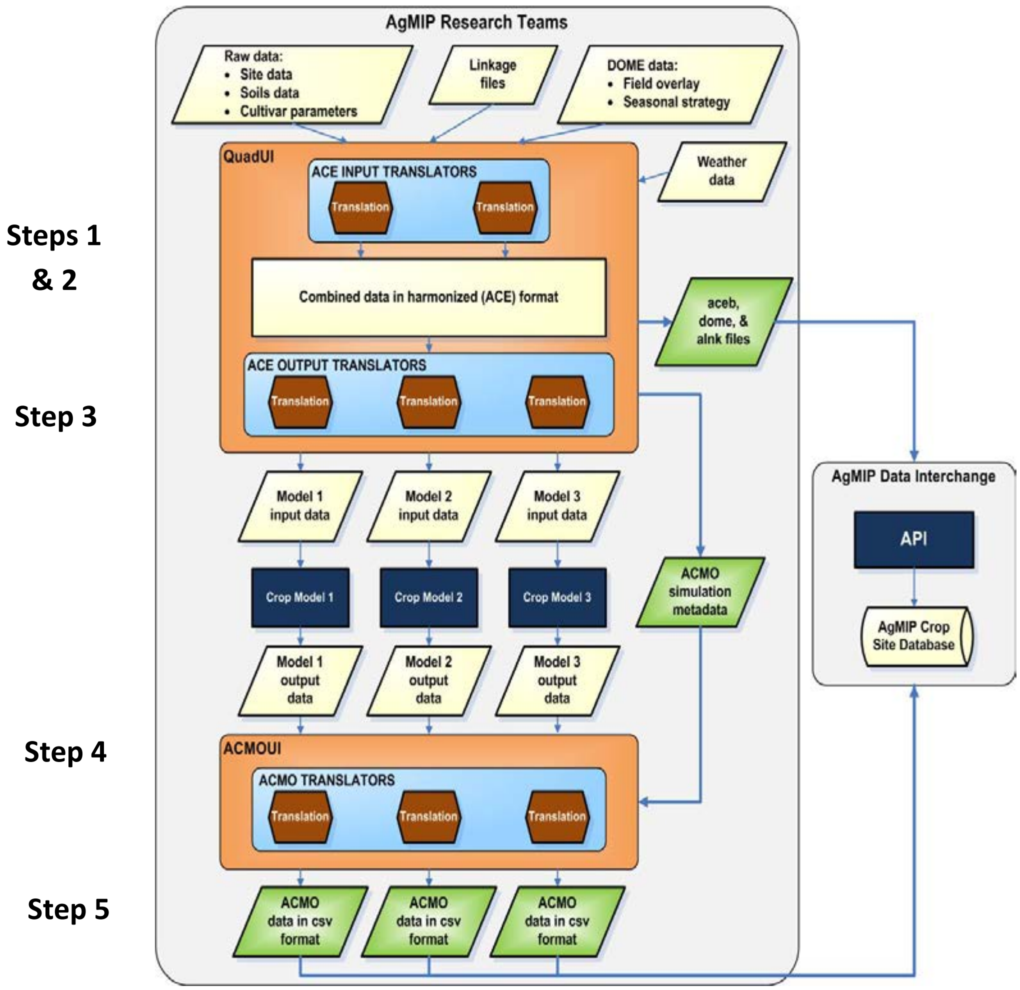
Table A3.4 Sample "roadmap" of files used in Crop Modeling analyses. The survey data and field overlay files are used in all simulations.

File Name	**CM0 Historical**	**CM1 Current**	**CM2 Future**	**CM3 Current, Adapted**	**CM4 Current, RAP**	**CM5 Future, RAP**	**CM6 Future, RAP, Adapted**
Survey_data-Region-MAZ.zip	X	X	X	X	X	X	X
Weather-Region-0XFX.zip	X	X		X	X		
Weather-Region-IxFA.zip			X			X	X
Field_Overlay-Region-MAZ.zip	X	X	X	X	X	X	X
Field_Overlay-Region-MAZ-0XFX-0-A0.zip (optional)				(X)			
Field_Overlay-Region-MAZ-0-R1-0.zip (optional)					(X)	(X)	(X)
Field_Overlay-Region-MAZ-X-0-R1-A1.zip (optional)							(X)
Seasonal_strategy-Region-MAZ-X-0-0.zip		X	X	(X)	(X)	(X)	(X)
Seasonal_strategy-Region-MAZ-0XFX-0-A0.zip (optional)				(X)			
Seasonal_strategy-Region-MAZ-X-R1-0.zip (optional)					(X)	(X)	(X)
Seasonal_strategy-Region-MAZ-X-R1-A1.zip (optional)							(X)
Linkage-Region-MAZ-historical.csv	X						
Linkage-Region-MAZ-X-0-0.csv		X	X				
Linkage-Region-MAZ-0XFX-0-A0.zip				X			
Linkage-Region-MAZ-X-R1-0.zip					X	X	
Linkage-Region-MAZ-X-R1-A1.zip							X
Climate_Batch.csv		X	X		X	X	X

Notes: For current climate conditions, CLIM_ID = "0XFX" (for this sample).
For future climate scenarios, CLIM_ID is represented generically as "IxFA", where the "x" represents the GCM used for the analysis.
Red highlight indicates files that are repeated for multiple Climate scenarios / GCMs.
Adaptation scenarios are Identified by "Ax", which represents the ID of the adaptation package.
All file names use the convention: *File_type-Region-Crop-ClimID-RAP_ID-MgmtID.ext*

Figure A3.3 presents the workflow for producing a single simulation dataset for the AgMIP Regional Integrated Assessment. The steps correspond to the more detailed descriptions below. In summary, raw data, weather data, linkage files and DOME files are used as inputs to QuadUI, which translates the data, first to ACE format, then to model-ready formats for multiple models. Model simulations are done manually. ACMOUI is run to gather crop model outputs and generate harmonized ACMO files, using the ACMO metadata file created by QuadUI.

Figure A3.3. Schematic data flow diagram for AgMIP RIA Crop modeling data translation using QuadUI and ACMO UI



applications .

Step 1. Gather, assemble and enter data (survey and expert)

- Download data translation tools from http://tools.agmip.org/
    - QuadUI – desktop application for data translation
    - ADA – converts from Excel to csv format for import to QuadUI
    - ACMO_UI – converts model output to ACMO format
    - Sample spreadsheet templates for survey data and DOME data ICASA Variables List– list of variables to extend the survey data template, if needed (http://tinyurl.com/ICASA-MVL)
- Enter survey data into one of the survey data templates, Additional columns can be added to the survey data import template for those data. Note that dates are entered using ISO compliant format: YYYY-MM-DD. Note also the units

for all variables. Conversions can be done in the spreadsheet, and unneeded data "commented out" as shown in the template files.

- If some data are missing, one or more Field Overlay templates should be used to FILL in the missing data (examples are dates of N fertilization or manure application). There can be multiple field overlays, if soils and soil initial conditions vary across farms.
- Visit with Soil Scientist experts from the region: Find the appropriate soil for each farm (linking to latitude-longitude or village information), and enter the soils information by soil layer in the soil tab in the Survey_Data file. The soil name is also listed in the field section of the Survey_Data file.

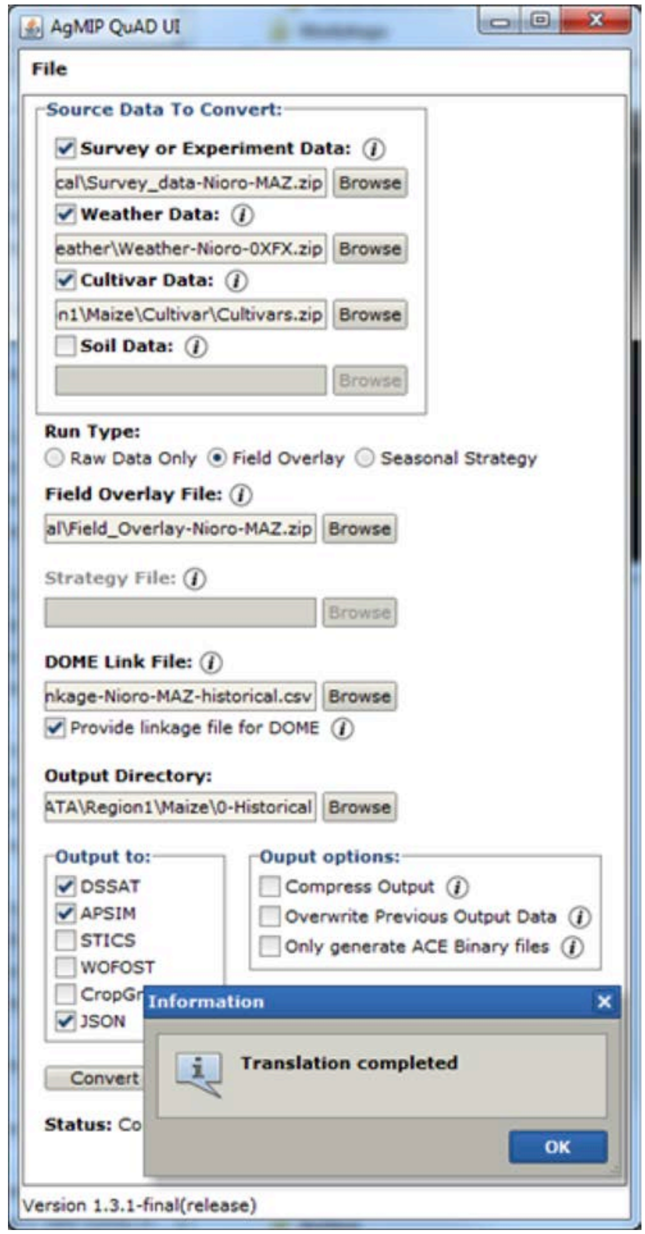


Figure A3.4. QuadUI screenshot showing selection of raw data, DOME data, Linkage file, Output directory and Model formats.

Step 2. Save Survey_Data and Field_Overlay Data to csv format

- Using the ADA utility, save Survey_Data, and field overlay sheets in comma delimited (csv) format. Caution: Do not open the *.csv files again with Excel, as they ARE NOT true spreadsheets and do not correctly convert back into the correct date formats.

Step 3. Translate data files to model-ready formats

- Run QuadUI by double-clicking on the QuadUI.bat file. Respond to the on-screen requests for location of the following data as depicted in Figure A3.3.
    - survey data (zipped csv),
    - weather data (zipped csv, .AgMIP or WTH files),
    - cultivar data (optional, zipped model-specific files),
    - soil data (optional, zipped csv),
    - field overlay DOMEs (optional, zipped csv),
    - seasonal strategy DOMEs (optional, zipped csv),
    - DOME linkage files (csv, not zipped)
    - Output file location (optional)
- QuadUI will generate files for running crop models, i.e., Files X, A, SOIL.SOL, *.CUL, *.WTH for DSSAT, and .APSIM and met files for APSIM. In the case of DSSAT or APSIM, simulations can be run by double-clicking the DOS batch file that is created with the translations.
- Excel files; the date format can revert to a machine-dependent format.

Step 4. Check and correct missing/invalid model input data and run simulations

- Run the crop model.
- Troubleshooting
    - DSSAT: Look at the Error.OUT and the Warning.OUT files.
    - APSIM: Load the simulation and view the log. Also review the *.sum files.
    - Look for missing climate or cultivar files found,
    - Look for missing data such as sowing date or plant population. Typically this means that these were not supplied in the DOME or that the linkage file does not correctly link the field overlay to the experiment or field.

    - Revise the Survey_Data and Field_Overlay files as needed.
- Evaluate the outputs. In DSSAT, look at the Evaluate.Out file which will list both the simulated and the observed yield. In APSIM, there is a single line output for each simulation. The APSIM-simulated yield values will need to be aggregated (assembled) into one file. The observed yields are in the Survey_Data file and will need to be matched per field.

Step 5. Create AgMIP Crop Model Output (ACMO) File for use by Economic Team Members

The ACMO file is partially created by QuadUI at translation time in the form of the ACMO_meta.dat file which contains metadata for all of the simulations. Running ACMOUI, a desktop utility, will complete the ACMO file with the selected crop model simulated outputs.

Note that the ACMO files contain raw simulated results for each field, not aggregated or adjusted in any way. This will ensure integrity of both inputs and model outputs.

Create Crop Model-Ready Input Files for Simulating Multi-Year Current Climate, Future Climate, RAPs and Adaptation Scenarios (Using Seasonal strategy DOMEs)

Notes on Use of Seasonal_Strategy Files

A Seasonal_Strategy DOME file allows the single year survey data to be used for multi-year simulations for current and future climate scenarios, both with and without RAPs and Adaptation Packages. Examples of DOME functions for seasonal strategy are:

- Auto-sowing rules,
- Links to future scenario Climate IDs,
- Translating RAPS into management using the Seasonal Strategy (RAPs can led to improved crop and soil management practices including improved genetic technology). Specifics include:
    - Auto-sowing, possibly modified for earlier/shorter sowing window because of better machinery
    - changed plant population,
    - improved or alternative crop cultivar,
    - changed N fertilization,
    - increased prior root and surface residue (because of better fertilization-population-cultivar)
    - other adaptation strategies, as needed

To run a Seasonal Strategy, repeat all of the steps for baseline analysis, using the Survey_Data and Field_Overlay files, but now the Seasonal_Strategy DOME files are also prepared using the spreadsheet templates.

Notes on Use of Field_Overlay Files

- Function and Purpose of multiple Field_Overlay files
    - Fill in data that crop models need that are almost never available in household survey, such as initial soil water, initial soil nitrate and ammonium, soil organic carbon pools (SOM3 for DSSAT-CENTURY, and inert SOC for APSIM), and rooting depth (see Table A3.4). The shape of these listed variables by soil depth is also generally missing.
    - Fill in needed data missing from household survey, such as root residue from prior crop, surface residue from prior crop, sowing date, sowing depth, plant population, amounts and dates of fertilizer or manure applied.
    - Link to cultivar ID and model specific cultivar ID
    - It can be used to set automatic sowing rules for each field in the survey, if planting dates were not recorded.
- Where to get Field_Overlay information? First, DO NOT use crop model defaults, as the model defaults are wrong and differ among crop models. Often defaults use zero or unity values when not appropriate and these are not region-specific. Secondly, this must be done in close collaboration with local agronomists and soil scientists.

- Agronomists and soil scientists in the region who know production practices for the crop and region in question. Although they may not give specific values for the needed inputs, they will likely provide very useful ranges of likely missing input information.
- Soil survey information (linking to latitude-longitude coordinates for field).
- Country-wide statistics (amount of N fertilization per hectare), but this is not region- or crop-specific.
- Pre-simulations with crop models with correct soil organic carbon and SOM3 (or inert SOC) pools, for setting SOM3 and inert SOC to mimic the low non-fertilized non-legume yields for the region (requires knowledge of unfertilized yield for region). Take the mineral nitrate and ammonium from the values simulated at the end of the "prior" season.
- Make sure that the assumed values that you use in the Field_Overlay file are consistent with all of the expert knowledge and soil survey information, and document how these values were developed.

Guidelines for Analysis of Crop Model Simulated Outputs for Matched Fields

Crop modelers should analyze model outputs prior to use of the data in the regional economic analysis. This is very important to ensure quality control of the process and that crop modelers are able to understand the variability in results. It is also important that crop modelers will be able to conclude that simulated yields are reasonable representations of water and nitrogen-limited yields, recognizing that other factors, such as other soil nutrients and pests, are likely to contribute to actual yields in a region and that these factors could vary considerably over space and time. We have provided suggestions for analyzing crop model outputs, including computation of means, distribution of observed and simulated yields, computation of mean bias between observed and simulated yields, and analysis of outliers.

- Evaluate Simulated and Observed Yields for Mean, Bias, and Distribution
- Place simulated yield and observed yields into a spreadsheet, computing means and standard deviation. Compute bias of the mean observed yield divided by mean simulated yield. We do not recommend computing bias of individual fields if there are any zero simulated yield values, as that will give error.
- Rank the observed yields and simulated yields from high to low and compute cumulative probability distributions of observed and simulated yields.
- Attempt to identify outliers and reasons for high mean bias as well as large differences between cumulative distributions of simulated and observed yields. These analyses may help crop modelers critically evaluate some of the input assumptions in the Field_Overlay file, for example, relative to the information from regional agronomists and other sources that were used to set the values. If there is a large bias, it would be good to review the inputs and results with agronomists. Be cautious in types of calibration for reducing the bias and base this on knowledge of the soils, initial conditions, and cultivars used. This is intended to improve the reliability of the process and results. These analyses may be useful in reporting and in publishing actual crop model results, although the economists will only be using change ratios described earlier. Some ideas to consider as you analyze results are:
    - If bias (observed over simulated) is dramatically different from 1.00 (for example 0.5 or 1.5), there may be problems in Field_Overlay assumptions. Bias is driven by the mean simulated and observed yields. For example, a high bias of 1.5 or more (model simulates low) could indicate that soil N availability (SOM3, initial nitrate, initial ammonium) or soil water availability (initial or capacity) is not high enough. A low bias of 0.5 (model simulates too high) could indicate too much soil N availability or too much water availability.
    - The full range of the cumulative distribution is driven not just by the management and climate, but also by the extent of range of initial nitrate, ammonium, SOC, SOM, DUL-LL, and initial soil water found across all the farms. If that range of inputs (and soil variability) is small (because of inadequate Field_Overlay entry), then the simulated distribution of yields could be insufficient.
    - Strong left tails in simulated distribution (or observed) are indicators of crop failures (zero and very low yields). If left tails is too strong in simulated, then you may need to increase initial soil water content to reduce the instance of simulated germination failures, or increase rooting depth or DUL-LL to minimize crop failures during reproductive growth.

- Strong right tails in simulated or observed distributions are indicators of high yields. If simulated right tails are too strong (or too little) where the water and N stresses are minimum, one can make the case that genetic yield potential of the cultivar is too high (or too low). Farmers' cultivars are often not as good as those used in research experiments.

These "indicator" problems are given, not for the purpose of re-calibrating the crop models to fit the distribution, but for the purpose of highlighting the need for obtaining correct Field_Overlay information in the first place.

Appendix 4

Fast-Track Activities to Demonstrate Integrated Framework

Because of the coordination needed among different science disciplines in the AgMIP regional integrated assessment efforts, each new AgMIP regional team should perform a “proof of concept” assessment on a fast track to help everyone on the regional teams to understand their roles and the interactions that must take place among different disciplines. Accomplishing this will ensure that the mechanics of the process are understood and functioning, at which point it will be easier for all teams to proceed with their further, more detailed assessments.

To do the fast track integrated assessment exercise, the team should select only one sub-region, one crop, one crop model, and one climate site location; then simulate crop yields using the historical climate data for that one location and also simulate crop yields for one climate change scenario for the time period of 2040 – 2069 using the methods described above. Additional details are:

a. The entire regional team should identify one small sub-region where the fast track assessment will be performed. Ideally, the sub-region should be an area in which household survey data are available with at least one climate data site within the area and where there are experimental data available in or nearby the area that can be used for calibrating one (or more) crop models.
b. The crop modelers will parameterize the crop models using available data from experiments, if this has not already been done. This will provide parameters for cultivar types that are currently being used in the region.
c. The economists should describe the site characteristics, including a map showing the farms and including management and farm characteristics.
d. Economists will provide the socioeconomic data, including farm site locations, to the crop modelers so that they can assemble the needed crop model inputs to run the crop models. Ideally, the socioeconomic survey data would have data on crop management practices (planting date, N application amounts) and on crop yield. For example, there may have been 80 farms surveyed with such data, and those farms would be used to assemble crop model input data for each farm, similar to the Machakos example that has been used by AgMIP to demonstrate the approach.
e. The climate team members in the region will prepare and quality-control the historical climate series for one station in the region. This site will act as the baseline climate series for all crop modeling and analysis in the fast-track (including surrounding farms), and will also serve as the basis of one climate change scenario generated using the basic delta method that represents projected GCM changes. These climate series may be used in the crop model runs to compute the impacts of climate change (assuming no adaptation for this fast track).
f. The regional crop modelers will prepare input files for running one selected crop model (DSSAT or APSIM preferably) for each farm location in the selected study site/area. This includes assembling representative soils for the sites. The crop modelers will simulate each of the fields in the farm surveys, analyze simulated results relative to observed yields to evaluate reliability of results, and prepare a model output file ACMO) for documenting model inputs and outputs for use by economists in the TOA-MD analyses.
g. If socioeconomic data do not include farm site yields, then the crop modeling team members will use the procedures for calibrating and evaluating crop models for use in simulating mean yields for district or other administrative unit (see section 5c in this handbook). This alternate procedure will provide crop models ready for use in the region with estimates of average bias.

h. The crop modelers will then simulate yields for each of the farm sites in the selected area using historical climate data (1980-2009 planting years) and repeat the simulations using the one selected climate scenario's climate file. The modelers will assess yield results, evaluating how reasonable they are and produce an AgMIP Crop Model Output file (ACMO) that will be used by the economists in the TOA-MD analysis.
i. The economic team members will take crop model results and use the TOA-MD model to analyze the impacts of the climate change scenario on the distribution of economic impacts for the area.
j. The entire team will meet to evaluate the entire process and to discuss and interpret the results.
k. After the proof of concept study, the team will be ready to design its assessments of impacts and adaptation options based on the RAPs, more advanced climate scenarios, and a better representation of climate and crop model uncertainties.