The Object Modeling System (OMS)

An Advanced, Object-Oriented, Modular Modeling Computer Technology for Agricultural Production and Natural Resources Systems

White Paper

USDA-ARS:
Dick Amerman
Will Blackburn
Laj Ahuja
Olaf David

USDA-NRCS:
Jon Werner
Jack Carlson

USDA-CSREES:
Ray Knighton

USGS:
George Leavesley

February 20, 2002

INTRODUCTION

Integrated Computerized Technical Guides – Essential to Agricultural Science and Technology in the 21st Century

In the 20th century, we made tremendous advances in discovering fundamental principles in different scientific disciplines that created major breakthroughs in management and technology for agricultural systems, mostly by empirical means. However, as we enter the 21st Century, agricultural research has more difficult and complex problems to solve. The environmental consciousness of the general public is requiring us to modify farm management to protect water, air, and soil quality, and minimize global climate change effects, while staying economically profitable. At the same time, market-based global competition in agricultural products is challenging economic viability of the traditional agricultural systems, and requires the development of new and dynamic production systems. Our customers, the agricultural producers, are asking for a quicker transfer of research results in usable form to meet these challenges.

To achieve all the above goals requires us to look at the whole system. The agricultural systems involve highly complex interactions of its soil-plant-weather-management components that are beyond a human brain to comprehend quantitatively. The modern computer technology can complement and assist the human brain in this process. Therefore, the scientific community has to present their research results in the context of the whole agricultural system in a computer. The only way to do this is to synthesize the experimental data into integrated scientific models of agricultural systems. Effective analysis and management of agricultural production systems and the environment requires integration of tools and data types that now exist in an array of individual disparate models.

Combining system modeling with good experimental data has been a vital step in many scientific achievements. We would not have gone to the moon successfully without the combined use of good data and models. Models have been used extensively in designing and managing water resource reservoirs and distribution systems, and in analyzing waste disposal sites. Although a lot more work is needed to bring models of agricultural systems to the level of physics and hydraulic system models, agricultural system models have gone through a series of breakthroughs and can be used for practical applications, with some good data. These applications will further improve the models over time. An integration of field research and models will take agricultural research and technology to the great higher plateau

A need for a whole-system approach does not mean that we will always need a whole-system model as a guide to management. After an initial analysis of the system has shown that only a certain component, such as soil erosion, or a certain insect or disease, is the main problem that needs to be corrected, a model or a technology guide that addresses only this component of the system may be adequate. An important requirement, however, is that there is a consistency in the science used among components and whole-system models.

Relevance to National Issues and Problems:

The Council of Agricultural Science and Technology 1992 report entitled, *Water Quality: Agriculture's Role*, explicitly stated that agriculture should address nonpoint source pollution problems at field, farm, and watershed scales. The National Science and Technology Council, Committee on Environmental and Natural Resources, 1997 report entitled, *Integrating the Nations Environmental Monitoring and Research Network and Programs: A Proposed Framework*, recommended a national framework that links systematic observations and monitoring with predictive modeling and process research. The 1999 report of the National Academy of Sciences and Engineering on "*New Strategies for America's Watersheds*" identified several critical support functions; one of these is the integration of theory, data, simulation models, and expert judgment to solve practical problems and provide a scientific basis for decision making.

Agricultural land management has to deal with a high degree of spatial variability of land and temporal variability of weather even at field scale. This complexity increases a great deal further at farm and watershed scales. In the last few years, there has been a nation-wide interest in so-called "precision agriculture", which consists of site-specific, spatially variable, management to optimize production and minimize both on-site and off-site adverse effects on water quality and quantity (Vanden Heuvel, R. M. 1996. *The Promise of Precision Agriculture.* Journal of Soil and Water Conservation 51:38-40). However, no methods and tools are available to guide this management of landscape and climate variability. For this purpose and to aggregate expected results from plots to field, farm, and watershed scales, we need a framework for integrated spatial analysis, and distributed modeling, and up- and down-scaling for different land areas. This framework will then allow downscaling of watershed level observations, such as TMDLS, to farm and field levels for the purpose of site-specific management. A scientifically robust scaling framework is the greatest need for making breakthroughs in transferring research knowledge across scales, and in understanding and managing large areas (National Research Council, 1991: *Opportunities in Hydrologic sciences*).

Relevance to NRCS and USDA strategic Goals and Objectives:

Increased capability to assess baseline conditions and simulate various practice alternatives of watersheds and ecosystems are necessary to accomplish the specific goals and objectives of the NRCS and USDA 2000-2005 Strategic Plans. The NRCS needs the development of new knowledge, data management and access technologies, improved modeling technology and decision support tools for guiding management on field, farm and watershed scales. The broad objective will be to develop a decision support system that integrates models, databases and spatial tools to evaluate both economic and environmental impacts of on-field and off-field management practices, guide precision agriculture and enhance watershed sustainability.

CURRENT STATUS AND CUSTOMER NEEDS

At present, the ARS and university scientists have more than a hundred small to large models or software packages to help evaluate or guide management practices; some of them are system-wide models whereas others are component models or software. Examples of large system models are: cotton and soybean models, Gossym and Glycim or Soygro; the crop models of CERES and CROPSGRO family; erosion models, WEPP and WEPS; water quality models, GLEAMS, RZWQM, AGNPS, and SWAT; and general purpose models, EPIC and CROPSYST. Examples of smaller, single-purpose, models are NLEAP and WEEDCOM. Unfortunately, these tools now exist as an array of individual disparate models. The large system models have been extremely expensive to develop ($15-30 million each).

At ARS customer workshops, the NRCS and other users of models or software have reported that different models or software tools do not give the same results for, say, soil erosion or crop yield, because these tools use different science in key areas like hydrology. There may also be conflicting data, scales, and methodologies.

Moreover, it is becoming extremely difficult to maintain and support such a large number of packages. In particular, yesterday's monolithic models are very difficult to update, add to, or interface with other models, have diminishing technical support, and lack the flexibility to meet today's needs for more integrated analysis of changing resource issues. It is difficult to transfer the new knowledge to the customers quickly through the use of these old tools.

All of the above reasons indicate a need for a new framework of model development that integrates all existing models and all future models into a common, collaborative, and flexible system. Such a system will maintain modularity, reusability, and interoperabililty or compatibility of both science and axiliary components. The system will also recognize the fact that different categories of applications may require different levels of scientific detail and comprehensiveness, as driven by problem objectives, scale of application, and data constraints. These functionalities of the system will be obtained by establishing standard libraries of inter-operable science and axiliary components or modules that provide the building blocks for a number of similar applications. Nodule libraries have been successfully used in several domains, such as the manufacturing, transport, and other systems (Top et al, 1997; Breunese et al, 1998; Praehofer, 1996). One of the earliest modular model developments was done for SHE, the European Hydrologic System Model (Abbot et al, 1986; Ulgen et al, 1991). Leavesley et al (1996 reported the conversion of the Precipitation Runoff Modeling System (PRMS) to a Unix-based Modular Modeling System (MMS) for hydrologic modeling. Leavesley et al (2002) presented some successful applications of this concept.

In 1997, and ARS-NRCS-USGS Interagency Workshop reviewed the MMS and other similar approaches and unanimously endorsed the development of an advanced Modular Modeling Framework for Agricultural and Natural Resource Systems that will:

- reduce duplication of effort in agricultural and natural resource modeling
- improve the quality and currency of model code
- make simulation models much easier to build, access, understand and use
- facilitate long term maintainability of existing and new models
- lead to greater consistency of modeling for particular problems and scales
- enhance response and delivery times in scientific modeling projects
- ensure creditability and security of model implementations
- function on any major computing platform

There is no software commercially available to build models from modules composed of existing source code. Some technology is available which uses a common interface for running different pre-built models, such as groundwater and surface water models. These systems cannot build models from modules.

Most current investments in model development programs by our university partners (e.g., The University of Florida DSSAT Modeling Group), other government agencies (Corps of Engineers, Environmental Protection Agency), and foreign countries (e.g., the Cooperative Research Centre (CRC) for Catchment Hydrology in Australia; the European Commission Framework Programme 5: Harmon IT) are in the area of modular or object modeling, all of which use a basic concept of providing proven and sound scientific modules of the past made available for use in custom-designed integrated analyses.

The Object Modeling System, described below, meets the above criteria. In 2001, several federal agencies – Nuclear Regulatory Commission, NOAA, EPS, COE, USGS, and USDA – entered into a Memorandum of Understanding to collaborate in developing and using models.   These agencies enthusiastically endorsed the development and deployment of the Object Modeling System.

THE OBJECT MODELING SYSTEM (OMS)

The OMS project was initiated in October 2000 as an interagency project between the USDA-ARS, USGS, and USDA-NRCS, with financial support from ARS and in-kind support from the partners. The financial support was used to invite Dr. Olaf David, Computer Scientist from Jena, Germany, who has been a key to OMS development. During the past 18 months, we have jointly completed the development of core components of OMS. The vision of OMS, described below, is close to being realized:

> *"The OMS is a computer framework consisting of: a library of science, control, and database modules; a means to assemble the selected modules into a modeling package customized to the problem, data constraints, and*

scale of application; an automatic generation of a friendly user interface; and creation of a compiled, ready-to-run, version of the package. The framework is supported by utility modules such as data dictionary, data retrieval, GIS, graphical visualization, and statistical analysis. The framework employs the latest technology for all its components. The science modules are also quickly updated or replaced as new knowledge and data become available. The OMS will be supported from a central server for use by all ARS scientists, NRCS specialists, USGS, and other collaborators."

The objectives of the OMS project included the development of generic software tools to glean modules from existing non-modular simulation models, and to enter them into the OMS framework with standard OMS descriptions. This tool has been developed, but needs further tests and improvement.

The OMS framework has the following functional components:

1. A module-building component that will facilitate the integration of existing (legacy) code into the framework.
2. A module repository that will contain modules that can be readily utilized to assemble a working model (types of modules in the library will include science-, control-, utility-, assessment-, data access-, and system-modules.
3. A model builder that will assemble modules from the module library into executable models and verify data connectivity, and compatibility in scale and comprehensiveness
4. A dictionary framework that will manage extended modeling data type information and provide extended semantics checking for module connectivity verification
5. An extensible user interface that will facilitate an appropriate user interaction for general model development and application (it will be supported by a number of contributing software packages for database management, visualization, and model deployment)

The components have the following architecture or characteristics:

1. OMS models are treated as hierarchical assembled components representing building blocks. Components are independent and reusable software units implementing processing objects for simulation models. They reside in a model library and are categorized into data access components, science components, control components, utility components, and system components.
2. OMS is able to integrate legacy code components. By an automated JAVA wrapper generation for legacy code, components written in languages such as Fortran can be embedded into OMS at the function level.
3. The "knowledge"-backbone of OMS is the dictionary framework. It enables OMS to verify state variables and parameters according to scientific nomenclatures during model development and application. Dictionaries are also

used to specify parameter sets, model control information and the component connectivity. They are implemented in the Extensible Markup Language (XML).

4. OMS is extensible. Extension packages exist for different aspects in model development and application. Extension packages are used for visual model assembly, model application, an interface to the dictionary framework, and GIS.

5. OMS scales from a full-featured, stand-alone development system with tools for model assembly, visualization, and analysis to a runtime Web service environment.

Each of the above features is described in more detail by David et al (2002).

Why should ARS-NRCS-CSREES deploy OMS?

1. **Tremendous Cost Savings in Building New Software Technology for all National Programs (Simple Decision Support Tools, Models, Decision Support Systems):** In the past ARS has had a number of individuals and teams build software technology and simulation models. Large system-level software packages required large teams of scientists, such as for the cotton and soybean models Gossym and Glycim, the erosion models WEPP and WEPS, and water quality models GLEAMS, RZWQM, AGWPS, and SWAT. Each of these packages cost ARS between $15 to $35 million dollars to develop, including scientist and support time, that contained significant duplication of work. The ARS can now leverage that investment by putting the science in those packages as modules in OMS to build new customized software packages at a small fraction of the cost. First, preliminary results in the core OMS development phase on the modularization of existing models RZWQM and PRMS showed a code reduction of OMS versions of these two by 20 – 33% while keeping the same modeling results. In this age of information technology, the demand for such software packages will increase tremendously. If we were to develop ten customized large new system packages to meet this demand over the next ten years, we will save at least $100 million dollars, assuming 80% of the science comes from existing modules already put in OMS and 20% is new code.

2. **Tremendous Cost Savings in Upkeep, Maintenance, and Customer Support for Software Technology:** It is generally agreed that over long-term these items cost three times as much as for initial development of the software packages. At present, the ARS has more than a hundred small to large software packages that need to be maintained and supported. A huge amount of SY time is being spent on this process. This time will at least double if twenty new large packages are added on over the next ten years. In fact, yesterday's monolithic models are becoming very expensive to use, difficult to update, have diminishing technical support, and lack the flexibility to meet today's needs for more integrated analysis of resource issues. These problems can be overcome if all of the existing packages where transferred to OMS and all new packages were developed within OMS; all the SY time will be saved to do research, as just 3-4 CAT IV staff in

OMS takes care of all this work. It is hard to estimate, but our guess is that the OMS will save at least $100 million dollars of SY time over ten years. With an earlier version of OMS, called MMS, the USGS has shown the huge cost savings in upkeep, maintenance, and customer support. Under interagency MOLL, several federal agencies, NRC, EPA, DOE, COE, USGS, and ARS have endorsed the development and deployment of OMS as a common platform for all agencies.

3. **Integrated Analysis of Whole Agricultural Systems for Production and Resource Issues:** Effective analysis and management of agricultural production systems and the environment requires integration of tools, and data types that now exist in an array of individual disparate models. The OMS will provide customized, whole-system, tools for environmental quality and global climate change management in crop and animal production systems.

4. **Easier and Faster Transfer of Technology:** For the transfer of technology tools to commodity associations and action agencies, like NRCS, Extension, and EPA, the OMS will serve as a platform for different software tools, from crops to animals to natural resources. The associations or agencies will then develop deployment links to only OMS, rather to develop links to each of the separate software tools as is done now. This will result in a faster transfer of ARS technology. The NRCS, Information Technology Center in Fort Collins, has been a partner in the OMS development and is fully committed to using OMS as a means to provide technical tools to 2,500 field offices for conservation planning. The NRCS-ITC is already developing links to OMS. Customized OMS models are being used by USGS, Denver (George Leavesley), BLM and Fish and Wild Life Service for water and land management in a variety of ways.

5. **National Coordination in New Science Module Development and Publication:** Through the internet, the OMS will serve a common platform for all ARS scientists to contribute their findings as modules to the OMS library. The ARS-NPS could coordinate this development, and provide a mechanism for peer review and quality control. The module contribution to the library will be considered a publication by scientists that will have world-wide impact.

6. **Using the Most Appropriate Science for Each Problem:** The OMS will allow the selection of the best and most appropriate science modules available to-date, depending upon the nature of the problem and answers required, availability of input data, and scale of application. The OMS library may have different modules for a research model versus a management decision tool. Similarly, a watershed-scale management model may require simpler science modules than a field-scale model.

7. **Assure Consistency in Results From Software Tools for Similar Applications:** At ARS customer workshops, the NRCS and other software users have reported that different software tools gave different results, say for soil erosion or crop yield, because the tools used different science in key areas like hydrology. The

OMS will significantly reduce this problem by utilizing evaluated, documented, and standardized modules for the basic science components for a given category of applications.

8. **Reduced Start-Up Time for Future Modelers and Lower Training Costs for Users:** Due to a common model building platform and a common user interface for all models, the OMS will result in those savings as well.

9. **OMS Library as a Reference and Coordination Mechanism for Future Research and Development:** The OMS library will be a repository of current, quantitative knowledge in different areas of agricultural science. Future scientists could look to this library to determine where further research and development are needed.

10. **OMS Certification mechanism for approved "science building blocks."** The OMS is supporting the technical certification of library components based on X.509 Certificates and the validation of such certificates. This will allow the agency to certify approved modeling components and models.

11. **The ARS Information Technology Strategic Plan:** Calls for such advanced technology packages for increasing innovation and efficiency in technology transfer.

12. **Increase in Productivity of ARS Scientists:** The customized, best-quality, software tools developed through OMS will help field scientists quantify their results and transfer them to other soils and climates very rapidly. The gaps identified in the process will make further research more focused. Overall, the productivity of scientists and the quality of their science will increase. Scientists can really focus on science module implementation rather than User Interface design, software deployment, packaging and maintenance.

The OMS implementation at ARS is illustrated in Figure 1.
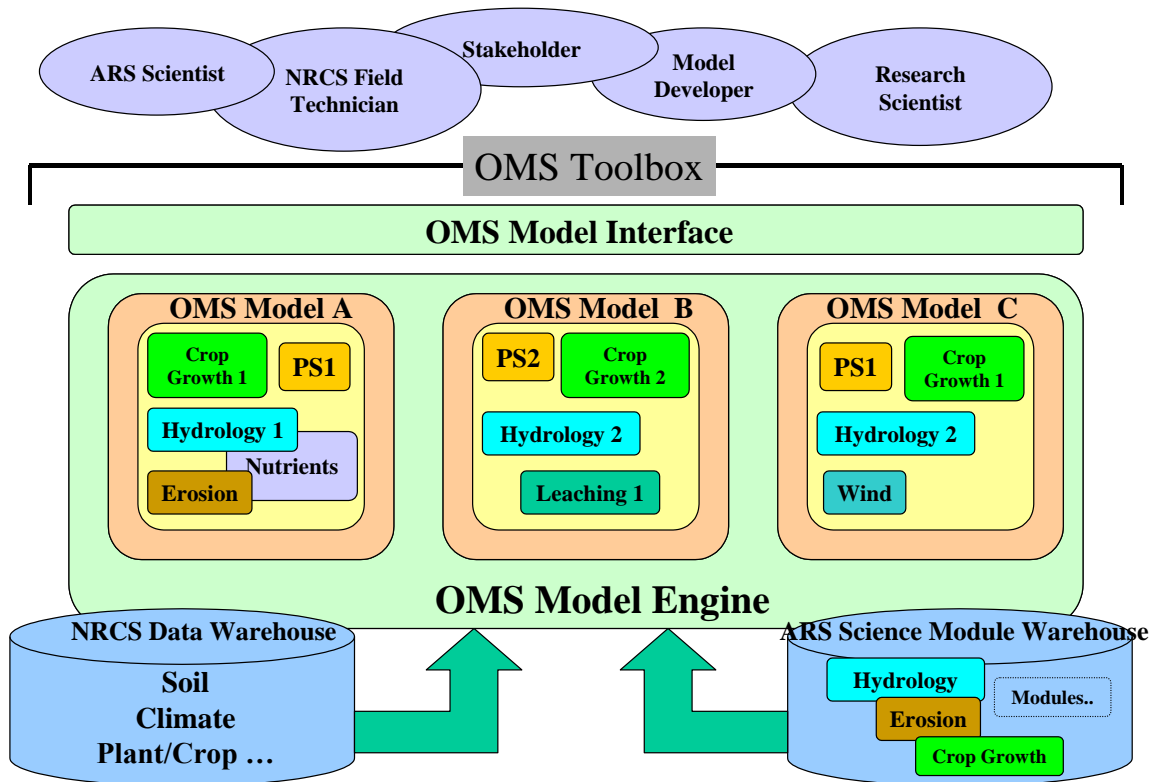
OMS Implementation at ARS

Figure 1: OMS Implementation at ARS

Infrastructure to Deploy OMS ARS-Wide – The Concept of OMS Central

The effective deliver of customer requested decision support tools based on best science for application in agriculture will be based on common use of the Object Modeling System infrastructure. It will promote a closer interaction between the science and operations communities in carrying out their respective responsibilities. To allow maximum utility and flexibility, simulation models will be constructed from affordable, reusable components interoperating through an open-systems architecture. Based on the recommendations of an ARS-NRCS-USGS National Guidance Committee, we propose that the ARS establish a new entity, call "OMS Central" to provide maximum support.

The mission of "OMS Central" will be to: (1) work with ARS scientists and customers to develop and deliver requested software tools based on the best and most appropriate sciences for application in agriculture; (2) update and enhance the computer technology used in OMS with time; (3) develop and maintain a state-of-the-knowledge library of well-documented science and utility modules; and (4) collaborate with scientists and institutions world-wide in the above endeavors.

To achieve the mission OMS Central will undertake multidisciplinary research and implement several programs delivering products and services:

Model Development and Integration Program.
- Design, Development and Integration of new models by using the Object Modeling System.
- Improving the design existing models for wider application; maintenance; agency wide model investment saving.
- Maintenance of existing models.
- Developing Libraries of Modules.

OMS Framework Research Program.
- Research in Object-oriented Modeling Frameworks, design and modeling principles, implementation into OMS,
- Establishing an Internet Collaboration Infrastructure.

Model Implementation Standards and Certification Program.
- Developing model implementation standards by means of model for model implementation
- Model certification based on model testing, validation and verification; applying X.509 Certificates
- Providing Model development guidelines.

Scientific Collaboration / Seminar and Workshop Program.
- Improving the communication in the scientific community by providing a collaboration platform for project related cooperation among researches, research units and the general scientific community.
- Seminars usually focus on a specific issue or area of research and often include general implications of the research outcomes to date for the land and water management industry. The speakers provide detailed information about the research techniques, analysis and results.

Deployment Program.
- That will result in the development, support and improvement of model deployment paths for the scientific community and the NRCS field offices.

OMS Central Structure

We envision that "OMS Central" will develop gradually in phases. It will grow with time as the customer base and customer demands increase and its value to ARS programs is demonstrated. In Phase 1, we propose only a limited number of core staff.

1. System Architect & Manager: One computer/information technology scientist, Cat. IV, to (i) organize and develop the program; (ii) supervise all operations and staff; (iii) lead efforts in enhancing and updating the computer technology used in OMS; (iv) train staff and ARS modelers in the use of OMS technology; (v) establish national priorities in consultation with NPS; and (vi) develop and interface with collaborators in other federal agencies and world-wide.

2. <u>System Developer</u>: One information technology specialist to: (i) work with the computer scientist to maintain, update, and enhance the computer technology used in OMS; and (ii) help train and then work with and support ARS users, NRCS, USGS, and other collaborators, in the use of OMS technology.

3. <u>Customer Programs</u>: Two OMS application support specialists, IT, Cat. III, or Cat. IV, to work with ARS scientists in their projects: (i) in modularizing their old code; (ii) creating new models by appropriate selection of modules from the library; (iii) customizing user interfaces; (iv) creating and supporting ready-to-run packages; (v) maintaining and enhancing the module library; (vi) helping ARS scientists put their new research as modules; and (vii) other such customer support tasks.

4. <u>Part-time, temporary,</u> IT students to help in above programs.

The above ARS core staff will be augmented by in-kind support from our NRCS and USGS partners:

1. The NRCS-ITC, Fort Collins, will detail one IT specialist to "OMS Central" to work with OMS core staff on developing and maintaining interfaces to NRCS databases, and for deployment of OMS for their Field Electronic Technical Guide, FSA Webfarm, and other applications.

2. The USGS, Denver, will detail one IT specialist intermittently to work with OMS core staff on enhancing OMS capabilities for watershed and spatial applications, remote data retrieval, creating management decision interfaces to OMS simulation results, and other such customer applications.

DEPLOYMENT OF OMS AT NRCS-ITC

The deployment of OMS Runtime packages at the NRCS, Information Technology Center, Fort Collins, Colorado, is illustrated in Figure 2.

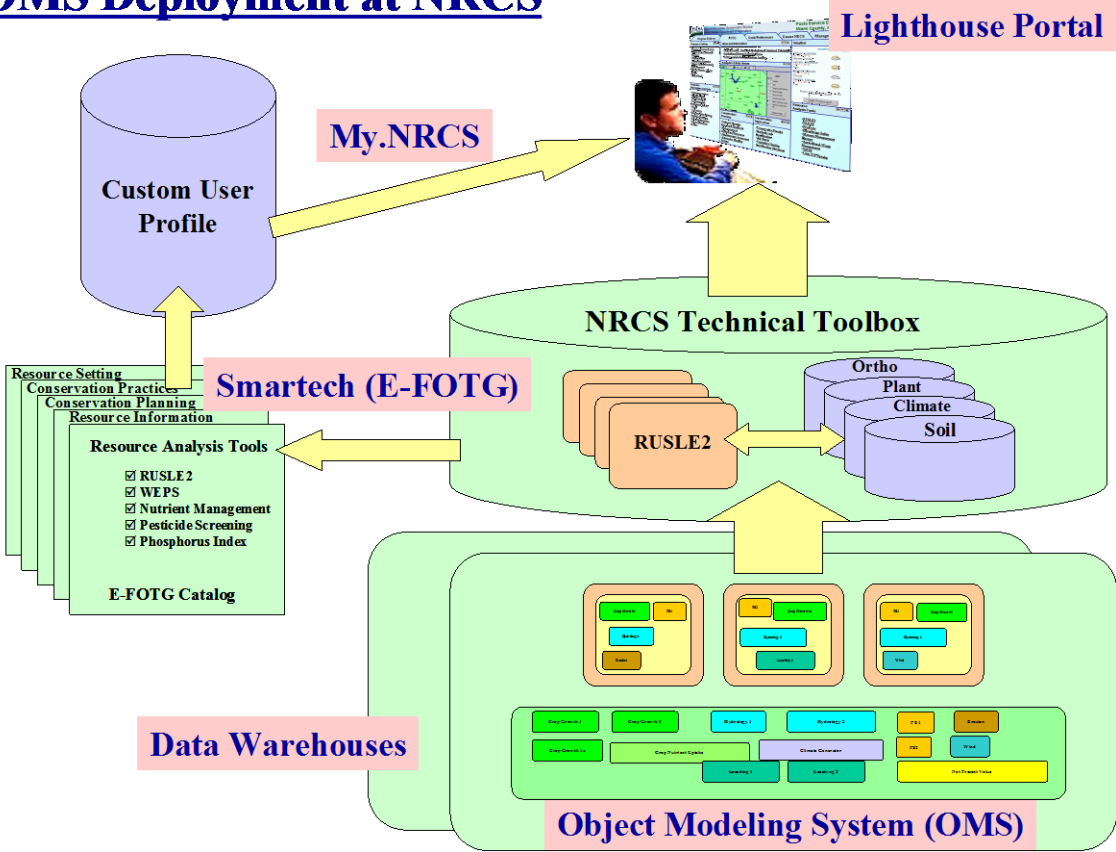The deployment will result in a much faster transfer of new technology to NRCS and an integrated analysis of environment and resource issues at field, farm, and watershed scales.

Figure 2: OMS Deployment at ITC