Conceptualizing a Collaborative Problem Solving Environment for Regional Climate Modeling and Assessment of Climate Impacts

George Chin Jr., L. Ruby Leung, Karen Schuchardt, and Debbie Gracio

Pacific Northwest National Laboratory, P.O. Box 999,
Richland, Washington 99352 USA
{George.Chin, Ruby.Leung, Karen.Schuchardt,
Debbie.Gracio}@pnl.gov

Abstract. Computational scientists at Pacific Northwest National Laboratory (PNNL) are designing a collaborative problem solving environment (CPSE) to support regional climate modeling and assessment of climate impacts. Where most climate computational science research and development projects focus at the level of the scientific codes, file systems, data archives, and networked computers, our analysis and design efforts are aimed at designing enabling technologies that are directly meaningful and relevant to climate researchers at the level of the practice and the science. We seek to characterize the nature of scientific problem solving and look for innovative ways to improve it. Moreover, we aim to glimpse beyond current systems and technical limitations to derive a design that expresses the regional climate or impact assessment modeler's own perspective on research activities, processes, and resources. The product of our analysis and design work is a conceptual regional climate and impact assessment CPSE prototype that specifies a complete simulation and modeling user environment and a suite of high-level problem solving tools.

1 Introduction

At Pacific Northwest National Laboratory (PNNL), a group of computational and earth scientists have been collaborating on the design of a collaborative problem solving environment (CPSE) for regional climate modeling and assessment of climate impacts. This collaborative effort is at the crosscurrents of two fields of specialty at PNNL. In the area of climate modeling and impact assessment, PNNL atmospheric scientists have developed a regional climate model based on the Penn State/NCAR Mesoscale Model (MM5) that is capable of simulating climate conditions over topographically diverse regions with spatial details down to one square kilometer. The model, PNNL-RCM [7, 9], has been applied to several geographic regions to elucidate the impacts of climate change on water resources and demonstrate the potential use of seasonal climate forecasts for water resources management [6, 8, 10]. In the area of CPSEs, PNNL computational scientists have experience and expertise in the research and development of CPSEs—most notable is the Extensible Computational Chemistry

Environment (ECCE) [4]. ECCE is a domain encompassing problem solving environment that enables chemists to transparently utilize complex computational chemistry modeling software and access high-performance compute resources from their desktop workstations.

Today, the earth sciences community is intensely focused on the rapid advancement and deployment of climate modeling capabilities as the potential hazards of global warming have been more widely recognized and investigated [15]. The general development approach is to build advanced climate modeling capabilities by leveraging existing modeling and analysis tools as well as high performance computational infrastructures [13]. As such, computational science research among the climate modeling community emphasizes largely on immediate issues such as the development of more sophisticated numerical algorithms, improving or automating the application of models, accessing large amounts of data, and running simulations across distributed networks. The focus is one of near-term, practical concerns as computational scientists incrementally provide advanced scientific simulation capabilities to climate researchers.

Computational scientists at PNNL are following a similar path in the advancement of climate modeling capabilities as we focus on incremental developments such as the parallelization of regional climate modeling codes, enhancement of analysis and visualization techniques and tools, and support for real-time computational steering. At the same time, we are also keeping an eye towards the future by envisioning how climate researchers might collaborate and work in a more ideal computational environment. Where an incremental development approach forces us to couch the climate modeling domain in view of existing computational infrastructures and technologies, the goal of envisionment is to project the advancement of computer technologies and infrastructures to meet the research functions and needs of the climate domain. In essence, envisionment guides development towards an inspiring, future vision. Both incremental and envisionment approaches are important to provide a strategy for climate and computational science to progress towards an optimal medium.

2 Conceptual Prototype Development

The climate CPSE prototype was the product of working with various domain scientists over the course of approximately six months. In the initial phase of our study, we met with five different groups of domain scientists consisting of computational chemists, regional climate modelers, nuclear magnetic resonance experimentalists, automotive engineers, and fluid dynamics modelers [2]. Through a series of participatory analysis sessions, the domain scientists elaborated their basic research and problem solving processes, functions, and needs. Our analysis with domain scientists revealed the following three areas where scientific problem solving could be enhanced through computational support:

- Easy and effective access to computational resources – Scientists need ready access to computational resources such as applications, data, data archives, and computers

in order to run their models. Resources should be represented in a way that is comprehensible and intuitive to the domain scientist.

- Experimental design and execution support – Scientists need to be able to intelligently apply computational resources in the context of a scientific process or experiment. The design and execution of computational experiments is often complex and intricate, and scientists could use better tools to assist them in defining, managing, executing, analyzing, interpreting, and sharing experiments.
- Domain and procedural knowledge management and dissemination – The ability of scientists to solve problems hinge on their knowledge of domain concepts and theories as well as the operational steps required to identify solutions. By making this knowledge explicit and concrete, scientists may be able to better maintain and evolve this knowledge, as well as share it with others.

Following analysis, the conceptual design effort was a modest undertaking involving two software designers, a regional climate modeler, and a hydrology researcher. In both analysis and design, we adhered closely to principles of participatory design [3, 14], which allowed scientists to actively participate in the actual analysis and design work. During several participatory design sessions, the collaborators engaged in the construction of a paper prototype, using large sheets of paper, pencils and pens, color markers, and Post-it™ notes. Through the use of such low-tech tools, the climate researchers were given the capacity to express their needs and ideas in the conceptual design. The paper prototyping technique we employed was based on the PICTIVE (Plastic Interface for Collaborative Technology Initiatives through Video Exploration) design method [12]. The paper prototype was then transformed into a computer prototype.

The scientific content of the conceptual prototype is based on a pilot project funded by the U.S. Department of Energy (DOE) Accelerated Climate Prediction Initiative (ACPI) [13]. Using existing climate modeling and impact assessment tools and capabilities, the ACPI pilot project aims to demonstrate an “end-to-end” approach that begins with the current state of the global climate system and ends with predicted impacts of anthropogenic climate change on water resources at the local and regional levels in the western United States. This approach requires data flow from global climate simulations to regional climate models to various independent assessment models of surface and groundwater, water management, fish habitat, and fire weather. This approach can be extended to integrated assessments of climate change impacts where data flow is also required among the assessment models for horizontal integration of impacts and feedbacks.

In the remainder of this paper, we describe specific features of the regional climate conceptual prototype and relate them to the three aforementioned scientific problem solving themes. We begin by describing CPSE capabilities for designing and executing experiments because the computational experiment is the central concept in the prototype around which other features and capabilities are organized.

3 Experiment Design and Execution

The scientific experiment is the vehicle and framework through which climate researchers attack their research problems and goals. For climate researchers, the practice of designing and executing climate models has certain tendencies and characteristics. Among those we uncovered in our analysis include the following:

- Climate modelers design their experiments by defining sequential steps that utilize the model, observational data, application tools, computers, and other miscellaneous resources.
- The experimentation process is highly repetitive. The climate modeler repeats a cycle of steps that includes modifying the configuration and initial/boundary conditions of the computational experiment, executing the experiment, and evaluating the generated output and its convergence to observed or theorized results.
- As climate modelers conduct computational experiments, they generally perform long sequences of computer operations such as logging into machines, querying for and collecting data from on-line databases and repositories, transferring data files between machines, running applications on distributed computers, capturing experiment output to files, applying translators to convert data formats, and executing analysis and visualization packages on specific data sets.
- In designing and executing their experiments, climate modelers typically maintain the design and execution processes in their heads and in notes placed on loose paper or in laboratory notebooks.

To support the experiment design and execution process, we introduce a variety of experiment management tools that allow climate researchers to better design, organize, manage, execute, and document their experimental research processes.

3.1 Scientific Workflow Execution

To support experiment design and execution in a climate CPSE, we initially focused on deriving a high-level visual representation that climate researchers could construct and utilize to manage their experimental research processes. This objective led to the investigation of scientific workflow representations and visual programming paradigms. As shown in Figure 1, we designed a graphical experiment management representation that is centered on concepts in data flow and flowchart diagrams [16].

The top workflow diagram in Figure 1 presents a high-level view of the steps or tasks of the ACPI project. The arrows of the diagram denote the sequential relationships as well as the data dependencies among the different tasks. As constructed, the diagram presents an execution model for the overall project. As the tasks are performed and completed, their status changes as indicated by the color of the graph nodes. The specific data sets that are passed from one task to another are identified along the arrows of the graph. The workflow diagram provides an abstract view of the design model and the execution state of the experiment.

Fig. 1. Scientific workflow diagrams are provided in the climate CPSE prototype to support experimental processes. The two workflow diagrams represent different levels of abstraction. The top workflow diagram depicts the high-level tasks of the ACPI project. The lower workflow diagram expands the Run Regional Climate Model node of the ACPI workflow to expose the specific computational resources that are accessed when the regional climate model is executed. The Execution History tool shows all past executions of a particular experiment. One specific branch in the history tree represents a particular thread of execution.

Graph nodes representing the high-level tasks of a project or experiment may be decomposed into specific experimental and computational resources. For example, in Figure 1, the Run Regional Climate Model task in the top workflow diagram decomposes into the more detailed network of resources shown in the lower workflow diagram. The lower-level resources identify data archives, applications, simulation models, and other subtasks. The higher- and lower-level views of experimental processes represent varying levels of abstraction to the climate modeler. Higher-level views represent more of a scientific and cognitive perspective of the experimental process. They often reflect the experimental steps at the climate researcher’s natural level of

thinking. Conversely, lower-level views emphasize the mechanical or procedural steps that are encompassed in scientific tasks.

By encapsulating computational operations within higher-level tasks, the details of the tasks are conveniently packaged for distribution and reuse. The ability to develop multiple layers of abstraction is an important capability for climate researchers because it requires the researcher to specify the intricate operations on resources only once, and then those specifications may be re-used in different runs of the same project, on different projects, or by different researchers. For instance, a novice climate modeler may elect never to expand a specific climate task to view its inner workings, whereas an experienced climate modeler may choose to expand that same task to modify portions of the underlying mechanics to support a different assessment or project.

Figure 2 illustrates the features of the visual execution model. Borrowing concepts from visual debugging, the visual execution model allows climate researchers to set breakpoints and monitors. A breakpoint pauses the execution of the workflow when a particular step is reached. Execution may be resumed from the paused state. A monitor is attached to the output stream of a particular node of the workflow. When data is generated from a monitored node, a specified data viewer or visualization application

Fig. 2. The workflow execution model provides breakpoints and monitors. Through breakpoints, climate researchers may pause and resume the current workflow execution. Through monitors, visualization tools and data monitors may automatically be spawned once specific output data sets are generated.

is invoked to display some or all of the outputted data.

In our analysis of research processes, we found that climate researchers typically conduct experiments in a dynamic, uneven way. They often pause at critical junctures during the experiment to examine, analyze, or verify critical, intermediate results. During these suspensions, climate researchers may employ one or more data analysis or visualization tools. The development of breakpoints and monitors are intended to support this spontaneous, sporadic mode of experimentation by allowing climate researchers to pause the experiment and conduct on-the-fly analysis and visualization.

For example, in Figure 2, the execution of the PNNL-RCM model and its subsequent generation of the RCM Data Set trigger the invocation of a scientific visualization tool called the DataMiner. In addition to connecting to specific analysis and visualization tools, the climate CPSE prototype also provides a general monitoring tool that allows the climate researcher to select fields of an outputted data set to be tracked and displayed. As shown in the Monitor Variable tool in Figure 2, a simple graphical plot is generated and compared against observed or saved results. Thus, the researcher need not always apply sophisticated analysis or visualization tools to extract basic pieces of information from intermediate experimental results.

3.2 Scientific Workflow Design

An important quality for the visual workflow model is that it be intuitive to and usable by climate researchers and reflects the researchers' views of their own scientific processes. In early analysis sessions, we had climate researchers construct diagrams on paper to convey their research and work practices [2]. The diagrams that they developed were very similar to the workflow diagrams that are applied in the climate CPSE prototype. The fact that the climate researchers were able to effectively construct similar workflow diagrams with relative ease leads us to believe that workflow diagrams may serve as an effective visual modeling approach for climate researchers.

Similar to other component-based visual programming models such as those provided by Advanced Visual Systems AVS [1] and Khoral Research Inc. Khoros [5], the workflow model in the CPSE prototype requires climate researchers to drag and drop modules from a palette of components onto a common workspace and to link the modules using arrows to indicate direction of data flow (see Figure 3). Unlike AVS and Khoros, however, the granularity of the components and processes is at a higher level that deals with computational and experimental resources. The purpose of the workflow model is not to construct a specific application from existing functions or objects but rather to link together a coherent set of resources in a way that allows climate researchers to effectively execute their models and conduct their experiments.

As shown in Figure 3, the climate researcher designs the experiment by placing various resources on the modeling workbench. Along the palette to the left, resources are categorized into scientific models, applications, data archives, files, and encapsulated subtasks. For each resource dragged onto the workbench, the climate researcher is required to specify the resource's attributes, which includes its directory location, run command, run options, host machine, and input and output parameters. For each

Fig. 3. The climate researcher constructs a workflow model by dragging and dropping resources onto the workbench from the palette. For each resource, the climate researcher must identify its attributes such as the directory location, run command, run options, host machine, and input and output parameters.

link between two resources, the climate researcher must also define the mapping of the data outputs of one resource to the data inputs of another resource.

From our analysis, we found that climate researchers often do not know which applications, data, and resources they will apply until they are in the midst of running their experiments. Thus, the climate CPSE prototype must accommodate the dynamic design and evolution of scientific workflows. The workflow model in the CPSE prototype does not require the climate researcher to fully design and construct a workflow diagram before executing or accessing a computational resource. For instance, the climate researcher may simply drag and drop various resources onto the workbench without connecting them. The researcher may then execute the individual resources separately and manually manage intermediate data among all the resources. This approach is not far removed from the current manner in which climate researchers manually access and apply computational resources. The advantage of performing these operations within the CPSE is that researchers may eventually gain greater comfort and expertise in constructing workflow models over time, but in the short term, they immediately gain additional problem solving capabilities through the other high-level tools provided by the CPSE.

In the climate CPSE prototype, we are also exploring the notion of an adaptable environment that may be customized to the way a climate researcher or group of researchers experiments and collaborates. Monitoring facilities may be built into the user environment such that the environment would keep track of the applications, databases, computers, tools, and other resources that climate researchers apply over time. Based on collected information, the environment may provide automatic links to those resources within the user environment for immediate access. One step further,

the environment may also keep tabs on the specific sequence of actions that particular researchers and groups conduct and automatically construct workflow configurations based on the most common patterns of behavior. This way, climate researchers would not be required to construct workflow diagrams from scratch but could rely on the user environment to provide preliminary workflow models.

3.3 Historical Experiment Execution Record

For any climate experiment, climate researchers make numerous runs against the climate model as they investigate different climate properties, regions, and time periods. In various cases, researchers may wish to revisit previous runs in an effort to reuse the initial conditions or configurations of those runs, track down a particular error or anomaly, reproduce a previous result, or identify patterns or trends in the results. To support these functions, the climate CPSE prototype has defined automated facilities for capturing, managing, and re-instantiating historical records of previous experimental runs.

As shown with the Execution History tool in Figure 1, the climate researcher may highlight and re-instantiate particular paths of executions of the climate model. The graphical tree-based depiction presents a hierarchical view of past executions with nodes identifying points of iteration in the experiment. When a particular path branches out into multiple paths, the multiple paths indicate that several experimental runs were carried out from a common state in the experiment. As shown in Figure 1, the climate researcher may select an extended thread through the history tree to reproduce that particular experimental run. Additionally, the climate researcher may retrieve input and output data sets generated at any point in the execution history.

4 High-Level Access to Computational Resources

Climate researchers do not naturally think of computational resources as applications, computers, and files but rather as models, calculations, and spatial and temporal data. One of our objectives in designing a climate CPSE is to promote the appropriate level of abstraction such that climate researchers may utilize these resources in a form consistent with their specific domain concepts and views. For the most part, we have found that current climate researchers are intimately familiar with the details of the computational environment in which they run their models. Such a prerequisite, however, is a high barrier that may prohibit many researchers, who may have the knowledge and skills to develop climate concepts and models, from contributing due to their inability to navigate the complicated computational environment. In a more ideal world, the computational environment would not be a barrier to conducting valuable climate research, and any climate researcher with sound domain expertise should be able to apply their science and contribute to the expanding body of knowledge.

In the climate CPSE prototype, we have derived a variety of tools designed to provide climate researchers with high-level access and a high-level view of computational resources. Together, these resource tools allow climate researchers to perform what

we call “computational problem solving.” In contrast to scientific problem solving, the emphasis of computational problem solving is not based on the science but rather on the effective and in situ operation of computational resources that is required in applying the science. We now describe a few of the tools in the climate CPSE prototype that fall under this category.

4.1 Calculation Table

Climate researchers think about their experiments and models in terms of the calculations they need to run. Prior to executing a climate model, the researcher organizes his investigation as a set of calculations that may be partitioned into different combinations of geographical regions, time periods, earth properties, and parameters for physical processes. Defining a set of calculations allows the climate researcher to break down the experiment into smaller, more manageable pieces and be able to allocate the computational work across distributed and parallel machines. Furthermore, the set of calculations may also serve to identify and manage the division of labor among collaborating researchers. Thus, a group of collaborating climate researchers might partition the set of calculations among group members as individual tasks.

In the CPSE prototype, the Calculation Table allows climate researchers to define, view, and execute a set of calculations while hiding as many of the details as possible of the underlying computer infrastructure and network. As shown in Figure 4, the Calculation Table maintains a dynamic list of climate calculations that are identified by key properties such as geographical location, grid size and resolution, and time period of simulation. Operational properties of the calculations are also maintained such as the machine on which the calculation is to run, the time of scheduled execution, and the project member who is assigned to execute and manage the calculation. Information is inserted into the Calculation Table through a calculation entry form, or the Calculation Table may be fed by a separate graphical user interface (GUI) application that is tied to the climate model. In the case of the latter, the Calculation Table stands as an intermediary between the GUI and the climate model—diverting model input parameters into the Calculation Table for follow-up execution.

4.2 Spatial and Temporal Data Browsers

Current climate models produce large, multi-file data sets that may be in the order of terabytes to petabytes in size. In applying the data, climate modelers must deal with intricate details such as accessing data archives and systems, retrieving large volumes of data, collecting sets of files into an organized view, transferring or streaming data to different locations, and converting both standard and application-specific data formats. Thus, climate researchers currently work with climate data from the view of data formats, data organization across files, file systems, and network topographies. With the climate CPSE prototype, we sought to shift the climate researcher’s view of data from a structural or mechanical one to a more conceptual one.

Fig. 4. Climate CPSE tools provide high-level access to computational resources. The Calculation Table manages a list of calculations for a particular experiment. The Spatial Data Browser and Temporal Data Browser allow climate researchers to select data based on spatial and temporal properties. The Data Pedigree tool allows climate researchers to view and access the historical lineage of a data set.

Consistent with the way they define their calculation sets, climate researchers view and apply climate data largely along dimensions of time, space, and earth properties. In addition, the selection of data is usually preceded by some high-level analysis of the data space. To accommodate these usages, the conceptual prototype supports two types of data views as shown in Figure 4. From a geographical or spatial perspective, the Spatial Data Browser allows climate researchers to view, select, and manipulate data and data files in 2D space by selecting regions of a map or image. As directed by climate researchers, a spatial map is overlaid with information, attributes, or contours that illustrate different earth properties. The climate researcher examines features and patterns in the map to identify regions from which the actual data should be extracted. This extraction of data is particularly important for impact assessments as they are typically performed over smaller regions that are defined by climate regimes, watersheds, basins, eco-regions, or farms. From a temporal perspective, the Temporal Data Browser allows climate researchers to view, select, and manipulate data and data files from different parts of a timeline. The climate researcher graphs the values of different earth properties over time. Based on an analysis of the graph, the climate researcher may select a time period over which climate data should be retrieved.

4.3 Data Pedigree Management

As we previously described, the Execution History tool in the climate CPSE prototype allows the climate researcher to navigate through previous instances of models, applications, and other computational resources. Similarly, the climate researcher may also wish to navigate through and revisit previous versions of a data set. The climate researcher evolves and iterates a data set over the course of experimentation. The ability to manage large, complex data sets becomes more difficult over time as more experiments and runs take place. Thus, an automated tool for managing historical versions of data sets would provide a valuable service.

As shown in Figure 4, we introduce the concept of a Data Pedigree tool that provides data set versioning capabilities. The Data Pedigree tool provides a hierarchical tree depicting the “pedigree” or the historical lineage of a climate data set. Each branch in the tree represents a particular path of evolution for the data set. From the tree view, the climate researcher may select a particular node to access a specific version of the data set. Alternatively, the climate researcher may select an entire branch and apply various comparison functions to evaluate how a data set may have changed over time. Overall, the Data Pedigree tool will be useful to researchers for diagnosing the source of an error or anomaly in a data set by providing access to the data set’s full history.

5 Knowledge Management and Dissemination

Scientific problem solving is inherently a collaborative effort among researchers as they share information, models, tools, resources, and results. More than just sharing specific research artifacts, however, scientific problem solving also involves the sharing of one’s expertise and experiences. As climate researchers run models, they apply a vast amount of procedural and domain knowledge. Climate researchers may have much experience in running particular climate models—understanding the nuances and problems of those specific models. Researchers may also have a lot of knowledge of climate concepts and theories that serves as the basis of their research and experimentation. The ability to capture these kinds of knowledge and share them with others is yet another potential enabling feature of a climate CPSE.

To a certain extent, climate researchers already have mechanisms in place for capturing procedure and domain knowledge. Depending on the climate researcher’s particular research style and specific experimentation needs, the researcher may jot rough notes down on a notepad or place long, detailed descriptions in a laboratory notebook. In the climate CPSE prototype, we have provided additional, more advanced knowledge capture and dissemination mechanisms.

5.1 Electronic Laboratory Notebook

As shown on the left side of Figure 5, an Electronic Laboratory Notebook [11] may be used as a knowledge capture tool much like a physical laboratory notebook is applied.

The climate researcher may store theoretical and procedural information in the Electronic Laboratory Notebook during the course of a project. The amount and level of detail of the information stored is under the control of the climate researcher.

An electronic version of the laboratory notebook may have additional features that allow the climate researcher to better manage, locate, and share scientific information. For instance, an electronic version may extend the capabilities of a laboratory notebook by allowing climate researchers to capture information in many different forms (e.g., text, drawings, tables, images, audio files, video files), share content by allowing researchers to access the same notebook from multiple machines, utilize standard security and authentication mechanisms to ensure valid access, and employ indexing and searching mechanisms that allow the researchers to easily and quickly locate specific information. These are among the capabilities that make Electronic Laboratory Notebook attractive as an integrated knowledge management tool.

Fig. 5. A dynamic collaboration environment allows real-time collaboration to take place within the context of the climate researcher's experimental process. In the above scenario, the climate researcher works within the climate CPSE, using a workflow and visualization tool. When the climate researcher requests collaboration with project members, the collaboration environment establishes video connections with on-line project members and provides immediate sharing of the workflow and visualization windows. The collaborators may also bring up shared notes in the Electronic Laboratory Notebook and through the Annotation tool.

5.2 Free-Form Annotation

As climate researchers conduct their experiments, they commonly jot down miscellaneous notes pertaining to their current research state such as the line of investigation being followed, the input parameters to the current experiment, key findings and results, and encountered error conditions. The notes or comments that are made may have a temporal, spatial, or referential context. For instance, they may be associated with a particular data set, a phase or step of the experiment, a specific result of the experiment, or a particular resource such as a computer or data archive. In the climate CPSE prototype, many of these associated entities have visual representations. A free-form annotation mechanism would allow climate researchers to enter annotations and link them directly to the their visual referents.

As shown in Figure 5, the Annotation tool allows climate researchers to annotate any design or resource representation. Annotations for any particular object are collected over time. Besides text, we also envision that climate researchers may also create audio and video annotation entries. Once an annotation entry has been created, collaborators with authorized access may view or replay the annotation. The annotation mechanism provides a means for capturing and sharing contextual knowledge. A climate researcher may review the annotations of other collaborators to comprehend the concepts, usage, or problems associated with any particular entity. Through the annotation mechanism, contextual knowledge is embedded into the CPSE and expands over time as more annotations are entered. In due course, the CPSE matures—gaining more knowledge and experience through continual use.

5.3 Computer-Mediated Collaboration

Scientific collaboration does not occur in isolation but is driven by the functions of the scientific research. From our analysis and design studies, we found that climate researchers collaborate at key points during the research process to define the scientific problem and hypotheses to attack, partition research tasks, compare and analyze results, and conduct other expected collaborative activities. Climate researchers also collaborate at random points during the research process to troubleshoot problems, seek additional expertise, compare and validate intermediate results, and carry out other collaborations that are necessitated by unforeseen events occurring during the research. In both situations, collaboration is a natural and dynamic function of the research activity—not an independent, preplanned act.

In the climate CPSE prototype, collaboration is instantiated in the context of the project and the experiment. For example, in the scenario depicted in Figure 5, a climate researcher executes a particular climate model and brings up a visualization of intermediate results. Seeing an unexpected result, the climate researcher initiates a real-time discussion with other on-line project members—video conferencing starts and the executing workflow and the visualization tool become collaborative. During the discussion, collaborating researchers share notes through the Annotation tool and parameters, procedures, and results through the Electronic Laboratory Notebook.

When the conversation ends, the collaboration terminates, and the initial climate researcher returns to previous work.

As illustrated, collaboration capabilities are initiated from within the CPSE rather than from a separate, non-related environment. The collaboration tools and sessions are instantiated rapidly, dynamically, and within the context of the research activities.

6 Summary

The climate CPSE we have described in this paper is currently in the form of an early prototype with very limited functionality. It will be incrementally developed in conjunction with an underlying scientific computing infrastructure using a combined top-down, bottom-up development strategy. The strategy involves simultaneously evolving user and system capabilities until these capabilities converge to a unified, fully functional system. Emerging scientific computing infrastructure capabilities include services for scientific data transport and management, intelligent job launching and control, application integration, security and authentication, and event management. The overall two-pronged development approach is mutually beneficial, because the high-level prototype provides user and system requirements to the infrastructure development effort while the infrastructure work provides features and imposes technology constraints on the development of high-level features and tools.

Climate researchers live in a world of increasing complexity. With more advanced models, faster, more powerful computers, and higher-capacity data archives—all distributed across high-speed networks—researchers have a growing capability to run more sophisticated simulations and solve more challenging scientific problems. Unfortunately, many climate researchers are unable to harvest the vast computational power afforded to them. Rather, the capability is relegated to those select few who are able to comprehend and navigate the complicated computational environment.

In conceptualizing and designing a climate CPSE, we aim to abstract the science of climate modeling and impact assessment as much as possible from the underlying computational infrastructure. We wish to pierce the technological barrier by allowing those who are familiar with the computational environment the capability to execute their experiments more productively and efficiently while providing those who are not the very opportunity to exercise their scientific abilities using advanced capabilities and the latest inventions. For us, the operative term in the creation of CPSEs is “problem solving.” Our ultimate goal is to provide climate researchers with greater ability to solve their scientific problems. Thus, beyond allowing climate researchers to execute their models easier and faster, CPSEs should meaningfully and fundamentally transform and enhance the science and the practice of climate modeling and impact assessment as well as energize and empower an entire community of climate researchers to achieve greater scientific heights.

References

1. Advanced Visual Systems: AVS/5. Advanced Visual Systems, Waltham, MA 02154, USA (2000)
2. Chin, G., Schuchardt, K., Myers, J., Gracio, D.: Participatory Workflow Analysis: Unveiling Scientific Research Processes with Physical Scientists. In Proceedings of Participatory Design Conference (PDC 2000), (Nov. 28-Dec. 1, New York, NY), CPSR, Palo Alto, CA (2000) 30-39
3. Greenbaum, J., Kyng M.: Introduction: Situated Design. In Greenbaum, J., Kyng M. (eds.): Design at Work: Cooperative Design of Computer Systems. Lawrence Erlbaum Associates, Hillsdale, NJ (1993) 1-24
4. Jones, D.R., Gracio, D.K., Taylor, H., Keller, T.L., Schuchardt, K.L.: Extensible Computational Chemistry Environment (ECCE). In Fayad, M.E., Johnson, R. (eds.), Domain-Specific Application Frameworks: Frameworks Experience by Industry. Wiley & Sons, New York (2000)
5. Khoros Research Inc.: Khoros Pro 2000 Version 2. Khoros Research Inc., Albuquerque, NM 87110-4142, USA (2000)
6. Leung, L.R., Ghan, S.J.: Pacific Northwest Climate Sensitivity Simulated by a Regional Climate Model Driven by a GCM. Part I: Control Simulation. *J. Clim.* 12 (1999) 2010-2030
7. Leung, L.R., Ghan, S.J.: Parameterizing Subgrid Orographic Precipitation and Surface Cover in Climate Models. *Monthly Weather Review*, 126(12) (1998) 3271-3291
8. Leung, L.R., Hamlet, A.F., Lettenmaier, D.P., Kumar, A.: Simulations of the ENSO Hydroclimate Signals over the Pacific Northwest Columbia River Basin. *Bull. Amer. Meteorol. Soc.* 80(11) (1999) 2313-2329
9. Leung, L.R., Wigmosta, M.S., Ghan, S.J., Epstein, D.J., Vail, L.W.: Application of a Subgrid Orographic Precipitation/Surface Hydrology Scheme to a Mountain Watershed. *J. Geophys. Res.* 101 (1996) 12803-12817
10. Leung, L.R., Wigmosta, M.S.: Potential Climate Change Impacts on Mountain Watersheds in the Pacific Northwest. *J. Amer. Water Resour. Assoc.* 35(6) (1999) 1463-1471
11. Mendoza, E.S., Valdez, W.T., Harris, W.M., Auman, P., Gage, E., Myers, J.D.: EMSL's Electronic Laboratory Notebook. In Proceedings of the WebNet '98 World Conference, (Nov. 7-12, Orlando, FL), (1998)
12. Muller, M.J., Wildman, D.M., White, E.A.: 'Equal Opportunity' PD Using PICTIVE. *Communications of the ACM*, 36(4) (1993) 54-66
13. Patrinos, A.: The Accelerated Climate Prediction Initiative: Bringing the Promise of Simulation to the Challenge of Climate Change, [Online report]. Available URL: <http://www.epm.ornl.gov/ACPI/Documents/ACPIfinal.html>, (1998)
14. Schuler, D., Namioka, A. (eds.): Participatory Design: Principles and Practices. Erlbaum Associates, Hillsdale, NJ (1993)
15. US Global Change Research Program: The Potential Consequences of Climate Variability and Change: Foundation Report. Cambridge University Press (2000)
16. Yourdon, E.: Modern Structured Analysis. Prentice Hall, Englewood Cliffs, NJ (1988)