CIMMYT – ORNL Summit
20– 22 May 2014

Location: *CIMMYT Headquarters Offices located at Km. 45, Carretera México-Veracruz El Batán, Texcoco CP 56237, Edo. de México. Mexico*

Meeting report

Background

The International Maize and Wheat Improvement Center (commonly called by its Spanish acronym CIMMYT for Centro Internacional de Mejoramiento de Maíz y Trigo) is a non-profit research and training institution dedicated to both the development of improved varieties of wheat and maize, and introducing improved agricultural practices to farmers, thereby improving their livelihoods. It is also one of the first of the 15 non-profit, research and training institutions affiliated with the Consultative Group on International Agricultural Research (CGIAR). CIMMYT focuses on the conservation and utilization of maize and wheat genetic resources, developing and promoting improved maize and wheat varieties, testing and sharing sustainable farming systems, and analyzing the impact of its work and researching ways for further improvement. CIMMYT partners with national agriculture research institutions across the globe.

Oak Ridge National Laboratory (ORNL) is the largest multipurpose science laboratory in the U.S. Department of Energy's national laboratory system. The mission of ORNL is to deliver scientific discoveries and technical breakthroughs that will accelerate the development and deployment of solutions to meet pressing global challenges, namely energy, environment, and national security, aligned with the DOE's goals. This "science-to-solutions" mission depends on the integration and application of distinctive capabilities in basic and applied research, which include leadership capabilities in biological, environmental, and climate change sciences, high-performance computing (HPC), advanced data curation, integration, analysis, and visualization, knowledge management, spatial data mining, geographic data modeling, computational science and systems engineering.

Since 2011, CIMMYT and ORNL have been interacting within a collaborative framework, the Knowledge Systems for Sustainability (KSS) community of practice, supported by a number of leading national and international public, private, and non-governmental organizations focused on the broad issues of agricultural sustainability and food security. In 2013, CIMMYT leadership visited ORNL for review and discussion of ongoing initiatives and scientific capabilities that are relevant for CIMMYT's research and development focus areas and an agreement to explore a formal cooperation was established. This summit marks the realization of this commitment between CIMMYT and ORNL with the intention to mobilize cutting edge frontier science and technology toward this century's most urgent threats and opportunities focused on food, water and energy security.

Table of Contents

Progress Reports and Trajectories of CIMMYT/ORNL Team Tasks Prescribed by the WFO (Work for Others) MOU Approved by US DOE

Task 1. Development and implementation of data and knowledge management platform

Working group members

Budhendra Bhaduri (ORNL)*	Jesús Herrera (CIMMYT)	Horacio Rodriguez (CIMMYT)
Richard Fulss (CIMMYT)*	Darell Sison (CIMMYT)	Andrea Gardeazabal (CIMMYT)
Stan Wood (BMGF)	Nick Davis (CIMMYT)	Philippa Zamora (CIMMYT)

* = Group leader

Vision: The Knowledge Platform is a single, structured, open-access portal to a federated, collection of harmonized and prioritized CIMMYT and partner data and application resources delivering strategic and operational knowledge products and services of direct relevance to targeted user needs**.**

Scope: The project will be scaled from a CSISA pilot in Year 1, to CIMMYT-wide adoption by Year 3 to CGIAR level adoption after that.
The project will be developed through the following stages:

- Characterizing, indexing and generating metadata for CIMMYT's data landscape by defining core data sets and their accessibility (internal / external, open / private).
- Designing the knowledge platform through a process of analyzing user requirements, reviewing relevant tools and platforms, connecting new and legacy databases in a structures that mesh with project needs, and providing the desired functionality to users, including requirements from other systems projects like MasAgro and SIMLESA.
- Investigate how to best integrate the knowledge platform with current activities and incentivize people to use the system, share data and enhance their science.
- Determine the model and system requirements regarding the building, hosting, operations and interface with public and private users. This includes both the design of the system and considerations for short-term and long-term funding.
- Develop guidance document to implement data management plan and integrate with the Knowledge Discovery Framework. This document would help to guide the data management plan into project proposals from their conception and be able to scale to the CG level for use by other scientists. The document would also incorporate input from donors related to requirements for data sharing.
- Monitoring and evaluation protocols would be developed that define metrics of success and how data would be collected to evaluate those metrics.
- Linkages with other working groups, such as the Data Mining and Data Transformation and Landscape Scale Crop Assessment tool (LCAT) will be explored.

In the initial phase, which will be completed by the end of 2014, the project team will develop, design and implement a systems level prototype for the knowledge discovery platform. In order to meet this deadline, a project team will be assembled and concept of operations will be

developed in coordination with the CGIAR. The scope of the pilot will be determined, including timeframes, milestones and deliverables. The team will conduct an audit of existing data and systems, where interoperability and interactions are highlighted and create a readiness assessment from this audit process. User requirements will be analyzed and their needs incorporated into the system design.

CSISA will be the test use case for implementation of the knowledge discovery platform. The use case will include the development of a spiral implementation plant, platform requirement analysis, creation of user profiles, identification and incorporating data from other ready and available data assets from across the CIMMYT portfolio, determination of the set of functionality and features that will be implemented and designing feedback loops and adoption strategies for the targeted users.

The full report from the Task 1 working group can be downloaded here: https://uwmadison.box.com/s/6eevbsnls8708c9ui5he

The overview of the Task 1 collaboration can be downloaded here: https://uwmadison.box.com/s/3rqzub3j69ys1hp99kw0

Task 2. Development of novel data and information products (Landscape Scale Crop Assessment Tool)

Working group members

Suresh K.S. Vannan (ORNL)*	Urs Schulthess (CIMMYT)	Juan Arista (CIMMYT)
Bruno Gerard (CIMMYT) representing Andy McDonald*	Kate Schneider (BMGF)	* = Group leader

The goal of the Landscape Crop Assessment Tool (LCAT) is to develop a program that will compile and build a comprehensive data inventory needed for landscape crop assessment. These data inventories will be built on existing data holdings such as the MODIS subsetting tool but other remotely sensed data products. The tool will extend existing data mining algorithms to build a crop assessment capability for croplands and will generate the following:

- Agricultural land classification (Using ASTER data)
- Planting/Flowering/Peak of Canopy Dates
- Maturity Dates
- NDVI Anomalies
- Cumulative GPP
- Surface soil moisture

Modeled after ForWarn, the satellite-based change recognition and tracking tool to which ORNL has contributed, this tool (provisionally called “CropWarn”) will provide a comprehensive crop analysis/early warning summary for each field site that provides data, visualizations, and analysis techniques.

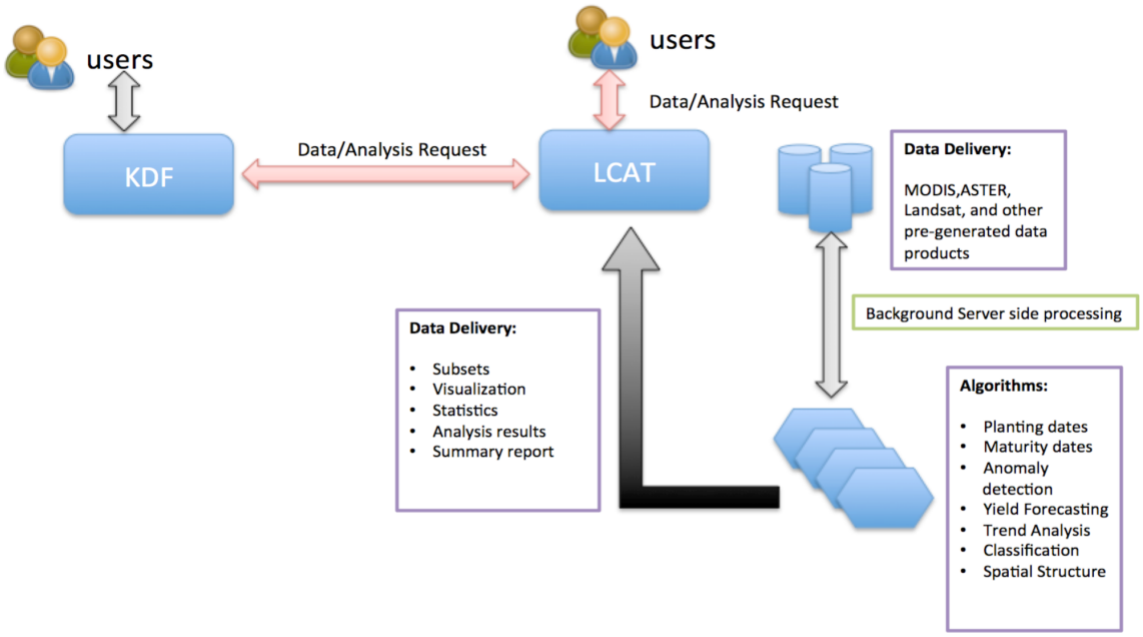
In a pilot activity, the LCAT will access and analyze large amounts of data for South Asia and CSISA sites of interest. The pilot will provide a decision support tool for better understanding of spatio-temporal variability and for technology targeting. It will also deliver an assessment of historical and current crop condition information. Key functions of this pilot activity will include:

- Repackaging existing data and tools to meet customer needs
- Exploratory research on new algorithms
- Provide the analysis summary and algorithms for crop assessment
- Provide ground-truthing for validation
- Explore funding opportunities to maintain/explore the system beyond the build period
- Establish interoperability with relevant projects (e.g. STARS and GeoGLAM)
- Explore connections with other tasks (e.g. Knowledge Discovery Platform)

Proposed Timelines:

In Year 1, the Core Tool Feature package will be designed, built and tested based on the MODIS data stream and Year 2 activities will focus on companion data integration into the Agricultural Knowledge Discovery Framework and tool applications.

Figure 1: System diagram of the LCAT



The full report from the Task 2 working group can be downloaded here:
https://uwmadison.box.com/s/dbfnrxebk2vssxb87szm

The overview of the Task 2 collaboration can be downloaded here:
https://uwmadison.box.com/s/krswaxun4p1pfdn9b8bn

Task 3. Incorporation of remote sensing data dissemination tools

Work products for this task will progress separate from the meeting.

Task 4. Development and integration of scalable data analysis and application tools

Working group members

Raju Vatsavai (ORNL)*	Gordon Huestis (CIMMYT)	Peter Wenzl (CIMMYT)
Kevin Pixley (CIMMYT)*	Xuecai Zhang (CIMMYT)	Sukhwinder Singh (CIMMYT)
Kate Dreher (CIMMYT)	Charles Chen (CIMMYT)	Sarah Hearne (CIMMYT)
Nele Verhulst (CIMMYT)	Tom Payne (CIMMYT)	Mike Olsen (CIMMYT)
Denise Costich (CIMMYT)	Zakaria Kehel (CIMMYT)	Jens Riis (CIMMYT)
Rosemary Shrestha (CIMMYT)		* = Group leader

Scope:

CIMMYT has generated and holds very detailed complex data (from the germplasm to the landscape). But tools are needed to extract the knowledge held within these critical data assets that can be used to better design breeding programs and larger research efforts. Potential solutions include specific data infrastructure that can be developed and introduced to specific data-mining strategies that can be applied to existing repositories and databases.

Detailed scope:

1. Data management

CIMMYT seeks a comprehensive approach to data management and the means to implement a data management plan, including the conversion of formats and accessibility of data in repositories. Products envisioned from this activity would include:

- Standards for data
- Data curation
- Metadata curation
- Database organization and construction
- Data ontology (mapping)
- Repositories (knowledge base, index based on metadata)
- Strategy to adopt these new tools and methodologies

The output of the data management tasks would be promoted to CIMMYT and CRP researchers and available to partners within and outside the CGIAR. The US Department of Energy Systems Biology Knowledgebase (KBase) is one potential platform that can assist in meeting the above stated goals.

Next steps for this task include: (1) a follow-up virtual discussion in June 2014 to be scheduled by ORNL Project Manager, Sue Heinz; (2) compilation of a draft Terms of Reference and audit of CIMMYT data management by October 2014; (3) evaluation and recommendation of existing standards and development of a strategy and action/work plan by December 2014.

2. Extracting value from CIMMYT data

Collaborative work that enhances the value of currently held CIMMYT data would include:

- Assess the capabilities and develop a plan to sequence data for core collections of accessions and elite germplasm wheat and maize
    - Determine how the ORNL user facility can be leveraged to assist in this task.
- Development of tools for data transformation (imputation, reformatting, BLUPs, etc.), including hardware and software needs to accelerate these transformations.
- Development of algorithms for data mining in collaboration with the Joint Genome Institute at ORNL for haplotypes or sequences associated with yield, stress tolerance, heterosis and responses to management practices.
- New data products/knowledge, for example, Identification of environments that are representative of Punjab in 2050 and identification of haplotypes and accessions with novel diversity for biofuel production.

Next steps for this task include the identification of the core team from CIMMYT and ORNL to work on the products listed above and refine the scope and objectives of these products by June 2014. A high-level data inventory will be compiled by August 2014 and representatives from CIMMYT propose to visit ORNL to learn more about their computational infrastructure and user facilities in September 2014. A work plan that identifies case studies will also be completed by September 2014. Possible cases include a) Soil fauna DNA influenced by management practices; b) Wheat international nurseries; and/or c) Assessing maize wild relatives / wheat landraces for traits such as perenniality or biomass production.

3. Adding value to CIMMYT data

Several mechanisms to leverage the vast collections of data at CIMMYT include:

- Customized search engine and user interface that allows users to go across the web and find available resources.
- Query tool that enables users to combine and mine information from those resources.
- Visualization tools.
- CIMMYT data visible and accessible to others (via open access pathways), with integrated use metrics collected
- An architectural scheme that shows interrelationships between various data mining and information retrieval subsystems. Example components include: large scale NoSQL storage, distributed web crawler, meta data harvesting and indexing, domain ontologies and data integration/fusion, data discovery and entity recognition, visual analytics, link discovery, correlation miner, annotator, etc.

The above products could be developed in coordination with partners such as the National Center for Biotechnology Information and iPlant.

Next steps for this task include finalizing the team and definition of scope by June 2014, identification of commonly used external resources by July 2014, identification of schema of resources to query and desired outputs including visualization tools by September 2014 and planning a meeting between CIMMYT and ORNL colleagues to discuss the customization opportunities of existing tools. These activities would be used to develop a proposal / plan for action by October 2014.

The full report from the Task 4 working group can be downloaded here:
https://uwmadison.box.com/s/jqn24fiewsyb1pn0fugl

The overview of the Task 4 collaboration can be downloaded here:
https://uwmadison.box.com/s/q1rjkk60igmfzzrvwd6k

Task 5. Sensors networks and measurement tools

Working group members

Ivan Ortiz-Monasterio (CIMMYT)*	Francelino Rodrigues (CIMMYT)	Samuel Trachsel (CIMMYT)
Yarom Polsky (ORNL)*	Urs Schulthess (CIMMYT)	David LeZaks (UW - Madison)
Maria Tattaris (CIMMYT)	Bruno Gerard (CIMMYT)	* = Group leader

Scope:
The goal of the collaboration will be to advance the following objectives:
1. Develop sensors to facilitate crop management and phenotyping for farmers and researchers
2. Improve on existing sensors and methods to reduce manual operations and increase measurement throughput

Sensor networks promise to transform and expand environmental science. However, many technical challenges must be overcome to achieve this potential. Technological issues include: innovation in new sensors and instrumentation system design; incorporation of power optimization schemes, power harvesting, or passive sensing approaches; integration of appropriate communication architectures and protocols; streamlining data management and access; use of innovative visualization and analytics; and enabling both scientists and the public access to results. Multi-disciplinary partnerships are making major contributions in measurement and sensing technology that informs improved modeling and simulation efforts for better preparation, use and reset of agriculture resources.

Specific sensor data to fulfill CIMMYT needs include:

Crop management
- Phosphorous measurement in plant or soil (available vs. total)
- Water content of plant and soil (applies to phenotyping as well)
- pH meter (lower priority, only if can leverage existing ORNL technology immediately)
- N, K spot measurements (inorganic N, nitrate)

Phenotyping
- Root depth and structure
- Multisensor packaging (e.g. NDVI, canopy temp, water content, etc.)
- Spatially distributed, time continuous light interception measurement w/wireless communication, data logging

- High throughput gas exchange measurement

Improvements to existing methods

- Automation of hyperspectral imaging workflow and feature extraction
- Better spike photosynthesis measurements
- Cheaper, more accurate GHG emission from soil measurements (potentially demonstrate ORNL tomographic CO2 measurement at CIMMYT site)
- Cost-effective temperature, precipitation and solar flux weather stations

In order to meet the needs of farmers and researchers, an approach that prioritizes future efforts based on the priority of the needed measurement and the confidence in the measurement strategy will be developed. This approach will also include the development of user specification and performance requirements, development of concept design, including user interface, data logging and transfer, and the development, testing and eventual deployment of the sensor that meets programmatic and budgetary requirements.

The full report from the Task 5 working group can be downloaded here:
https://uwmadison.box.com/s/rys8k7rx33mrlykd88ib

The overview of the Task 5 collaboration can be downloaded here:
https://uwmadison.box.com/s/3y8oriecm729gm6qsy65

Task 6. Identification and evaluation of ecosystem service indicators

Working group members

Carolina Camacho (CIMMYT)* representing Bram Govaerts	Victor Hernández (CIMMYT)	Matthew Thornton (CIMMYT)
Virginia Dale (ORNL)*	Rachael Cox (CIMMYT)	Victor López (CIMMYT)
Hans Braun (CIMMYT)	Nele Verhulst (CIMMYT)	* = Group leader

Scope:
Develop and implement a joint approach for sustainability assessment of farming systems that:

- adapts to decision making at different scales (farmers, farm advisors, policy makers and other stakeholders)
- recognizes which indicators are within our control and which are not

Detailed scope:
The development and implementation of indicators that are useful for assessing both socioeconomic and environmental sustainability of agricultural and food systems will be developed using the following steps:

- Elaborate a framework for assessing sustainability that considers approaches already developed by CIMMYT and ORNL and includes gathering input from stakeholders

- Select a case study, or a network of case studies, and perform a primary assessment that includes identification of data availability and gaps, scale of the assessment, boundary conditions and important stakeholders
- Identify potential indicators or attributes of the system and assess existing data assets and data gaps
- Develop and deploy capacity building initiatives to collect information for the sustainability assessment
- Collect and assimilate data from different sources
- Define incentives, best practices and alternative pathways for increasing biophysical, social and economic sustainability
- Develop a process for ongoing assessment and evaluation with stakeholders

Products and deliverables:
The process detailed above will result in the generation of sustainability assessment tools, including tools for identifying trade-offs and for incentivizing change and promoting best practices that result in improvements across a range of sustainability indicators. The tools developed will be broad enough to be applied over a range of geographies yet able to be tailored to specific locales or researchable questions. Inherent in the process is the identification and filling of data gaps (where necessary) and improvement in data management tools. The sum of these products and services will improve the capacity for CIMMYT to perform sustainability assessments and field evaluations internally and in collaboration with partner organizations.

Access and management of data and models is a key component of this process. Quantitative and qualitative data from both environmental and social sources will need to be collected, processed, stored and used in a way that supports the goals of the task. Data for these types of analyses will be derived from primary observation (e.g. field measurement, interviews), in-situ sensor and remote sensing platforms, modeling and the compilation of existing data. Data management systems will be needed to manage these diverse sets of data to enable better, faster science to be completed and more informed management decisions to be made.

The full report from the Task 6 working group can be downloaded here:
https://uwmadison.box.com/s/22akberho6464tuxque8

The overview of the Task 6 collaboration can be downloaded here:
https://uwmadison.box.com/s/1pwshpq5tql6fwsrheoa

Task 7. Coupling human and natural systems for landscape analysis

Work products for this task will progress separate from the meeting.

Task 8. Program coordination and management

An overall umbrella contract to allow facile movement of project-focused resources will be concluded in the next 2 weeks. CIMMYT and ORNL leadership will join CSIRO, IIASA, Columbia Earth Institute and U Oxford colleagues at a meeting sponsored by the Rockefeller Foundation at the Bellagio Conference Center in late August, 2014 to solidify the trilateral and multilateral

interactions and explore additional administrative and organizational collaborative models that can include committed private sector partners. Thomson Reuters will donate access their Eikon platform to researchers at CIMMYT and ORNL, as well as other collaborators. Those with access to Eikon will attend a meeting hosted by the Boston University Pardee Center for Longer Range Futures from 5-7 November, 2014 and led by our colleague, Tony Janetos. The meeting will be held in collaboration with GEOGLAM and AgMIP and will focus specifically on building methodologies and high quality approaches to estimating probabilistic risk and uncertainty related to the possible failure of multiple global breadbasket regions within a single annual cycle. The CIMMYT/ORNL alliance with other principal collaborating organizations will partner with a network of counterparts in the financial sector to announce an effort at the UN Climate Summit on September 23, 2014 on encoding disaster risk in finance; this effort will include systemic financial risks that could originate in food system failures. On September 24, 2014, also in NYC, collaborators will participate in the Future Earth Summit. CIMMYT colleagues who are interested in becoming better acquainted with our collaborators at the Earth Institute, Columbia University and AgMIP at NASA GISS, are invited to join us in New York City on these dates. In addition, a number of collaborators may gather at the annual AgMIP meeting in Long Beach, CA in late October 2014. Dialogues in both the G7 and G20 are advancing, thus showcasing this alliance and its intersections with GEOGLAM.

General Presentations

Slide decks from the meeting can be found here: https://uwmadison.box.com/ornl-cimmyt

Common Themes Across Tasks

While each task was designed to respond to specific programmatic objectives, it became evident that there are common needs that can be addressed by a single task and deployed across multiple research areas. Examples of these common themes are highlighted below.

Data management:
Design and development of data structures, open access data systems, cyberinfrastructures, metadata templates, ontologies, organization of field data collection tools, genetic and genomic data management, privacy and security functions, incorporation of legacy datasets into modern systems.

Integrating data from novel sources:
Digital collection and storage of quantitative and qualitative field data from survey mechanisms, new sensor technology and improved protocols for existing sensors for in-situ measurement of biophysical crop variables (aboveground and below ground), capture, processing and use of near-real-time remotely sensed imagery.

Modeling:
State-of-the-science crop, climate, earth-system and integrated modeling systems that run the gamut from genetic – microbial community – plot – landscape – global scale and incorporate socioeconomic and biophysical modeling within specific projects.

Creation of learning systems:
At each intervention (e.g. field site or landscape) system variables need to be captured that describe the size, location, production system, management history, social context and other relevant variables. It is possible to capture both "fast" and "slow" variables such as yield and price (fast) and soil health and equity (slow) that also describe the system. The outcome of each intervention can be synthesized in a way that informs future science and policy that is easily accessible.

Additional Areas for Collaboration

1. Measurement and related analytics of greenhouse gas measurements from microbial communities in agricultural soils

2. Opportunities to strengthen crop-modeling activities at CIMMYT.
    a. There are existing activities in the socioeconomics groups
    b. Peter Craufurd (CIMMYT), Strategic Leader Sustainable Intensification in Africa, Global Conservation Agriculture Program is actively engaged in AgMIP

3. Investigate the potential for *Zea perennis* and *Tripsicum dactyloides* as bioenergy crops.
    a. Similar genetics to varieties already used for bioenergy production, therefore similar genetic tools might be useful.

4. Idaho National Laboratory has developed a Bioenergy Landscape Design Tool and Residue Removal Assessment Framework that might be of interest to CIMMYT researchers
    a. More information available at: http://bioenergyldt.inl.gov/

5. Interdisciplinary assessment of atmospheric dispersal of crop disease spores
    a. Historical modeling of distribution of previously documented disease dispersal patterns
        i. Example of impacts of climate change on the dispersal of yellow rust in US Great Plains, North Africa, Southern Turkey
    b. Impacts of climate change on the disease dynamics and dispersion of crop diseases

Overview of Complementary Partners and Activities

CSIRO

Mike Grundy, CSIRO Sustainable Agriculture Flagship, attended the meeting and shared details about several of their projects that are complementary to the work being proposed by CIMMYT / ORNL teams. In his remarks, he stressed the need for integration across multiple systems such that specific initiatives will be designed to not only provide input to the challenge at hand, but

also create "learning systems" where the results of one study, tool development or intervention can be applied to the subsequent interventions at scale. Mike illustrated many of the projects and initiatives that CSIRO is leading or contributing to, and pointed out how the research community can come together to apply these tools and develop new ways of working together to address 21st century challenges.

Slides from this presentation can be found here:
https://uwmadison.box.com/s/mcth2ac15ldunsaarw0r

GEOGLAM

Inbal Becker- Reshef, University of Maryland Center for Global Agricultural Monitoring Research and the GEO Global Agricultural Monitoring Initiative (GEOGLAM), provided updates on the many products that she and her colleagues are producing, including production forecasts, crop condition monitoring, crop type mapping, land cover change mapping, and capacity building for smallholders. Many of the examples that are being developed are designed to answer specific research objectives, but are also being delivered operationally to stakeholders and decision-makers. Many of the GEOGLAM products are tested locally before being rolled out globally and across many commodities. For instance, the Crop Condition Monitoring System is currently being deployed to Australia, Mexico, Argentina and Brazil. Many of these products can be integrated into the CIMMYT – ORNL research product tasks.

Slides from this presentation can be found here:
https://uwmadison.box.com/s/77fvdgjapnkc5va7sgkp

AgMIP

The Agricultural Model Intercomparison and Improvement Project was not explicitly represented at this meeting, but there are many potential overlaps in the subject areas covered at this meeting and AgMIP's research. AgMIP's work scales from simulations at very small (plot) scale to global scale integrated analyses. They have a specific focus on the intersections between climate change, agriculture and economics and focus regionally in South Asia and Sub-Saharan Africa.

More Information about AgMIP can be found at http://www.agmip.org/

Next Steps

Task leaders and members of the working groups from each team will continue to refine their project ideas, scope and potential funding opportunities. ORNL has organized an internal all-hands meeting on 17 June to present the details of the CIMMYT – ORNL Summit to colleagues. Representatives from CIMMYT and ORNL are working to complete the remaining technical details related to the Work for Others agreement, as soon as possible.

Appendices

Appendix 1: Meeting Agenda

Monday 19 May 2014 – Oak Ridge National Lab Leadership and Scientific Team leads arrive
Tuesday 20 May 2014

<table>
<tr><th>Time</th><th>Topic</th></tr>
<tr><td>8:30-8:35
Sasakawa Room</td><td>Welcoming remarks
Tom Lumpkin & Martin Keller</td></tr>
<tr><td>8:40-9:00
Sasakawa Room</td><td>Overview and vision for the collaboration
Molly Jahn</td></tr>
<tr><td>9:00-9:15
Sasakawa Room</td><td>Knowledge Systems for Sustainability a Collaborative Effort using Big Data for Improved Agricultural Systems
Tom Lumpkin</td></tr>
<tr><td>9:15-9:30
Sasakawa Room</td><td>Knowledge Systems for Sustainability a Collaborative Effort using Big Data for Improved Agricultural Systems
Martin Keller</td></tr>
<tr><td>9:30-10:30
Sasakawa Room</td><td>Overview of key CIMMYT projects and research areas
<table>
<tr><td>• Global Maize Program (GMP)
Felix San Vicente</td><td>• Global Wheat Program (GWP)
Hans Braun</td></tr>
<tr><td>• Global Conservation Agriculture Program (GCAP)
Bruno Gerard</td><td>• Socioeconomics Program (SEP)
Olaf Erenstein</td></tr>
<tr><td>• Genetic Resources Program
Kevin Pixley</td><td></td></tr>
</table></td></tr>
<tr><td>10:30-10:45</td><td>Break</td></tr>
<tr><td>10:45-11:45
Sasakawa Room</td><td>Overview of key ORNL projects and research areas
<table>
<tr><td>• Environmental Intelligence for Global Sustainability (EnvInt)
Jay Gulledge</td><td>• Knowledge Discovery Framework (KDF)
Budhendra Bhaduri</td></tr>
<tr><td>• Environmental Data Science Systems (Novel data products and tools)
Raju Vatsavai and Suresh SanthanaVannan</td><td>• Sensors and Controls Research
Yarom Polsky</td></tr>
<tr><td>• Environmental and Socioeconomic Indicators for Sustainability
Virginia Dale</td><td>• Impact, Adaption and Vulnerability Research Science
Jay Gulledge</td></tr>
</table></td></tr>
<tr><td>11:45-12:30
Sasakawa Room</td><td>Overview of the ORNL – CIMMYT collaborations
Bram Govaerts
• Review of WFO themes
• Introduction of team leads
• Review of meeting goals and dynamics
o Task 1- Knowledge Platform Development & CIMMYT open access & data in the frame of the CGIAR
o Task 2- Landscape Scale Crop Assessment Tool
o Task 3- Incorporation of remote sensing data dissemination tools
o Task 4- Data mining and data transformation
o Task 5- Environmental sensors
o Task 6- Environmental and socioeconomic indicators for sustainability</td></tr>
</table>

	○ Task 7- Coupling human and natural systems for landscape analysis ○ Task 8- Program coordination & strategic planning • Comments from CIMMYT scientists (Bruno Gerard, Richard Fulss & Kevin Pixley)
12:30-1:30 *CIMMYT's Guesthouse*	**Lunch**
1:30-2:00 *CIMMYT's Genebank*	**Visit to Wellhausen-Anderson Plant Genetic Resources Center** *Kevin Pixley*
2:00-3:00 *D5 Experimental Platform and Machinery station*	**Long Term Sustainability Trial (Take it to The Farmer)** *Bram Govaerts*
3:00-3:45 *Sasakawa Room*	**Joint 5-minute presentation by each pair of technical/scientific teams that reviews progress and sets a vision for the collaborative work** • Knowledge Platform Development & CIMMYT open access & data in the frame of the CGIAR (Task 1) *Budhendra Bhaduri & Richard Fulss* • Landscape Scale Crop Assessment Tool (Task 2) *Suresh SanthanaVannan & Bruno Gerard* • Data mining and data transformation (Task 4) *Raju Vatsavai & Kevin Pixley* • Environmental sensors (Task 5) *Yarom Polsky & Iván Ortiz-Monasterio* • Environmental and socioeconomic indicators for sustainability (Task 6) *Virginia Dale & Carolina Camacho*
3:45-5:15	**Joint technical/scientific teams meet focused on WFO priorities**
Sasakawa Room	• Knowledge Platform Development & CIMMYT open access & data in the frame of the CGIAR (Task 1) ***Discussion Leaders:*** *Budhendra Bhaduri & Richard Fulss*
Room B113	• Landscape Scale Crop Assessment Tool (Task 2) ***Discussion Leaders:*** *Suresh SanthanaVannan & Bruno Gerard*
Room B114	• Data mining and data transformation (Task 4) ***Discussion Leaders:*** *Raju Vatsavai & Kevin Pixley*
Room B116	• Environmental sensors (Task 5) ***Discussion Leaders:*** *Yarom Polsky & Iván Ortiz-Monasterio*
Wheat Conference Room	• Environmental and socioeconomic indicators for sustainability (Task 6) ***Discussion Leaders:*** *Virginia Dale & Carolina Camacho*
3:45 – 5:15 *Board Room*	Program coordination & strategic planning *Thomas Lumpkin, Jay Gulledge, Martin Keller, Mike Grundy, Molly Jahn and Bram Govaerts*
5:15 – 5:30	Break
5:30 *Meeting Point: Main Building's lobby*	Transfer to the restaurant
6:30 *Restaurant in Teotihuacán Pyramids*	**Dinner with social event close to the pyramids**

Wednesday 21 May 2014

Time	Topic
8:30	**Meeting convenes**
8:30-12:00	**Joint technical/scientific teams meet focused on WFO priorities**
Sasakawa Room	• Knowledge Platform Development & CIMMYT open access & data in the frame of the CGIAR (Task 1) ***Discussion Leaders:*** *Budhendra Bhaduri & Richard Fulss*
Room B113	• Landscape Scale Crop Assessment Tool (Task 2) ***Discussion Leaders:*** *Suresh SanthanaVannan & Bruno Gerard*
Room B114	• Data mining and data transformation (Task 4) ***Discussion Leaders:*** *Raju Vatsavai & Kevin Pixley*
Room B116	• Environmental sensors (Task 5) ***Discussion Leaders:*** *Yarom Polsky & Iván Ortiz-Monasterio*
Wheat Conference Room	• Environmental and socioeconomic indicators for sustainability (Task 6) ***Discussion Leaders:*** *Virginia Dale & Carolina Camacho*
12:00-1:00	**Lunch**
1:00- 2:30 *Sasakawa Room*	**Report back from break-out groups and discussion** (Plenary)
2:30 – 3:00 *Sasakawa Room*	**Discussion about the Multiple Breadbasket Failure initiative** *Molly Jahn* Development of a critical route towards clear goals and outcomes • Data mining CIMMYT's vast holdings of agricultural and climate data • ORNL data, modeling and knowledge management • Collaborations with GEOGLAM and AgMIP
3:00- 3:15	**Break**
3:15- 4:15 *Sasakawa Room*	**Overview of CSIRO activities** *Mike Grundy* • Computational Sciences Knowledge and Technology Exchange • Opportunities for landscape science • CSIRO – ORNL MOU review and next steps
4:15 – 4:45 *Sasakawa Room*	**Building the Knowledge Systems for Sustainability Co. (Plenary - Molly Jahn)** • Report back from program coordination and strategic planning sessions • Partners, their roles and the structure of the organization • Enabling a global collaboration platform to support research and operational capacity • Discussion
4:45- 5:15 *Sasakawa Room*	**Concluding comments** *Tom Lumpkin and Martin Keller*
5:15- 6:00 *Sasakawa Room*	**Perspectives from the broader community** *Stan Wood*
6:00- 6:30	**Break**
6:30 *CIMMYT's Guesthouse*	**Signing Ceremony and Dinner**

Thursday 22 May 2014

Time	Topic
8:30	**Oak Ridge National Lab Leadership and Scientific Team leads departs** *A voluntary social program is foreseen*

Appendix 2: Task List

Task 1. Development and implementation of data and knowledge management platform

This task will focus on developing an enterprise wide or project specific data and knowledge management platform to increase the efficiencies of various project activities and facilitate compliance with the principles of open data initiative. The overall goal is to create a capability beyond a simple data warehouse or web-mapping application to include data integration for analysis and not just overlay. Utilizing the example of the Bioenergy Knowledge Discovery Framework (Bioenergy KDF; http://Bioenergykdf.net), developed for the Bioenergy Technology Office of the U.S. Department of Energy, the proposed knowledge discovery platform for CIMMYT will provide a shared environment for collaboration and facilitates informed decision making by providing a means to synthesize, analyze, and visualize vast amounts of information in a relevant and succinct manner. Incorporation of a GIS-based framework is designed to comprehensively analyze the economic and environmental impacts of various development options for biophysical and socioeconomic infrastructures. This platform will be designed to connect data, people, and knowledge to build a Community of Practice. The KDF will be built upon a standards-based dynamic and scalable architecture that integrates, from distributed archives, agricultural infrastructure related data, models, and tools developed by government, academic, and private sector partners. A robust geospatial technology framework provides efficient data collection, integration, management, and analysis through Geographic Information Systems; visualization through Geographic Information and Exploration Systems; and dissemination through Geographic Information Services. Web-enabled and role-based interactive access ensures wide accessibility and usability. The long-term goal is to develop a KDF for CIMMYT by leveraging the core functionalities of the Bioenergy KDF technology platform with appropriate customizations guided by focused requirements and feedback from the CIMMYT team. However, in the near term, we propose to demonstrate the utility of the KDF platform for a CIMMYT led project (such as the Cereal Systems Initiative in South Asia or CSISA) through a prototype application. This prototype will integrate a small subset of basic components (databases and spatial analysis tools) and functionalities relevant for the CSISA community.

Task 2. Development of novel data and information products

ORNL will explore novel spatio-temporal data generation and integration methods that allow exploration and understanding of geographic characteristics of environmental, physiographic, and socioeconomic resources including climate, weather, topography, and population from local to global scales. Observation and measurement, as well as social and physical survey based data, provided by CIMMYT or independently acquired, will be compared and analyzed to develop novel higher level data products and provide plausible physical and socio-cultural insights. Such data sets will also be used to develop regional models leading to novel physical and socioeconomic data products that could be spatially enabled for analysis. Observation and commercial activity (such as trading) data will be utilized to augment such models and will inform multi-scale and/or hybrid modeling of regional and sub-regional economic, political, and social stability. For example, spatial aggregation and interpolation based modeling could be applied to available surface weather observation data to create high resolution spatiotemporal estimates of rainfall and other weather parameters. It would then be possible, based on our

current understanding of the available data resources, to use the Daymet method (http://daymet.ornl.gov) to generate gridded daily surface weather data at 1 km resolution, with quantified uncertainty information. Similarly, spatial disaggregation or smart interpolation based modeling could be applied to available data at a coarser scale (such as from the Indian Meteorological Department) to develop higher resolution weather information. ORNL has already developed and demonstrated the success of this approach by developing very high resolution population distribution and dynamics model and data called LandScan (www.ornl.gov/LandScan). Through focused interaction with the CIMMYT team, we expect to be able to customize the format, summarization, and delivery of novel data output to optimize its usefulness for land management policy and practice in the regions of interest. Additionally, ORNL will collaborate with CIMMYT to identify key data and information gaps and explore novel ways of generating desired data through applications of large scale data mining from the open source. For example, structured and unstructured data from disparate sources, including internet portal and social media platforms, could be mined. In the context of weather information, ORNL could employ web mining activity to scour daily weather reports from public news sources to create a real time product for a country of region.

Task 3. Incorporation of remote sensing data dissemination tools

This task will facilitate the development of remote sensing data exploration and dissemination services through open standard based technology that allows user friendly access to a range of remote sensing based data products to CIMMYT scientists and partnering organizations. ORNL will leverage the experience and expertise in research user support service for three large data archives (DOE Atmospheric Radiation Measurement/ARM and Carbon Dioxide Information and Analysis Center/CDIAC and the NASA biogeochemical DAAC/ Distributed Active Archive Center) to design, develop and operationalize of remote sensing data delivery and sub-setting tools and services for landscape assessment, specifically, crop assessment studies. Data quality filtering and visualization services for data products, in particular MODIS (Moderate-resolution Imaging Spectroradiodmeter) terra and aqua data products will be provided. Integration of diverse temporal, spatial and spectral remote sensing products and development of web based MODIS sub-setting and visualization interface for data delivery will be integrated and exposed through open standards based with the data and knowledge discovery framework.

Task 4. Development and integration of scalable data analysis and application tools

Monitoring crop biomass at global scale for identifying changes and recognizing early warning signatures for important crop diseases is a challenging task. So far, the state of art systems are limited to visualizing NASA MODIS time-series data for a given plot, or identifying static changes at small spatial footprints. More complex systems, for example, Gaussian Process (GP) based change detection systems, in theory can be applied to global scale problems, but they are not scalable on traditional computing infrastructure. Research in this area will address these challenges and include algorithm and model development related to high resolution spatiotemporal mapping of vegetation, infrastructures, socio-economic indices, and environment-infrastructure correlations. Research activities encompass geographic data analysis, spatiotemporal data mining to explore spatiotemporal database capabilities with management of uncertainties. Drawing expertise from interdisciplinary areas such as statistics, data mining, nonlinear dynamics, risk analysis, and optimization, we focus on knowledge

discovery from multisensory, diverse, and disparate data streams, and advanced statistical analysis of scientific data using scalable computing. Data driven analysis of the earth observation data for improved understanding of the physical phenomena requires novel feature generation and pattern recognition techniques. In this task we will develop methods to characterize complex facilities by semantic modeling of the geospatial objects, develop spectral and structural features to represent the geospatial objects, develop and demonstrate a comprehensive system for mapping, monitoring, and visualizing physical and social characteristics of agricultural and food systems components, and develop data derived features to characterize the infrastructure and other socio-economic indicators from satellite imagery.

Task 5. Sensors networks and measurement tools

High fidelity data sourcing is critical for developing data driven biophysical and biogeochemical models and underpins all analytical tools. ORNL has a long history of innovation in the areas of advanced measurements, sensor networking, embedded intelligence, power management, and secure and robust communication. This unique combination of expertise enables a systems approach to advanced measurement observatories that will provide opportunities to correlate data across geospatial scales. This task will focus on evaluating, developing, and deploying novel sensor technologies to collect observation and measurement data at field and experimental scales that will allow increasing sensitivity and fidelity of the sensors coupled with communication technologies for time sensitive coordination during the data collection and aggregation process. Integration of sensors for accurate, field scale measurement and reporting of biophysical parameters utilizing commodity mobile devices in the developing farming societies could enable unprecedented rate of data collection for science and practice. Such locally calibrated sensor data will facilitate development of advanced crop dynamics at micro scale that could be correlated with agriculture dynamics models of vegetative stress, sub-optimal resource utilization or meteorological factors developed from observations at mess and macro scales. Moreover, it is anticipated that ORNL expertise will equip agriculture sustainability practitioners with ground-based sensor networks to complement existing satellite-based remote sensing information.

Task 6. Identification and evaluation of ecosystem service indicators

Adoption of more sustainable agricultural practices entails defining sustainability, developing easily measured indicators of sustainability, moving toward integrated agricultural systems, and offering incentives or imposing regulations to affect farmer behavior. This task will identify indicators that are useful for assessing both socioeconomic and environmental sustainability of agricultural and food systems. Effective indicators can help to identify, quantify, compare and establish the sustainability attributes of particular agricultural options. This effort will build from ORNL's work that identified indicators of bioenergy sustainability to establish an indicator-to-assessment approach using context-specific targets as a way to improve agriculture sustainability. The analysis will identify indicators that are practical, doable, useful to decision makers, applicable across the entire agricultural supply chain, and incorporate key areas of interest to science. The advanced sensor networks and measurement tools described in task 5 will be used to inform the selected indictors. Targets for indicators will be established based on scientific analysis, regulations, or consensus processes involving negotiation. Establishment of context-specific targets for each indicator and prioritization of goals by stakeholders provides a

way to be explicit about progress toward enhancement of benefits provided by ecosystem services.

Task 7. Coupling human and natural systems for landscape analysis

This task will specifically focus on leveraging ongoing research at ORNL to address important scientific, technical, and policy related challenges in the CIMMYT mission areas. One key interest is to develop a general modeling framework for capturing dynamic interactions within and/or across coupled human/environmental systems in terms of natural capital (ecosystem services), human capital, social capital and built infrastructures, and how those distributions and interactions influence agricultural production at different spatial and temporal scales. Current research is focusing on the development of empirical and process-based methods to understand the effects of climate variability and human system dynamics on agricultural systems, including exploiting remote sensing data, multivariate data integration, and ensemble crop modeling. This task will therefore focus on extending this framework to explore the future impacts of climate change, the benefits of different management or development interventions, as well as applications of the framework in other global regions. Future research within the task could also focus on the development of methods for modeling complex interactions between climate change, human systems, and other economic sectors and/or infrastructure systems (e.g., water resources, energy systems).

A second key interest is analysis and assessment of effects from current agriculture practices in order to devise better management practices for particular situations. This work builds from ORNL's expertise in devising landscape design approaches for enhancing ecosystem services associated with bioenergy. Appropriately applied, these approaches can guide choices toward more sustainable provisioning of social and ecosystem services. Topics that need further attention at local and regional scales include protocols for quantifying material and energy flows through agricultural systems; standard specifications for management practices and corresponding effects; incentives and disincentives for enhancing economic, environmental, and social conditions (including financial, regulatory and other behavioral motivations); integrated landscape planning and management; monitoring and assessment; effects of societal demand on agriculture; and integrative policies for promoting agricultural sustainability.

Task 8. Program coordination and management

ORNL will work closely with CIMMYT in developing technical concepts for short and long-term strategies for this collaborative program. These plans will address the impacts that natural and human-directed environmental changes are having on landscape management practices in the context of sustainability science. ORNL personnel will provide routine technical and programmatic consultation for coordinating and aligning the mission driven priorities with ORNL capabilities, engage key ORNL technical activities and researchers, and communicate between the two organizations to ensure a successful execution of the annual operating schedule. ORNL will help CIMMYT establish relationships with stakeholders, and assist in developing cooperative and collaborative multi-agency research and development activities. ORNL will assess and evaluate available technologies for optimal match with CIMMYT customer requirements.

Appendix 3: Meeting Participants

The group photo taken 21 May 2014 can be found at:
https://uwmadison.box.com/s/iyy2me8pbkdqushxk9jz

Participant list ordered alphabetically by the participant's last name

Last Name	First Name	Institution	Position	E-mail
Arista	Juan	CIMMYT	GIS Technician Jr.	j.arista@cgiar.org
Becker-Reshef	Inbal	University of Maryland	Research Assistant Professor, U Maryland	ireshef@umd.edu
Bhaduri	Budhendra	ORNL	Corporate Fellow, / Group Leader, Geographic Information Science and Technology	bhaduribl@ornl.gov
Braun	Hans	CIMMYT	Director Global Wheat Program	h.j.braun@cgiar.org
Camacho	Carolina	CIMMYT	Principal Researcher, SEP	c.camacho@cgiar.org
Carreiro	Celso	CIMMYT	Lead Software Developer	C.Carreiro@cgiar.org
Chen	Charles	CIMMYT	Statistical Geneticist, Seeds of Discovery	Charles.Chen@cgiar.org
Cox	Rachael	CIMMYT	Platform Coordinator	r.cox@cgiar.org
Dale	Virginia	ORNL	Corporate Fellow / Science Team Leader, Landscape Ecology	dalevh@ornl.gov
Dendooven	Luc	CINVESTAV	Associate Scientist	dendooven@me.com
Dreher	Kate	CIMMYT	Germplasm Data Coordinator	k.dreher@cgiar.org
Fulss	Richard	CIMMYT	Head Knowledge Management	r.fulss@cgiar.org
Gardeazabal	Andrea	CIMMYT	ICT Specialist	andreagardeazabal@gmail.com
Gerard	Bruno	CIMMYT	Director Global Conservation Program	B.Gerard@cgiar.org
Govaerts	Bram	CIMMYT	Associate Director Global Conservation Program	b.govaerts@CGIAR.ORG
Grundy	Mike	CSIRO	Theme Leader, Landscape Systems and Trends	Mike.Grundy@csiro.au
Gulledge	Jay	ORNL	Director, Environmental Sciences Division	jgulledge@ornl.gov
Hearne	Sarah	CIMMYT	Maize Molecular Geneticist	S.Hearne@cgiar.org
Hernández	Victor	CIMMYT	Junior GIS Analyst	vmhernandez@cgiar.org

Herrera	Jesus	CIMMYT	Information and Knowledge Management	j.herrera@cgiar.org
Huestis	Gordon	CIMMYT	Maize Molecular Breeder	G.HUESTIS@cgiar.org
Jahn	Molly	UW Madison	Professor, University of Madison-Wisconsin	mjahn@cals.wisc.edu
Keller	Martin	ORNL	Associate Laboratory Director, Energy & Environmental Sciences Directorate	kellerm@ornl.gov
LeZaks	David	UW Madison	Managing Scientist - Knowledge Systems for Sustainability (KSS)	lezaks@wisc.edu
Lopez	Victor	CIMMYT	Manager of Institutional Relations	v.lopez@cgiar.org
Lumpkin	Thomas	CIMMYT	General Director	T.Lumpkin@cgiar.org
McDonald	Andrew	CIMMYT	Cropping Systems Agronomist	A.McDonald@cgiar.org
Ortiz - Monasterio	Iván	CIMMYT	Fertility Coordinator	i.ortiz-monasterio@cgiar.org
Payne	Thomas	CIMMYT	Head GRC & Wheat Genetic Resources Bank	t.payne@cgiar.org
Pixley	Kevin	CIMMYT	Director Genetic Resources Program	k.pixley@cgiar.org
Polsky	Yarom	ORNL	Group Leader, Sensors & Controls Research	polskyy@ornl.gov
Rodriguez	Francelino	CIMMYT	Biometrician	F.A.Rodrigues@cgiar.org
Rodriguez	Horacio	CIMMYT	Extension Partners Coordinator	H.Rodriguez@cgiar.org
San Vicente	Felix	CIMMYT	Maize Breeder - Latin America	F.SanVicente@cgiar.org
Schneider	Kate	Gates Foundation	Associate Program Officer	Kate.Schneider@gatesfoundation.org
Schulthess	Urs	CIMMYT	Remote sensing consultant	U.Schulthess@cgiar.org
Shrestha	Rosemary	CIMMYT	Acting Data Coordinator	r.shrestha2@cgiar.org
Singh	Sukhwinder	CIMMYT	Wheat lead for Seed	Suk.Singh@cgiar.org
Sison	Darell	CIMMYT	Head of Web Development	D.Sison@cgiar.org
Sonder	Kai	CIMMYT	Head, GIS Laboratory (Geographic Information Systems)	k.sonder@cgiar.org
Tattaris	Maria	CIMMYT	Remote Sensing	m.tattaris@cgiar.org
Thornton	Matthew	CIMMYT	Hubs Coordinator	m.thornton@cgiar.org
Traschsel	Sam	CIMMYT	Maize Physiology	S.Trachsel@cgiar.org

Vannan	Suresh K.S.	ORNL	DAAC Manager	santhanavans@ornl.gov
Vatsavai	Raju	ORNL	Research Scientist, GIS	vatsavairr@ornl.gov
Vega	Daniela	CIMMYT	Projects Coordinator and Liaison Officer	D.Vega@cgiar.org
Verhulst	Nele	CIMMYT	Strategic Research Coordinator	n.verhulst@cgiar.org
Wenzl	Peter	CIMMYT	Seed Project Leader	p.wenzl@cgiar.org
Wood	Stanley	Gates Foundation	Senior Program Officer, Data and Diagnostics	Stanley.Wood@gatesfoundation.org
Zamora	Philippa	CIMMYT	MasAgro Movil Leader	p.zamora@cgiar.org
Zhang	Xuecai	CIMMYT	Maize Genomic Selection Breeder	XC.Zhang@cgiar.org