Model Reduction In Groundwater Hydrology

P.T.M.Vermeulen[a], A.W. Heemink[a] and C.B.M.Testroet[b]

[a] *Delft University of Technology, Faculty of Information Technology and Systems, Department of Applied Mathematical Analysis, Delft, the Netherlands (p.vermeulen@nitg.tno.nl)*
[b] *Netherlands Institute of Applied Geoscience TNO - National Geological Survey, Delft, the Netherlands*

Abstract: Due to the increased computer speed and memory nowadays available, very large numerical groundwater models (with finite difference and finite elements) can be constructed. Those model networks can be time-efficiently parameterised using GIS systems. These advantages, both the computer speed and the accessibility of digital data, change into a disadvantage when there is a need for many model computations. For model calibration and optimisation of hydrologic cases, many impulse-response computations are necessary. Although the computer speed will definitely increase in the future, it appears that the need for model evaluations increases more rapidly than this computer speed. A recent study has shown possibilities to reduce complex computational models to relatively simple models. In essence, a reduced space is captured which contains most of the relevant information of the complex model. The dimension of this reduced space is often small compared to the original dimension and models which compute in a low-dimensional space will require less CPU time. To find the denominators in a large-dimensional space, a particular data-driven pattern identification technique, Empirical Orthogonal Functions (EOFs), is used. These EOFs are the eigenvectors of the system covariance matrix; the accompanying eigenvalue indicates the importance of that particular eigenvector. Those EOFs which contribute the most to the explanation of the complex model behaviour are included in the reduced model; the others are left out. The computational efficiency and accuracy of this reduction method are discussed.

Keywords: Groundwater; Model reduction; MODFLOW; Empirical Orthogonal Functions (EOFs)

1. INTRODUCTION

The increased computer speed has led to numerical groundwater models dimensioned by networks with an enormous amount of grid points. Due to GIS systems those schemes can be parameterised accurately and time-efficiently, justifying the amount of grid points required. To compute this groundwater model, it is necessary to solve a set of equations equal to the number of grid points. More grid points will definitely increase the required computational time, but introducing less grid points, declines the accuracy on a local scale which is undesirable. The question is: "*How can model reduction be applied without the loss of information on a local scale?*" The key thought for this kind of model reduction is to catch patterns which together represent the space in which the model acts (Hooimeijer [2001]). The most important patterns are physical processes, less important patterns (*noise*) can be neglected. Model reduction without the loss of detail is implemented as a weighted sum of those distinguished important patterns, see equation (1).

$$\mathbf{h}_k = \sum_{i=1}^{n} \mathbf{r}_k \mathbf{P}^i \qquad (1)$$

where $\mathbf{h}_k$ (L) is the result for time step k, $\mathbf{r}_k$ these in reduced space (L) and $\mathbf{P}^i$ is the i^{th} selected important pattern (dimensionless) out of n. The entire reduction process is summarized in Figure 1.

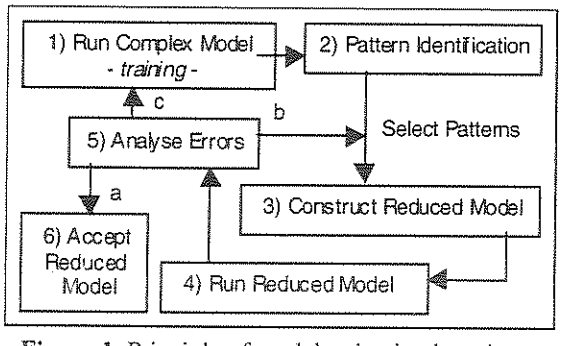

Figure 1. Principle of model reduction based on data mining from a complex model.

From Figure 1 it appears that model reduction can be

an iterative process. First, the complex model should be *trained* in such a manner that all future dynamics are captured by patterns. Secondly, the results from the reduced model are analysed and the error can lead to a) accepting the reduced model, b) adding more patterns or c) training the complex model for a different or longer period.

This paper describes the method proposed in Figure 1. A test model will illustrate the mathematics underlying the method, Figure 2.

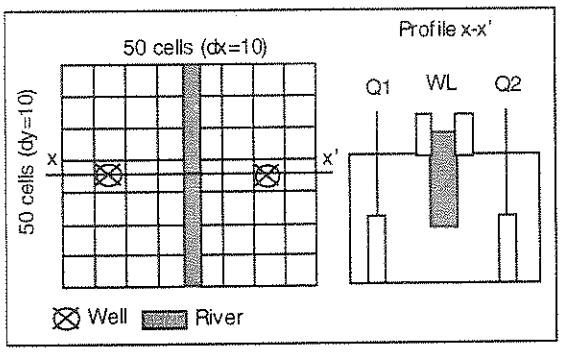

Figure 2. Layout of the test model.

The test model (2500 cells) has a Dirichlet boundary condition on the left and right model edges. There are 2 independent wells with randomly an infiltration and/or exfiltration rate of maximum 500 m^3 day^{-1}. There is a river heading from north to south with a water level fluctuation of 2 meter. There is recharge randomly between −10 and 10 mm day^{-1}. The model is simulated for 60 time steps, each 1 day long.

2. PATTERN IDENTIFICATION

2.1 What is a Pattern ?

A pattern is a combination of distinguished physical processes which denominates the model. Therefore a pattern is comparable with analytical elements used and proposed by Strack [1989]. Instead of representing only one feature (well, river e.g.), a patterns can represent combined features within distorted parameterisations.
There are two basic types of pattern identifications techniques, data driven and analytical driven techniques. This paper only describes the data driven technique and is usual called Empirical Orthogonal Functions (Richman et al. [1986]).

2.2 Empirical Orthogonal Functions (EOFs)

To identify the patterns (called EOFs) in the results of the groundwater model we created a data set. The groundwater model has a number of grid points N_G and time steps N_T. The results are arranged in a matrix $\mathbf{X}$ with vectors (columns in $\mathbf{X}$), $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_N$ ($N=N_T$). Each $\mathbf{x}_k$ contains the results at each grid location at a given time step k. Locations that act time-independently are excluded from $\mathbf{X}$. The pattern value for these locations is zero.

To quantify the dynamics in the model one should zero average each $\mathbf{x}_k$ by subtracting the mean vector:

$$\mathbf{d}_k = \mathbf{x}_k - \frac{1}{N_T}\sum_{i=1}^{N_T}\mathbf{x}_{k,i} \qquad (2)$$

The main difference between $\mathbf{D}$ and $\mathbf{X}$ is that all grid points within $\mathbf{d}_k$ are comparable to each other in terms of differences (variances) instead of absolute magnitudes. To express the correlation of each grid point to the other ones, we computed the covariance as the major product of $\mathbf{D}$:

$$\mathbf{C}_{[N_G x N_G]} = \frac{1}{N_T - 1}\mathbf{D}_{[N_G x N_T]}\mathbf{D}^{\mathbf{T}}_{[N_T x N_G]} \qquad (3)$$

where $\mathbf{C}[N_G x N_G]$ is the covariance matrix with dimension $N_G x N_G$.

The eigenvalue decomposition factorises the symmetric matrix $\mathbf{C}$ into real normalized eigenvectors $\mathbf{v}$ ($\mathbf{v}^T\mathbf{v}=1.0$) and eigenvalues λ ($\lambda >= 0.0$), for $k=1,N_G$:

$$\mathbf{C}\mathbf{v}_k = \mathbf{v}_k\lambda_k \qquad (4)$$

where $\mathbf{v}_k$ is the k^{th} eigenvector go with the k^{th} eigenvalue λ_k. The eigenvalues are scaled to represent the explained variance, for $k=1,N_G$:

$$\lambda_k = \frac{\lambda_k * 100}{\mathrm{trace}(\mathbf{C})} \qquad (5)$$

where **trace**($\mathbf{C}$) is the total variance formulated as the summation of the main diagonal in $\mathbf{C}$.

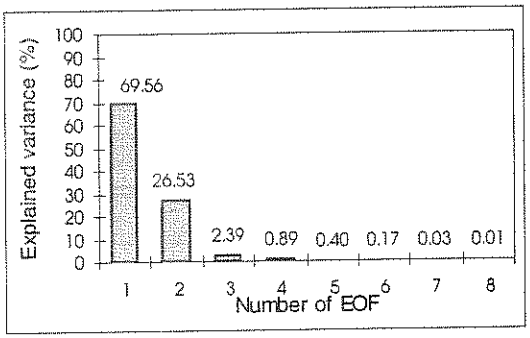

Figure 3. Explained variance by the important EOFs.

In Figure 3 the largest EOFs are presented for the test model. Apparently the complex model can be described for 98% by 3 patterns and for 99.9% by 6 patterns only.

2.3 Efficient Computation of EOFs

In practice $N_T << N_G$ and therefore the rank of the covariance matrix **C** is equal to N_T. This results for $N > N_T$ in eigenvalues λ_N equal 0.0 and therefore eigenvectors $\mathbf{v}_N$ equal 0.0. Instead of calculating the eigenvalues for the major product of **D** (equation 3), which is a time-consuming process, we compute the eigenvalues for the minor product of **D**:

$$\mathbf{D}^T_{[N_T x N_G]} \mathbf{D}_{[N_G x N_T]} = (\mathbf{U} \Lambda \mathbf{U}^T)_{[N_T x N_T]} \qquad (6)$$

where **U** and **Λ** are matrices with the eigenvectors and eigenvalues, respectively, but with a small dimension N_T instead of a large dimension N_G. To expand the eigenvectors to dimension N_G we applied the following formulation (Golub et al. [1989]):

$$\mathbf{V}_{[N_G x N_T]} = \mathbf{D}_{[N_G x N_T]} \mathbf{U}_{[N_T x N_T]} \Lambda^{-1/2}_{[N_T x N_T]} \qquad (7)$$

where it is easy to compute $\Lambda^{-1/2}$ (square root of the inverse of **Λ**) because **Λ** is a main diagonal matrix with only positive eigenvalues.

2.4 How Many Patterns ?

Theory prescribes that the rank of the covariance matrix **C** will provide the minimum number of patterns necessary to describe the data completely. It is the challenge to increase the computational speed by reducing the number of patterns such that an acceptable reconstruction can be applied. In Figure 4 the number of EOFs is plotted against the reconstruction error between the reduced model and the complex model.

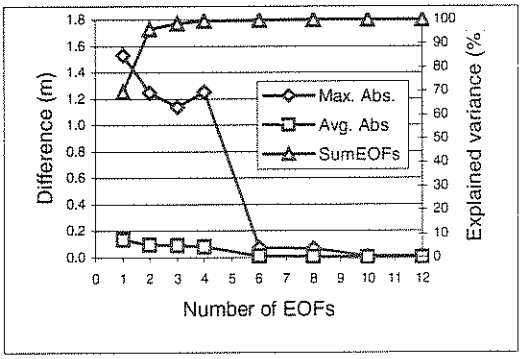

Figure 4. Calculated differences (m) in groundwater levels for a different number of EOFs.

In Figure 4 the maximum absolute error (Max.Abs) shows a significant improvement between 4 (99.37%) and 6 (99.94%) EOFs. The average absolute error (Avg.Abs) shows the same behaviour but at a magnitude smaller. Adding more EOFs (>6) will increase the computational time more than it will improve the model results.

In the literature different kinds of strategies are mentioned to obtain the amount of EOFs (i.e. Reyment et al. [1993] and Hooimeijer [2001]). One philosophy for gathering EOFs is to collect those patterns which describe a physical process and to leave out all others. In Figure 5 the patterns of the main 6 EOFs from Figure 3 are presented. All patterns are clearly related to a physical process (wells and river) within the model.

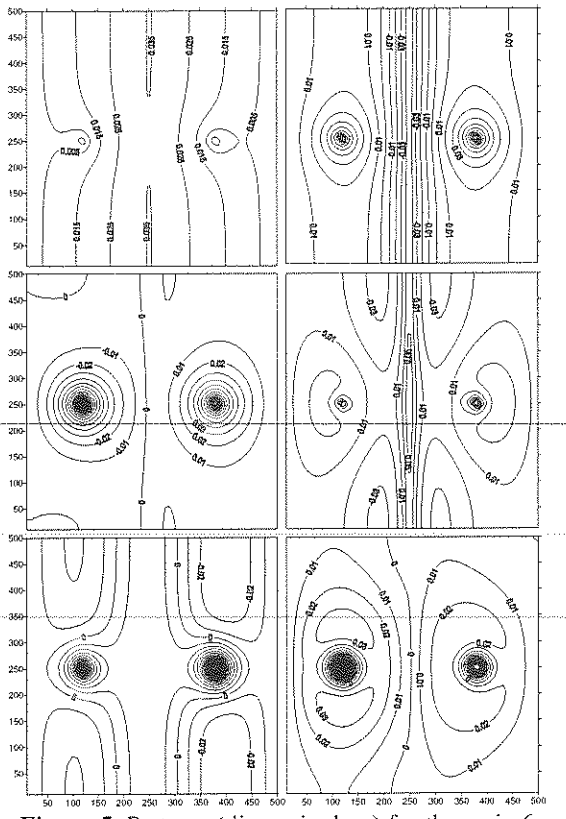

Figure 5. Patterns (dimensionless) for the main 6 EOFs explaining 99.94% variance in total. The upper left box is EOF no.1, the lower right box is EOF no. 6.

Statistically speaking, patterns will become more related to physical processes if they are based on large data sets. On the other hand patterns can be obtained time-efficiently if the model impulse for the model (*training*) is chosen wisely. Finally, patterns can be obtained by a covariance matrix which is based only on results from certain areas of interest within the model. In this manner patterns will become more representative for the interested area.

3. CREATING A REDUCED MODEL FOR GROUNDWATER FLOW

3.1 Numerical Modelling of Groundwater Flow

To describe groundwater flow in a two-dimensional plane we use the groundwater-flow equation based on Darcy's law and the equation of continuity [Strack, 1989]:

$$\frac{\partial}{\partial x}(T_x \frac{\partial h}{\partial x}) + \frac{\partial}{\partial y}(T_y \frac{\partial h}{\partial y}) = S \frac{\partial h}{\partial t} - q \qquad (8)$$

where h is the groundwater head (L), x,y are the coordinates in space (L), t is the time (T), $T_{x,y}$ is the transmissivity in x and y direction (L^2T^{-1}), S is the storage coefficient (-) and q is the discharge or recharge term (L^3T^{-1}). In models such as MODFLOW [McDonald and Harbaugh, 1988] equation (8) is discretized for a mesh of grid cells into an equation with finite differences:

$$T_{i-1,j}h_{i-1,j} + T_{i,j-1}h_{i,j-1} + T_{i+1,j}h_{i+1,j} + T_{i,j+1}h_{i,j+1} + (-T_{i-1,j} - T_{i,j-1} - T_{i+1,j} - T_{i,j+1} - \frac{S}{\Delta t})h^k_{i,j} = -S\frac{h^{k-1}_{i,j}}{\Delta t} - q \qquad (9)$$

where $T_{i-1,j}$ is the transmissivity over the cell face between the grid cells with indices i-1,j and i,j (i=row;j=column number). For grid cell i,j the groundwater head of the previous time step is expressed as h^{k-1}. The equation can be written in matrix notation to be solved implicitly:

$$\mathbf{A}\mathbf{h}_k = \mathbf{S}\mathbf{h}_{k-1} + \mathbf{q} \qquad (10)$$

where

$$\mathbf{A} = \begin{bmatrix} -T_{1,2} - T_{2,1} - \frac{S_{1,1}}{\Delta t} & T_{1,2} & T_{2,1} & 0 \\ T_{1,1} & -T_{1,1} - T_{2,2} - \frac{S_{1,2}}{\Delta t} & 0 & T_{2,2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & T_{1,2} & T_{2,1} & -T_{1,2} - T_{2,1} - \frac{S_{2,2}}{\Delta t} \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} -\frac{S_1}{\Delta t} & 0 & 0 & 0 \\ 0 & -\frac{S_2}{\Delta t} & 0 & 0 \\ 0 & 0 & -\frac{S_3}{\Delta t} & 0 \\ 0 & 0 & 0 & -\frac{S_4}{\Delta t} \end{bmatrix}, \mathbf{h}_{k-1} = \begin{bmatrix} h_1^{k-1} \\ h_2^{k-1} \\ h_3^{k-1} \\ h_4^{k-1} \end{bmatrix}, \mathbf{q} = \begin{bmatrix} -q_1 \\ -q_2 \\ -q_3 \\ -q_4 \end{bmatrix}$$

The symmetric system matrix **A** is a very sparse matrix consisting, for this two-dimensional case, out of five diagonals with non-zero elements. Rearranging **A** in equation (10) to the other side of the equation will result in a formulation for the groundwater flow to be solved explicitly:

$$\mathbf{h}_k = \mathbf{A}^{-1}\mathbf{S}\mathbf{h}_{k-1} + \mathbf{A}^{-1}\mathbf{q} \qquad (11)$$

To solve equation (11) we should compute $\mathbf{A}^{-1}$ (the inverse of **A**). This is a very time-consuming process and should be avoided. Due to the incorporation of EOFs this problem can be overcome in reduced space.

3.2 Reduced Model

The selected EOFs (§2.4) are stored as vectors **p** in matrix **P** (called the *reductor*). The first vector $\mathbf{p}_1$ in **P** is the one with the largest eigenvalue λ, the second vector $\mathbf{p}_2$ is the one with the second-largest eigenvalue and so on. The reductor **P** is orthonormal; all vectors have a length of 1.0 and are perpendicular to each other ($\mathbf{P}^T\mathbf{P}=\mathbf{I}$). **P** has as many columns as selected EOFs and the number of rows is equal to the amount of *active* grid points in the model; so $\mathbf{P}[N_P \times N_G]$.

In the operator approach, **P** can operate as reductor ($\mathbf{P}^T$) and reconstructor (**P**). The reductor reduces a high-dimensional vector $[N_G]$ by vector multiplication to a low-dimensional space $[N_P]$; the reconstructor performs the reverse operation. To reduce a matrix we should first multiply it with the reductor and thereafter with the reconstructor. Equation (11) can be projected into reduced space by applying the reductor $\mathbf{P}^T$ and reconstructor **P**:

$$\mathbf{r}_k = \mathbf{P}^T\mathbf{A}^{-1}\mathbf{S}\mathbf{P}\mathbf{P}^T\mathbf{h}_{k-1} + \mathbf{P}^T\mathbf{A}^{-1}\mathbf{q} \qquad (12)$$

where $\mathbf{r}_k$ $[N_P]$ is the groundwater head (L) for time step k in reduced space. To reconstruct the groundwater head to the original space $[N_G]$, apply:

$$\mathbf{h}_k = \mathbf{P}\mathbf{r}_k \qquad (13)$$

3.3 Solving the Reduced Model

It is not recommended, and even unnecessary to solve the reduced model completely for each time step k. Though most of the terms in equation (12) are time-independent (matrices $\mathbf{P}^T$,$\mathbf{A}^{-1}$,**S** and **P**), the reduced model can be simplified by carrying out some matrix multiplications beforehand. The main issue here is avoiding the computation of $\mathbf{A}^{-1}$.

Due to the symmetric form of **A**, equation (12) can be rewritten as:

$$\mathbf{r}_k = (\mathbf{A}^{-1}\mathbf{P})^T \mathbf{S}\mathbf{P}\mathbf{P}^T\mathbf{h}_{k-1} + (\mathbf{A}^{-1}\mathbf{P})^T\mathbf{q} \quad (14)$$

Instead of computing $\mathbf{A}^{-1}$ completely, we introduce a matrix **M** as:

$$\mathbf{A}^{-1}\mathbf{P} = \mathbf{M} \Leftrightarrow \mathbf{A}\mathbf{M} = \mathbf{P} \quad (15)$$

Each vector **p** in **P** is solved sequentially to find the vectors **m** in **M**; for i=1,N_P:

$$\mathbf{A}\mathbf{m}_i = \mathbf{p}_i \quad (16)$$

The reduced model (12) can be minimized to

$$\mathbf{r}_k = \mathbf{N}\mathbf{r}_{k-1} + \mathbf{M}^T\mathbf{q} \quad (17)$$

where

$$\mathbf{N} = \mathbf{M}^T\mathbf{S}\mathbf{P} \quad (18)$$

The reduced model consists finally of three time independent matrices: $\mathbf{P}[N_G x N_P]$, $\mathbf{M}[N_P x N_G]$ and $\mathbf{N}[N_P x N_P]$ and two time-dependent vectors $\mathbf{r}_{k-1}[N_P]$ and $\mathbf{q}[N_G]$. To compute the groundwater heads in reduced space one needs a total of $(N_P*N_P+N_G*N_P)+N_P$ FLOPS (floating-point operations) for each time step k. For the test model with 2500 equations this results in 15042 FLOPS. In Figure 6 the results are presented as differences between the complex model and the reduced model with 6 EOFs (Figure 5), explaining 99.94% variance.

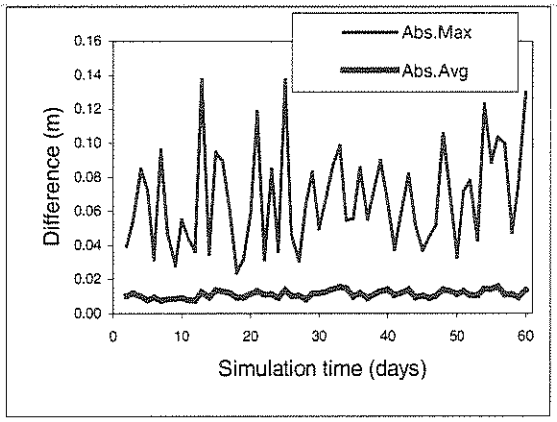

Figure 6. Calculated differences in groundwater levels (m) as the absolute maximum and the absolute average.

The absolute maximum error (Abs.Max) and the absolute average (Abs.Avg) both show a trend explainable by the fluctuations of the model forces only (wells, recharge and river). The error does not have the tendency to grow over time due to missing patterns or wrongly defined patterns. The reduced model is stable and can therefore be used for longer simulations different than the one used to *train* the complex model. In Figure 7 a plane view is given of the absolute maximum differences after 60 time steps.

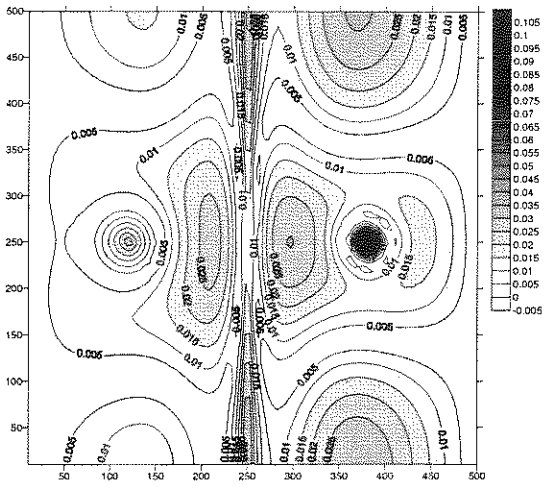

Figure 7. Plane view of absolute maximum differences after 60 time steps of simulation.

The unacceptable major error (0.11m) only occurs locally in the middle of the right well. Except this error, the errors in the remaining area are acceptably small (<0.025m) and can be ascribed to those zones in which the groundwater fluctuation is maximum, caused by wells, the river and recharge.

3.4 Initial conditions

Groundwater models often have initial conditions different from zero. The reduced model is based on the superposition of patterns, and for the first time step, no patterns haven been applied yet. Therefore the initial conditions for the reduced model are always equal to zero, despite the original initial conditions. To include those initial values correctly we should add the results from equation (13) at the end of the simulation to the initial values $h_{initial}$, so:

$$\mathbf{h}_k = \mathbf{P}\mathbf{r}_k + \mathbf{h}_{initial} \quad (19)$$

For features like surface water and other head-resistance relationships within a groundwater model, absolute levels herein (water levels, drainage base) should be corrected by subtracting the initial condition at that specific location before adding it to the force vector **q** within equation (12) or (17).

3.5 Verification and Suitability of the Reduced Model

One of the main reasons to create a reduced model is to speed up computational time by applying the reduced model to other cases. These models should not possess more dynamic or different locations with forcing-terms (wells, rivers e.g.), another time discretisation and/or a different system matrix. What may differ in such other models, are the forces on the same locations, for example the extraction rates or the water levels. In Figure 8 results are given for a model with different forces than the one used for the model *training*. The distribution of the errors are comparable with those in Figure 7.

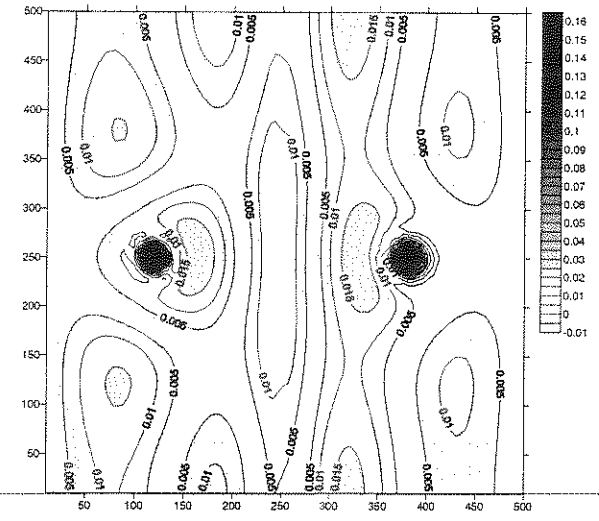

Figure 8. Maximum difference in groundwater levels by verification of the reduced model with different model forces.

The time efforts derived with the reduced model are shown in Figure 9. In these linear cases the increased CPU time is linear for all processes mentioned in Figure 9. The reduced model does gain more time effort for larger model schemes.

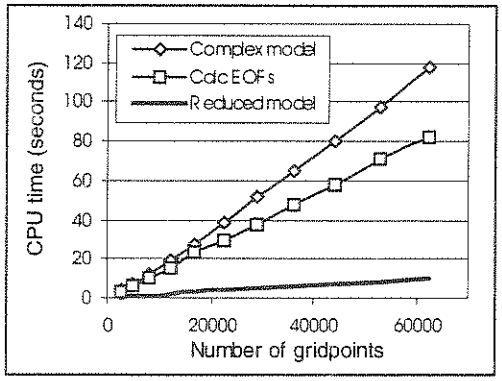

Figure 9. CPU times for the computation of the complex model, the EOFs and the reduced model.

4. CONCLUSIONS

For many years the evolution of computers was much faster than that of the accessibility of digital data. Nowadays it seems the other way around, and numerical models grow larger and larger, swallowing computational time. This paper describes a different strategy in solving those model by searching for patterns. Those patterns (Empirical Orthogonal Functions), are the most important eigenvectors of the covariance matrix, which is computed from one transient model simulation. EOFs can be compared with the analytical elements used and proposed by Strack [1989]. Instead of representing one feature (analytical element), EOFs can represent combined features within distorted parameterisations.

The described EOF method reduces the complex model to a smaller space which still can predict the behaviour of the complex model. Instead of solving millions of equations iteratively, the reduced model computes the groundwater heads by only three matrix multiplications. The computational effort is thus reduced enormously.

Besides gaining time, the discussed model also makes it possible to select an area of interest from the reduced (regional) model. Without reconstructing the model and concerning about model edges, more time can be saved because only a limited part of the reduced model has to be multiplied. This practical phenomenon can be implemented within Decision Support Systems (DSS).

5. REFERENCES

Golub, G, and C. van Loan, *Matrix Computations*, John Hopkins University Press, 2nd edition, 1989.

Hooimeijer, M.A., *Reduction of Complex Computational Models*, PhD thesis, Delft University of Technology, in press, 2001.

McDonald, M.G., and A.W. Harbaugh, *A Modular Three-Dimensional Finite-Difference Groundwater Flow Model*, U.S. Geological Survey, Open-File Report 83-875, Book 6, Chapter A1, 1988.

Reyment, R.A., and K.G. Jöreskog, *Applied Factor Analysis in the Natural Sciences*, Cambridge University Press, 1993.

Richman, M.B., Review Article, Rotation of Principal Components, *Journal of Climatology*, 6, 293-335, 1986.

Strack, O.D.L, *Groundwater Mechanics*, Prentice Hall, 1989.