Handling and linking data and hydrological models – experiences from the Danish national water resources model (DK-model)

A. L. Højberg, L. Troldborg, P. Nyegaard, M. Ondracek & S. Stisen

Geological Survey of Denmark and Greenland (GEUS), 1350 Copenhagen, Denmark

ABSTRACT: A national water resources model (DK-model) has been established covering the entire 43.000km^2 of Denmark. The model is presently updated including a revised hydrological interpretation based on existing local and regional models, a refinement of the model setup and resolution and a more detailed description of model input data. The present model update will be finalized by 2009, but visions are that the model will be updated regularly in the future, based on new detailed studies. Formulation of a nationwide hydrological model provides the opportunity to integrate all hydrological data in a consistent framework, i.e. linking raw data in databases to a conceptual understanding of the physical system. Such relation between data and the physical system is required for water management, particularly in European where water status must be reported at the level of (ground-)water bodies. Coupling models and databases are additionally beneficial for rapid model update by new input data, and for utilizing the model for quality assurance of data. During the model update, much emphasize have therefore been given to link the model to national databases with data relevant for hydrological modeling.

1 INTRODUCTION

Monitoring data are essential to water resource management to monitor the quantitative and qualitative status of the aquatic system. Data themselves do, however, only represent point observations of the current state of the system at the point and time where the measurement was taken. To extract valuable information from monitoring data they must be analyzed on the basis of a conceptual understanding of the physical system observed. Models have proven valuable tools to integrate the various data and link them in a conceptual framework. For model constructions descriptive data on the physical system are required. For hydrological modeling such data include geological and geophysical data, river network descriptions, soil physics, land use, farming praxis etc. Using data actively in modeling, errors or inconsistencies in data and databases may be revealed, which may not be identifiable by standard quality checks on the databases alone. Combining data with modeling thus have several advantages (Højberg et al., 2007; Jørgensen et al., 2007). A prerequisite for the combined use of models and data are, however, an optimal dataflow between databases and the models.

Denmark has, supported by adequate legislation, developed comprehensive national databases storing geo-related data and monitoring data. By 2007 the national databases have been supplemented by data previously stored by regional water authorities, completing the national databases to include all data relevant for hydrological modeling. The storage of data in central databases makes it very attractive to develop standardized protocols for data extraction and processing. Such protocols will ensure that data are correctly extracted from the databases, have been subject to documented data processing and quality checks. Furthermore, an automated links to databases, data processing and generation of input files will significantly reduce the resources spent in updating a model with new input data, such as climate and data abstraction.

The hydrogeologic model is essential to groundwater modeling, and the hydrogeologic modeling often comprises a significant part of the total resources spent on modeling. Much may thus be gained by reuse and updating of existing models, when new data and knowledge become available, compared to the nowadays tradition of starting from scratch each time. Although the models themselves often are stored electronically, they are typically not documented in sufficient detail with respect to the assumptions and simplifications made during the interpretation, as well as the storage of digital data used in the construction of the final model.

2 NATIONAL DATABASES

Data relevant for integrated hydrological modeling are stored in several databases hosted by different national research institutions, Figure 1. All databases are stored digitally but vary greatly in their organization and ease and terms by which data can be assessed and extracted. While some data can be downloaded free of charge via the internet by user defined filtering to target the data extraction, other data are subject to a fee, or are only accessible by personal contact to the institutions.

In addition to databases storing raw data, the Geological Survey of Denmark and Greenland (GEUS) has developed a model database for storage of geological and hydrological models. The model database are capable of storing grid data to which various parameters can be associated, e.g. surfaces of geological or model layers and geological and hydrological parameters, etc. Basis data used in model construction, such as digitized points, can similarly be stored in the database as well as geological profiles interpreted during the geological modeling. The overall perspective is to facilitate reuse and continued upgrading of models.

Borehole data • Geology • Groundwater quality • Groundwater levels • Abstraction	Geophysical data	Geological models Hydrostratigraphic models	Land use Farm data	Soil data • Texture • Chemical properties • Hydraulic properties Topography	Surface water quality	Hydrometric data	Agricultural management data from LOOP	Wetland data	Climatic data
Jupiter	GERDA	Model DB	FRJOR	Geodata	ODA	HYMER	AGRI	WET	
(GEUS)	(GEUS)	(GEUS)	(DJF)	(DJF)	(DMU)	(DMU)	(DMU)	(DMU)	(DMI)

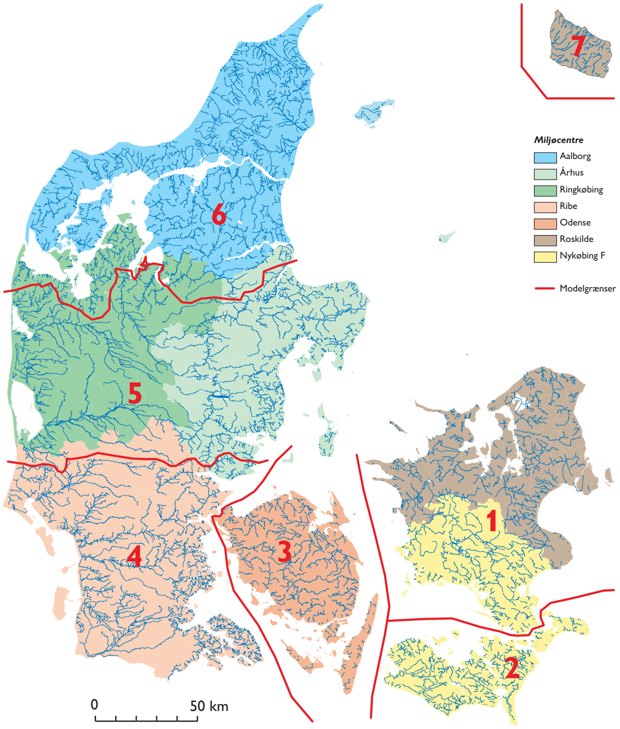
GEUS: Geological Survey of Denmark and Greenland; DJF: Faculty of Agricultural Science
NERI: National Environemtal Research Institute; DMI: Danish Meterological Institute

Figure 1 National databases storing data relevant to hydrological modelling.

3 THE NATIONAL WATER RESOURCES MODEL (DK-MODEL)

Starting in 1996 a national hydrological model (DK-model) has been developed for the entire country (Henriksen et al. 2003; Sonnenborg et al., 2003). The model is composed of a relatively simple description of the unsaturated zone, a comprehensive three-dimensional groundwater component and a river component for stream flow routing and calculation of stream-aquifer interaction. The model is constructed on the basis of the MIKE SHE and MIKE 11 codes.

The DK-model is presently updated. A significant element of the updating is improvement of the hydrogeologic interpretation, based on the recent extensive geologic mapping. The update has been carried out in close collaboration with the regional water authorities, where more than 50 geologic and hydrologic models constructed by the water authorities have been compared with the DK-model. In order to support the details in the local models, the hydrogeologic model in the DK-model is constructed in a 100m × 100m grid, while the numerical simulations are carried out in a 500m×500m grid. Other model updates include a more detailed description of the river network and more spatially detailed input data on e.g. climate and groundwater abstraction.



Figur 2 Sub-models in the DK-model. Colors represent administrative boundaries for regional water authorities

Key issues in the update of the DK-model have been to construct a transparent model and link the model input data to the national databases when possible. Transparency in the model setup was considered important both for the sake of using the model as a communication platform for parties involved in groundwater related issues and to ensure an understandable model setup that can be used as an offset for the development of more detailed models and future model developments based on new information.

Future visions are to sustain an ongoing process in which the national water resources model is regularly updated and adapted as more knowledge becomes available from the continuing geological mapping, the national monitoring program and research activities. For a consistent estimate of the water resource the DK-model will provide the overall hydrogeological reference model for all levels of water resource management, i.e. from a national overview to the definition of actions plans at a local level, where the DK- model can be used as off-set for more detailed model studies.

4 LINKING DATABASES AND THE DK-MODEL

An immense challenge in the model update has been the handling of the vast amount of data existing at a national level including both raw data and existing geological interpretation and model setups. Prior to the project, knowledge and experience in model updating including the reuse of existing models was very limited. The model database has only recently been developed and did not contain models useable for updating the DK-model. The model update process was therefore very laborious and too lengthy to be feasible for a regularly update. A key issue in the project has therefore been to link the model to the national databases hosted by GEUS where possible, and developed procedures and tools to link the databases and the model, and identify possibly problems.

4.1 *Database to model*

The DK-model includes all abstractions, which amounts to withdrawal from more than 40,000 well screens. Data for abstraction are processed on the basis of a standardized query from the national database Jupiter, from which data are converted to GIS-format and processes by standardized GIS-routines. A significant part of the data processing is quality checking including identification of missing data vital for geo-referencing the screen locations and obvious errors in the dataset. More advanced quality checks include checks on reported water abstraction and coupling between well-fields and intakes. Results of the data quality checks are documented in the GIS-coverage and sent to the regional water authorities. Based on their local knowledge, the regional water authorities will be able to supply the missing data, and thus provide a feedback and correction of the national database. A similar procedure is applied to process hydraulic head data also stored in Jupiter used for model calibration and validation.

4.2 *Model to database*

By linking the national databases to the DK-model the model acts as a three-dimensional database, linking the raw data in the national databases to the conceptual understanding of the physical system. This is particularly advantageous in the administration of the freshwater system within Europe. According to the European Water Framework Directive (WFD) (EC, 2000) the qualitative and quantitative status of the groundwater must be reported at the level of groundwater bodies. Groundwater monitoring may be carried out by grouping groundwater bodies with similar characteristics, but it must be documented that the groundwater is monitored in sufficient detail to cover all types of groundwater bodies. In order to do so, it is essential that all groundwater bodies are delineated in three dimensions and that monitoring data can be associated to a specific groundwater body.

4.3 *Model to model*

The DK-model consists of seven submodels, Figure 2. All submodels will be uploaded to the national model database both in form of the conceptual (100m×100m grid cells) and the numerical (500 m×500m grid cells) model. Furthermore, key model results such as simulated hydraulic heads, groundwater recharge and streamflows will possibly be uploaded. With the built-in viewer facilities of the national model database, it is possible for stakeholders lacking modeling expertise to explore the models. All models and model results are freely available. Only restriction in data use is that new models based on the DK-model must be uploaded to the national model databases, and made available for future updates of the national model.

4.4 *Quality assurance by modeling*

The Jupiter database has only recently been expanded to include observations on hydraulic heads collected by the regional water authorities and water companies. Prior to this, hydraulic head data had to be collected from a variety of local databases varying greatly in structure and quality. The procedure has been to collect, pre-process and perform quality checks on data for each new model study, often carried out in very different ways. As a result, quality assurance of data has been a considerable task in all model studies, and in areas with previous model setups data have often been subject to several quality checks. A doubtful and opaque procedure of data collecting, pre-processing and quality checks often makes it difficult to evaluate the data basis and quality, and thereby results in a reduced credibility of the models. Furthermore, it has not been possible to provide feedback to Jupiter on inconsistencies or errors, due to insufficient documentation of the quality check. Storage of all hydraulic data in a central database like Jupiter combined with standardized methods for data processing to link the databases more directly to the models is a major achievement.

Although well structured and subject to quality checks on data upload, the data processing revealed different problems with data stored in the Jupiter database, primarily related to①) lack of information vital for model use (e.g. geo-references), and ② ambiguous information. The main reasons for the imperfect storage of data can be attributed to the continuous development of the database and the conditions for uploading new data. The database was established long ago (Jupiter was established in paper version more than 100 years ago) and have gradually been extended to include more information as this has been required. Data reported to Jupiter at earlier times do thus not contain the information added to the database at a later stage. The Danish legislations prescribe, at an overall level, the data that must be reported and uploaded to the database. It is, however, the data owners who are responsible for the actual upload. To assure a high rate of data uploaded to the database, the amount of mandatory information are balanced between the completeness of the database and the effort put on the data owner for upload, and some information beneficial for model constructions, but not crucial to the water management at local or national level, are thus not mandatory. Another implication is due to the comprehensiveness of the database, allowing a detailed description of e.g. relation between screens, well-fields and administrative plants encompassing several well-fields. While this allows a detailing of the abstraction history, it also opens up for storage of information in different, and thereby ambiguous, ways. As a result, some errors or inconsistency in the data storage in the databases can only be identified and corrected by people with local knowledge. Finally, rigorous data quality checks on upload have only been standard in recent years.

Precipitation and evaporation data are similarly subject to extensive quality assurance on data sampling and use a scientific based state-of-the-art approach to correct the precipitation data (accounting for wind effects and wetting) and estimate evaporation. From a scientific standpoint the data resemble thus the best and most reliable data set. But when data were applied in the national water resources model and model outputs compared to measured groundwater head and river discharge data, significant water balance errors were observed. This error could only be explained by inconsistency in the climatic input data. While the precipitation and evaporation data were sufficiently accurate when only the precipitation or the evaporation where of interest, the methods used were not adequate when the entire water balance was the target.

5 EXPERIENCES AND PERCPECTIVES

The update of the DK-model has provided valuable experience and knowledge in handling large amount of data and linking the model to the national databases. Combining the national model with all relevant data in the national databases has revealed errors and inconsistencies in data. Some inconsistencies are due to the way data are uploaded to the databases by the data owner, in a not fully consistent and detailed way. The current trend in Denmark is that water management decisions are strongly supported and based on hydrological modeling. Possibly unambiguous storing of data may have a server impact on model results, and thereby on the water management. It is therefore envisaged that storage of information relevant for modeling will receive much more focus in the near future.

Recently a new project has been initiated in which all well screens are associated to one of the 385 groundwater bodies in Denmark. The project will further provide suggestions on how to characterize and group the groundwater bodies, and evaluate the adequacy of the present monitoring network. The foundation for the project is the DK-model with its nationwide description of the conceptual understanding of the hydrogeology combined with its link to Jupiter. Once a screen has been associated to a specific groundwater body, it may be stored in the Jupiter database as additional information, allowing future queries at the groundwater body level.

Updating the hydrological part of the DK-model on basis of existing local models has highlighted the importance of thorough documentation and data storage when models are to be re-used in an update procedure. These experiences, combined with the newly developed national guidelines on construction of geologic models (Jørgensen et al. 2008), are being utilized to draw up a comprehensive list of data storage and documentation that should accompany any model. The list will be used to evaluate and possibly expand the present requirements for documentation and data storage in the national model database.

Traditionally, conceptual models are prepared by use of stratigraphical techniques resulting in models comprising geological layers and lenses. This methodology has proven to function well for many groundwater flow problems. However, it fails to account for local heterogeneities that often are important for describing the transport processes leading to an underestimation of the vertical mixing processes (Troldborg et al. 2007). With reactive transport becoming an important modeling objective in future there is a need to characterize and map geological heterogeneity for instance by combining geostatistical tools and geological knowledge. Storing of such information require an extension to the present database designs.

We consider it very likely that the future will bring a much more seamless integration between databases with different raw data relevant for hydrological modeling, databases with model data and models themselves. Future plans are to extend the national databases on raw data to include a description of the heterogeneity, and develop a common platform to all databases for accessing data and heterogeneity descriptions.

REFERENCES

EC Directive 2000/60/EC of the European Parliament and of the Council of October 23 2000 establishing a framework for Community action in the field of water policy. Official Journal of the European Communities, L327/1–L327/72, 22.12.2000.

Henriksen H J, Troldborg L, Nyegaard P, et al.. Methodology for construction, calibration and validation of a national hydrological model for Denmark. Journal of Hydrology, 2003, 280, 52~71.

Højberg A L, Refsgaard J C, van Geer F, Jørgensen L F. et al.. Use of Models to Support the Monitoring Requirements in the Water Framework Directive. Water Resources Management, 2007, 21, 1649~1672. http://dx.doi.org/10.1007/s11269-006-9119-y.

Jørgensen L F, Refsgaard J C, Højberg A L. Joint use of monitoring and modelling. Water Science & Technology, 2007, 8, 21~29.

Sonnenborg T O, Christensen B S B, Nyegaard P,et al.. Transient modeling of regional groundwater flow using parameter estimates from steady-state automatic calibration, Journal of Hydrology, 2003, 273, 188~204.

Troldborg L, Refsgaard J C, Jensen K H, et al.. The importance of alternative conceptual models for simulation of concentrations in multi-aquifer system. Hydrogeology Journal, 2007, 15, 843~860