iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Step by step calibration of an Integrated Model System for irrigation management

T. Arnold [a], **T. Berger**[a], **H. Uribe**[b]

[a]*Universität Hohenheim (490d), 70593 Stuttgart, Germany (tarnold@uni-hohenheim.de)*

[b]*INIA Quilamapu, Chillan, Chile*

Abstract To assess how the multiple processes that constitute irrigation land use interact, the basin-level hydrological model WASIM-ETH and a bioeconomic model MP-MAS were embedded into a common framework. Following conceptual integration of theories, we here describe the integration of data and joint calibration and validation of both models. Methodologically, interacting variables are first specified from data, then from model outputs, and then dynamically coupled. Interdisciplinary cross-checks and sensitivity analyses improved calibration. For irrigation management at basin scale, we indicate that physical scarcity and allocative scarcity coexist in one basin. In *a*heterogeneous population of farmers, constraints vary considerably so that effective policies require precise definition and in-depth analysis of the target groups at individual level.

Keywords: Catchment scale; Irrigation; Multi-agent System

1 INTRODUCTION

A comprehensive scientific theory of agricultural land use change and resource use (land, water) does not exist. Rather, a multitude of different theories coexist - not only amongst disciplines, but also within disciplines. Phenomena in the real-world are the product of multiple concurrent processes; ecological, soil, meteorological and hydrological conditions constrain the decisions of farmers, who must also deal with market conditions, institutional constraints and incentives. In addition, their own knowledge base also has impact on what is, cumulatively, land use change [Parker et al., 2003]. Multiple processes interact at different scales, each examined by different disciplines and institutions. This produces an abundance of theories and alternative explanations, but data tends to be scarce and disconnected.

The integration of processes into a single framework (and model) poses at least three types of conceptual challenges. The first challenge is the conceptual integration of theories, as a basis for the coupling of models and data, to generate scientifically sound computer code. This challenge includes formal descriptions of exchanged variables, inconsistencies in model assumptions across model components, and sequencing [Argent, 2004]. The second type of challenge, addressed in this paper, is the process of bringing data and integrated models together; the simultaneous calibration and validation of models with one data base, in order to understand interaction. A third challenge is epistemological, questioning under which conditions the resource-intense integration of models can actually generate new knowledge. Beven [2007] relates it to the context of a case study, as well as to the organization of the knowledge-generating research process.

1.1 The project 'Integrating Governance and Modeling'

The integrated model system presented here is being developed within the project *Integrating Governance and Modeling*, a CGIAR *Challenge Program on Water & Food* project. Its objective

is to explore policy options that improve the management of water resources at both the local and the regional levels, aided by integrated computer models that resolve the micro-scale. Furthermore, this project analyzes existing governance structures and their ability to distribute water equitably and efficiently (Project website http://www.igm.uni-hohenheim.de/igm).
The case study in Chile assessed challenges in water management through a participatory approach with key stakeholders, and identified policy options to address these challenges, along with policy evaluation criteria. Based on this, model use cases are being developed jointly with farmer organizations and with governmental institutions. An interdisciplinary data base was compiled, which combines GIS data, socio-economic data from census and farm surveys, crop production data, plant data, canal data, and registries on land use and water rights. Under a (semi-)predictive modeling paradigm, a model system was built that integrates the basin-scale distributed hydrological model WASIM-ETH [Schulla and Jasper, 2007] and a bio-economic, agent-based model MP-MAS used for agricultural water use analysis [Berger et al., 2007]. To study the relevance of interactions, models can either be used and calibrated separately [Uribe et al., 2008], or dynamically-coupled [Arnold et al., 2007].
The study area contains the watersheds of the Putagán, Ancoa, Achibueno and Longaví rivers, in the Maule Region of Chile. At a total area of 5300 km^2, this region contains approximately 100.000 hectares of agriculturally used lands, mostly with snow melt from the Andean mountains through a complex distribution system of canals. Water is distributed from river organizations, to smaller user organizations summarized as irrigation sectors, and finally to farmers, according to the water rights system of Chile.

1.2 Integrated modeling: A calibration and validation challenge

The data base of our case study contains maps, constant technical parameters, economic cross-sectional data for two time periods (1996/7 and 2006/7), time series on price- and cost developments, and hydrological and meteorological measurements (time series). Qualitative expert opinions on the transient period exist, but are not always coherent. The analytical challenge is to find a theory-based explanation that is consistent with data for 1996/7 and 2006/7, with transient time series, and which matches expert opinion as well as possible.
Alternative model formulations can lead to the same outcomes ('equifinal models', Beven [2001]), or they can look equifinal because outcome aggregates and coarse empirical data conceal micro-scale logical fallacies. Here, the integration of evidence from empirical data and from theories across disciplines and at multiple scales can greatly reduce uncertainty between alternative explanations, by proving hypothesis as inconsistent or incomplete.
The objective of this publication is to demonstrate the value of an integration framework as methodology, rather than offering policy implications. All data interpretation is still preliminary.

2 METHODS

2.1 The Integrated Modeling System (IMS)

The bio-economic, agent-based model MP-MAS describes and spatially captures land use decisions of farmers [Berger, 2001]. At individual level, a constrained-optimization model mimics production and farm investment decisions, and the resource use involved (land, water, credits, labor). *Inter alia*, modules include land and water markets, soil nutrient dynamics, diffusion of innovation in networks, water sharing within irrigation sectors, and a lumped routing model between sectors. At yearly time steps, MP-MAS recursively updates the asset endowment of farmers, so that aspects of heterogeneity and distribution can be analyzed.
The WASIM-ETH is a process-based, fully-distributed hydrological basin model [Schulla and Jasper, 2007]. For each grid cell, vegetation cover can be parameterized (single or multi-layered). Processes include interception, surface evaporation, infiltration into the top soil layer, surface

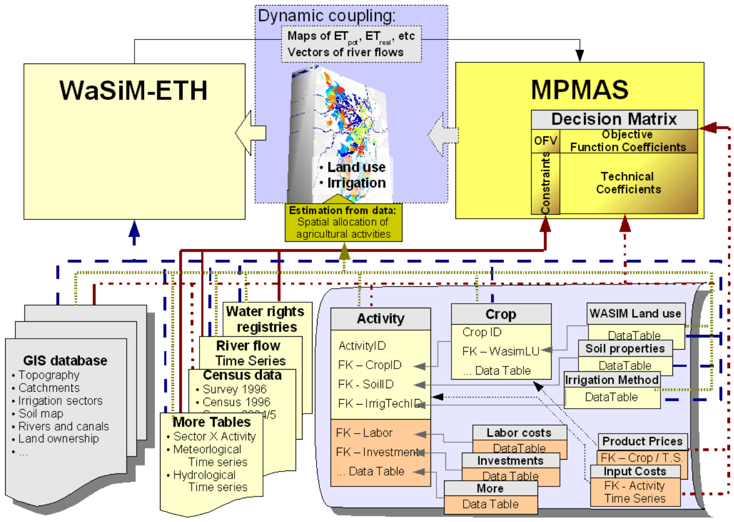

Figure 1: An integrated model system that allows standalone and coupled model runs

runoff. The model also includes an unsaturated 1D-vadose zone module, based on Richards' equation, in which plant transpiration is withdrawn while limited by ET_{pot}. Excess water percolates. Groundwater flow can be parameterized, or dynamically modeled with a 3-dimensional advection model. The representation of above-ground vegetation characteristics and the vadose zone make WASIM-ETH a good choice for modeling irrigation. Routing of surface water is based on a sub-basin approach, derived from a topographic analysis. Water channeled between sub-basins is externally parameterized as extractions, inflows or bypasses. The original irrigation module was extended by our working group [Arnold et al., 2008].
Both models MP-MAS and WASIM were integrated conceptually and technically (figure 1) and linked to an interdisciplinary data base. In standalone mode, models can be calibrated separately by domain experts. During this calibration phase, inconsistencies within the shared data base were identified and eliminated. The framework also allows for the dynamic coupling of the components, such that agricultural land use is computed from the socio-economic model at a yearly time step and reported to the hydrological model through a translation interface. For the monthly decision on irrigation quantities and its abstraction, the model is run twice for each month: First, WASIM estimates crop water demand and water availability *without* irrigation. Then, the economic model develops a monthly irrigation plan and reports it to WASIM. Agricultural production is continued within MP-MAS and the hydrology in WASIM. After each month, maps (e.g. real evapotranspiration) are reported to MP-MAS. The translation interface consists of a hierarchical sequencer and spatial and temporal rescaling functions, partly drawing on the relational data base. Data is passed between applications using the Typed Data Transfer library. Technical implementation has been finalized and the software code has been technically verified by showing that outputs from a reconstructed standalone version corresponds to coupled results [Arnold et al., 2007].

A key conceptual integration challenge is the cyclic reuse of irrigation water. Physical-based modeling requires micro-scale data (soil characteristics, methods of irrigation, drainage, scheduling, and the characteristics of the canal system). Hansen et al. [2007] discusses coupled approaches, and concludes that the problem of micro-level heterogeneity for irrigation has not been resolved. Within our IMS, we apply dual concepts [Arnold et al., 2008]: In WASIM, cyclic reuse within sub-basins is eliminated and *effective* irrigation efficiency η_{eff} is used. MP-MAS uses field-level efficiency η_{field} and estimates the reuse fraction within the EDIC module. This duality facilitates the modeling of water as an economic constraint for production, while maintaining a basin-wide perspective and ensuring consistency at all scales. Return flows from one sub-basin to another can

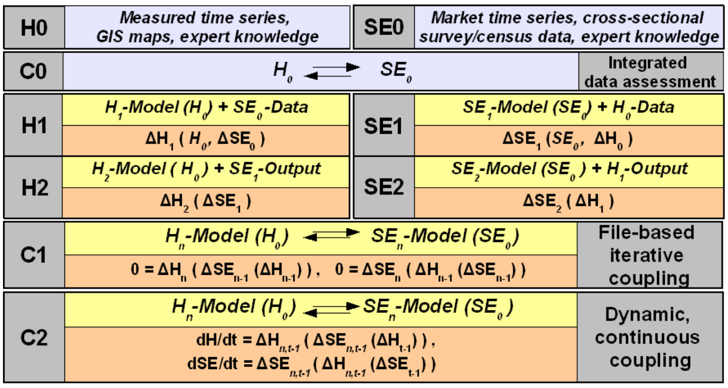

Figure 2: Stepwise calibration and validation

use the natural path, but also artificial canals, so modeling remains problematic.
Within the MP-MAS model, the handling of space was revised, and data handling improved was improved. The internal time stepping of the code was restructured, which required further modularization.
Within WASIM-ETH , the original irrigation module was extended. Potential evapotranspiration can now be computed with a crop-coefficient. Irrigation water can be applied as precipitation, but also directly to the surface. Access water may be directly applied as runoff, or as percolation through parameterization, and scheduled rotational schemes are avoided. Routing and the irrigation restriction was also modified [Arnold et al., 2008].

2.2 Framework for stepwise calibration

Three stages in the calibration and validation of our coupled model, which includes a hydrological component H and a socio-economic component SE, can be distinguished: (1) The integration of interdisciplinary data and consistency tests on data, (2) the calibration and validation (C&V) of each component H and SE with that data, and (3) the coupling C of calibrated models - first file-based, then dynamic. At each step, inconsistency might require revisiting concepts, further calibration or data improvement.
Both modules H and SE build on data, on expert knowledge, and on 'boundary conditions' from the other domain ('coupled variables'), which are uncertain and mutually dependent.
After data from both domains (H_0 and SE_0) is collected, the first step is the creation of a common information base (C_0). Already, this step requires full conceptual integration. Using C_0, baseline scenarios of coupled variables are then defined, as boundary conditions for standalone calibration. Both domains benefit from the exchange of data and expert knowledge. Within the usual, discipline-specific model analysis of that base line, the sensitivity analysis to the coupled variables gives $\Delta H_1(H_0, \Delta SE_0)$ and $\Delta SE_1(SE_0, \Delta H_0)$. When assessing biophysical-economic interaction, the team can focus analytical and data collection efforts to improving those coupled variables with a strong influence on model outcome indicators.
In the second step (H_2 and SE_2), boundary conditions are created from model results. Logical consistency (correlation structures) are then preserved and finer time resolution makes interpolation unnecessary. Recalibration improves models and model sensitivity $\Delta H_2(H_0, \Delta SE_1)$ and $\Delta SE_2(SE_0, \Delta H_1)$. Procedurally, steps H_2 and SE_2 require revisiting the conceptual integration and data. Technically, data exchange between both models is clarified, data formats are revisited, and data translation functions required for coupling become apparent and can be automatized.
The third coupling step $C1$ is a linear, file-based coupling of models, built on automatized data translation routines. Models are rerun and outputs are iteratively exchanged $SE_n(H_{n-1}), H_n(SE_{n-1})$, until convergence is reached ($H_n - H_{n-1} \approx 0$). Eventually, both

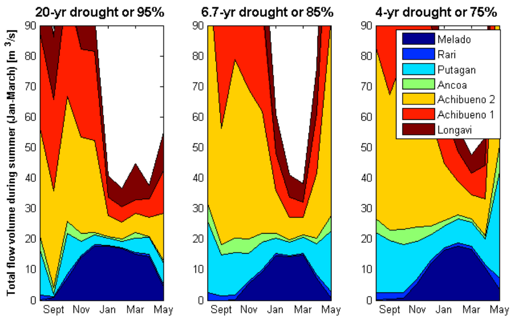

Figure 3: River flows at seven stations in dry years (20-year, 6.7-year, and 4-year drought).

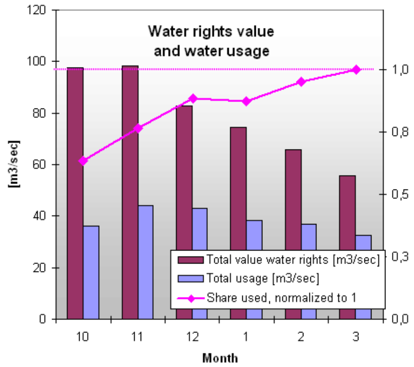

Figure 4: Relative share of water rights used, normalized to February share.

models are re-calibrated. It is recommendable to choose the time horizon such that the complexity of interactions remains reasonable. This step already allows for full validation against observed data. As an example, we analyze the impact of a new technology that is used by upstream farmers. What is the first-order impact on downstream farmers (first iterative step), and how will the adaptation of downstream farmers further modify the hydrology at convergence?
Finally, dynamic coupling allows for the assessment of the speed of change. Here, the continuous exchange of coupled variables between the two well-calibrated and validated models allows for direct interpretation of model scenarios. Using the above example, after upstream changes of farming practices, with what speed do hydrological system change and farmers adapt?

3 Step-wise model calibration and validation

With an integrated and technically verified coupling setup, and finalized data collection and data integration at appropriate scales, the dependencies of coupled variables were analyzed. Using indicative results, the step-wise and integrated calibration and validation procedure is outlined.

3.1 Data analysis for variations in hydrology ΔH_0

The following analysis integrates river flow time series, data on water rights registries, data on agricultural land use by irrigation sector, and technical coefficients on land use activities, as compiled in an interdisciplinary data base (fig. 1).
As input to SE_1, time series of river flow stations are analyzed for average and dry years (4-, 6.7-, and 20-year frequency, in which water flows exceed 75%, %85 or %95 of the time respectively). Monthly water availability fluctuates strongly over years, especially during the month of January (fig. 3). During Jan-Feb of "normal" droughts, e.g. during El Nino years, access to the Melado canal, fed from a neighboring catchment, buffers shortages. During extreme droughts, this dependence starts in Nov-Dec.
As an indicator for physical water scarcity per sector, the percentage of available water is calculated by multiplying agricultural land use data [ha/sector] at irrigation sector level with estimations of water requirements for those land uses [$\frac{\text{liter}}{\text{ha}}$]. Results are divided by water availability, calculated with water rights data for each sector [$\frac{\text{units}}{\text{sector}}$] times the value of these rights [$\frac{\text{liter}}{\text{unit}}$]. Results show full usage only in the downstream sector 04-f, which also receives water not used upstream. During average years, all other sectors under-exploit water resources. Possible explainations include over-supply, risk-aversion to droughts, and scarcity due to misallocation of water rights among farmers. The estimated change of water requirements in sectors 04-l and 05-e from 1997

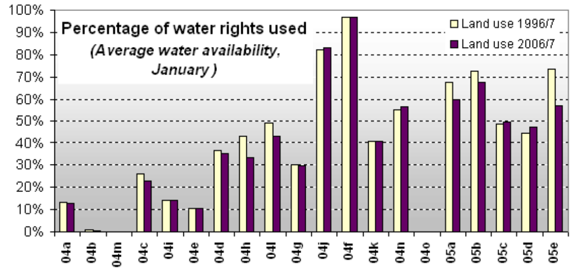

Figure 5: Usage of water rights during average January

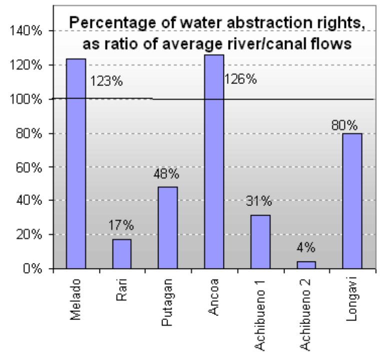

Figure 6: Ratio of river flows to equivalence value of added water rights.

to 2007 reflects the reduction of rice production, in line with falling prices in 2000 (fig. 5).
For MP-MAS , water rights are expressed as the percentage of river inflows. The irrigation sectors that suffer from the strongest annual fluctuations of inflows are the large rice-producing sector 04-l, which satisfies 20% of January water demand from the Melado canal, sector 04-f ($< 2\%$ Melado water) and the large 04-g ($\approx 45\%$ Melado water), while access to anti-cyclic water supplies (Melado, and to some extend the Longavi River) buffers drought conditions.

3.2 Sensitivity of MP-MAS to variations in hydrology - $\Delta SE_1(\Delta H_0)$

The economic model MP-MAS was used to assess the impact of water variability on individual farm households. Based on empirical data, we modeled farm decisions using Mathematical Programming, and validated results interactively with 10 selected farmers. A statistically consistent population of farmers was then created, and Schilling [2007] estimated shadow prices for water for different years. Shadow prices for water (SPW) range from \$Peso/m^3 $0.0 - 60.0$ and a maximum of $\$/m^3$ 120.0. During droughts, the percentage of farmers with low or very low SPW falls from 55% during normal years to 15%. These results clearly highlight that not only water availability, but also further constraints, determine water usage by farmers.

3.3 Data analysis for variations in economic impacts ΔSE_0

To model irrigation at basin scale, while maintaining a field-level perspective, cyclic water reuse needs to be captured properly. It is invisible at macro scales, and data scarcity requires a parametrization approach for return flows. Using the WASIM-ETH specification for 1996/7 land use [Uribe et al., 2008], the net (non-cyclic) abstraction of irrigation water was estimated at $\approx 38.2\frac{m^3}{sec}$, and a (net) irrigation table of $IRR_{\text{eff}} = 3.76\frac{\text{mm}}{\text{day,ha}}$.
The routing of irrigation water between sub-basins through canals was initially estimated from water rights. Yet, analysis of land use water requirement suggests that only during the peak month of February is the full routing realized, in other months only shares are utilized (fig. 4). Further, for January, a comparison of river/canal flows with the value of water rights flows reveals that abstraction rights exceed flow for both the (water-scarce) Ancoa river and the Melado canal (fig. 6). The additional 23% are explained as return flows, from irrigated fields back into main canals/rivers, which conforms closely to the 22% estimated by MOP [1992].
Using EDIC equations and land use data, field-level data for irrigation methods and their efficiency η $[\frac{\text{mm ETP}}{\text{mm applied}}]$ was aggregated to *effective* efficiencies (without direct reuse, but with return flows) to bridge field-scale and basin-scale.

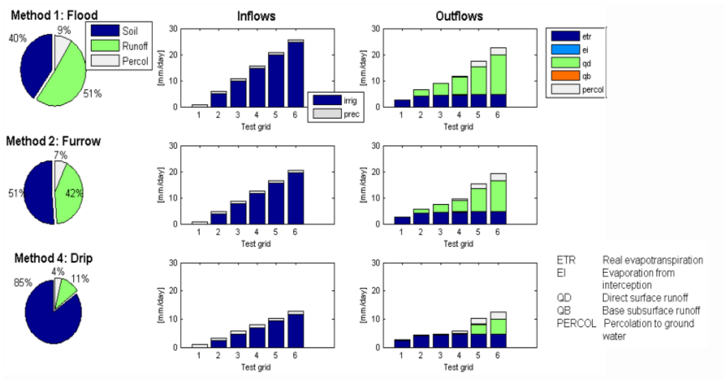

Figure 7: Sensitivity to flow component at grid cell level to irrigation methods, and irrigation table (at grid cell level)

3.4 Sensitivity of WASIM-ETH to variations in irrigation - $\Delta H_1(\Delta SE_0)$

The initial calibration gave a Pearson correlation of $R^2 > 0.9$ for large sub basins and $R^2 > 0.55$ for small sub basins [Uribe et al., 2008], before incorporating the above data. Grid-scale analysis revealed that even with acceptable reproduction of flow station measurements, relevant internal model processes were not properly represented. The heterogeneity of soil water dynamics and data scarcity for ground water dynamics, limits the physical scope of this study (see also [Hansen et al., 2007]). The soil module of WASIM is currently recalibrated at micro-scale, using the extended WASIM-ETH irrigation module with effective irrigation efficiencies, applying water directly to top-soil and using parameterized return flows and percolation. Figure 7 shows model experiments where otherwise homogeneous cells receive different irrigation tables, according to three different methods.

4 DISCUSSION OF PRELIMINARY RESULTS

In accordance with local knowledge, the analysis shows that hydrological constraints are economically relevant only during the summer months (Jan - March). Economic analysis shows that a basin perspective on absolute water scarcity does not suffice to explain why high-value irrigation agriculture (fruits & vegetables) is not expanding further. In average years, the total amount of water utilized is below 65%. Nevertheless, high shadow prices, especially during drought conditions, show that a significant portion of the farm population is constrained by water. The hydrological model requires a consistent representation of surface evaporation, plant water uptake from root zones, interflows and surface runoff at grid scale. Furthermore, the conceptual treatment of return flows, such that irrigation dynamics at the basin scale are correctly represented and properly understand, is also necessary.

Ongoing analysis focuses on consistency checks of data with local experts and on calibration of both the hydrological and the economic model, first at grid (and agent) scale, with grid (and agent) scale parameters, then at meso-scale, with sector-level parameters.

5 CONCLUSION

A calibration and validation framework for integrated, empirical and (semi-)predictive modeling was demonstrated using a case study. The integration of theories and concepts, of models and model source code, and ultimately of data, was shown to be dependent on and mutually beneficial for each other. Reasoning that builds on integrated concepts and data at respective scales reveals inconsistencies in otherwise behavioral model specifications, which are thus inadequate for in-depth or predictive analysis. For irrigation management at the basin scale, we indicate that physical scarcity and allocative scarcity coexist in one basin. In a heterogeneous population of farmers, constraints vary considerably. Therefore, effective policies require precise definition and in-depth analysis of the target groups. As a general note, this paper emphasizes that integrated modeling at the basin scale needs to be seen as a co-evolution of concept development beyond disciplinary boarders and theoretical, computational and data management aspects. This supports the vision of Beven [2007] for model tool boxes to be learning tools, to relate data, to identify knowledge gaps, and to value well-targeted monitoring for new information.

Acknowledgements

We thank the CGIAR Challenge Program on Water & Food and Alechile (DAAD/CONICYT) for financial support.

References

Argent, R. M. An overview of model integration for environmental applications—components, frameworks and semantics. *Environmental Modelling & Software*, 19(3):219–234, March 2004.

Arnold, T., H. Uribe, and T. Berger. Description of a multi-purpose model system for policy planning support, Nov 2007. Presented at International Conference on Adaptive & Integrated Water Management (CAIWA); Basel (Switzerland) 12.-15. Nov 2007.

Arnold, T., H. Uribe, and T. Berger. A distributed hydrological model of irrigation and agricultural land use at the basin scale: Application to the viith region chile, April 2008. Accepted for Conference "21st Century Watershed Technology (ASABE)".

Berger, T. Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. *Agricultural Economics*, 25(3):245–260, 2001.

Berger, T., R. Birner, J. Díaz, N. McCarthy, and H. Wittmer. Capturing the complexity of water uses and water users within a multi-agent framework. *Water Resources Management*, 21:, 129–148, 2007.

Beven, K. A manifesto for the equifinality hypothesis. *Journal of Hydrology*, 320:18–36, 2001.

Beven, K. Towards integrated environmental models of everywhere: uncertainty, data and modelling as learning process. *Hydrological & Earth System Sciences*, 11(1):460–467, 2007.

Hansen, J., J. C. Refsgaard, and V. Ernstsen. Problems with heterogeneity in physically based agricultural catchment models. *Journal of Hydrology*, 342(1-2):1–16, 2007.

MOP. Estudio de factibilidad del canal linares. Report, volumen 1 y anexo, Consorcio EDIC-CEDEC for Ministerio de Obras Públicas, Dirrección de Riego, Chile, 1992.

Parker, D., S. Manson, M. Janssen, M. Hoffman, , and P. Deadman. Multi-agents systems for the simulation of land-use and land-cover change: A review. *Annals of the Association of American Geographers*, 93(2):314–337, 2003.

Schilling, C. Interactive modeling of farm-level response to the ancoa dam project, chile. Master thesis, University of Hohenheim, Stuttgart, Germany, 2007.

Schulla, J. and K. Jasper. Model description wasim-eth, new version. Internal report, IAC, ETH Zürich, Switzerland, 2007.

Uribe, H., T. Arnold, J. Arumí, T. Berger, and D. Rivera. Aplicacion del modelo wasim-eth para el estudio de la hidrologia en un sistema de riego, 2008. submitted.