

ЗАПИСНИЦИ СРПСКОГ ГЕОЛОШКОГ ДРУШТВА за 2013. годину

COMPTE RENDUS DES SÉANCES DE LA SOCIÉTÉ SERBE DE GÉOLOGIE
pour les années 2013

REPORTS OF THE SERBIAN GEOLOGICAL SOCIETY
for the year 2013

Beograd, 2014.

АУТОМАТСКА КАЛИБРАЦИЈА ХИДРОДИНАМИЧКИХ МОДЕЛА ПОМОЋУ ПРОГРАМА PEST

(Збор Секције за хидрогеологију СГД, 29.11.2013.)

Д. Поломчић, Д. Бајић, М. Чокорило Илић

Департаман за хидрогеологију, Универзитет у Београду-Рударско-Геолошки факултет. Ђушина 7, Београд

Department of Hydrogeology, University of Belgrade-Faculty of Mining & Geology, Djusina 7, Belgrade, Serbia

Апстракт: У хидрогеологији, код хидродинамичког моделирања режима издани подземних вода, калибрација модела представља усклађивање резултата симулације са подацима који су регистровани на терену. У пракси, обично се то радило „ручно“ (мануелно) – „пробањем“ са различитим вредностима хидрогеолошких параметара и хидрауличких карактеристика граничних услова. Употребом програма PEST уводи се аутоматска калибрација модела којом се знатно смањује субјективни утицај аутора модела на резултат. Новим приступом у PEST-у, односно увођењем тзв. контролних тачака, превазилази се концепт хомогених зона са вредностима параметара хидрогеолошке средине или зона са задатим граничним условима. Међутим, последица оваквог начина аутоматске калибрације је знатно време потребно за обављање прорачуна. PEST садржи два модула за скраћење времена тог процеса - Parallel PEST и BeoPEST. Други начин превазилажења дужине трајања прорачуна је употреба SVD анализе са обавезном оценом неодређености модела уз модул NSMC.

Кључне речи: регуларизација са контролним тачкама, Parallel PEST, BeoPEST, Singular Value Decomposition, Null Space Monte Carlo

Увод

Процес еталонирања (калибрације) модела представља најделикатнију фазу израде модела. Он подразумева усклађивања резултата симулације (нпр. режима издани) са подацима регистрованим у природи (стања пијезометарског нивоа и елементе биланса подземних вода). То је осетљив и напоран процес који захтева доста времена за успешну реализацију.

Калибрација модела може се обављати мануелно (ручно) и аутоматски преко наменских програма. У домаћој пракси знатно је заступљенији процес мануелне калибрације који се заснива на низу пробања са различитим комбинацијама вредности параметара модела. Овај начин захтева велико искуство аутора модела, при чему остаје отворено питање да ли је добијено решење са оптималним скупом параметара.

У циљу избегавања учешћа субјективног суда истраживача као и услед великог броја могућих комбинација параметара модела јавила се потреба за ефикаснијим приступом процесу еталонирња модела. Аутоматска калибрација обезбеђује наведене захтеве. Њоме се оптимизирају вредности изабраних параметара хидрогеолошког система, при чему је подразумевани критеријум, односно циљна функција, минимизирање разлика између регистрованих вредности и резултата симулације на моделу.

Индустријски стандард за аутоматску калибрацију је програм PEST (Model-Independent Parameter Estimation and Uncertainty Analysis). У питању је најнапреднији светски програмски пакет за оптимизацију параметара било којег симулационог модела и анализу неизвесности прогнозних прорачуна. PEST се успешно користи у многим научним областима: хидрогеологија, хидрологија, хидраулика, геофизика, геомеханика, хемија, аеронаутика, машинско инжењерство, биологија, итд. (Doherty, 2013).

На основу задатих критеријума, PEST мења параметре модела све док се излазни резултати симулације приближно не поклопе са мереним подацима (нивои, протицаји) (Doherty, 2010). Оптимизациона метода се заснива на Gauss-Markar-Levenbergovom алгоритму проналажења минимума суму квадрата резидуала (разлика између мерених података и резултата симулације модела). Када су у питању хидродинамички модели, оптимизација параметара је итеративни процес.

Бројни параметри могу да се задају, односно да се одређују њихове оптималне нумеричке вредности у склопу постављених критеријума: за неограничени број зона хидрогеолошких параметара (коефицијената филтрације - хоризонталне и вертикалне вредности), специфичну издашност издани, специфично ускладиштење, и за неограничени број врста, карактеристика и деоница граничних услова (интензитет инфилтрације и евапотранспирације, дубину евапотранспирације, протицаје или пијезометарске нивое, кондуктивитете граничних „хед-депендент“ услова итд.).

Иако има могућност симултаног одређивања већег броја параметара, резултат који се добије употребом класичног PEST-а налази се у оквирима које је аутор модела поставио. То се односи на димензије, број и облик зона неког параметра које се не мењају током прорачуна. Вредности параметра (нпр. коефицијената филтрације) унутар унапред дефинисаних зона су константне вредности, па су и та унапред дефинисана поља са хомогеним вредностима задатих параметара, што у природи није случај (слика 1).



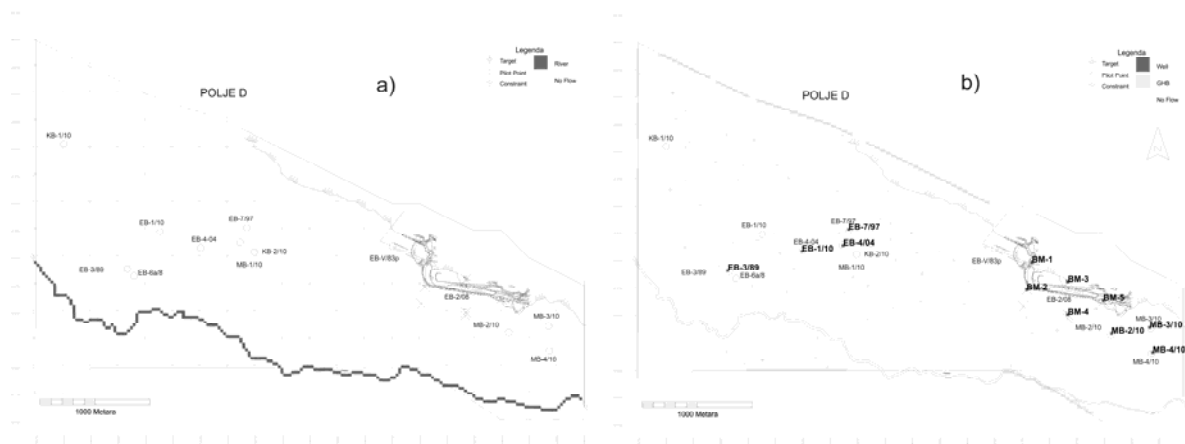
Сл. 1. Приказ хомогених зона параметра модела који се одређује (нпр. коефицијент филтрације)

Fig. 1. The display of homogeneous zones of model parameter which has been determined (e.g. hydraulic conductivity)

Пест са регуларизацијом

Један од начина да се наведени недостатак класичног PEST-а превазиђе направљен је уношењем потпуно новог концепта који подразумева геостатистичке обраде података и примену Кригинг методе у интерполацији између контролних тачака (Тонкин анд Дохерту, 2005). Те тачке се називају Пилот поинтс, и не означавају нужно фиктивне тачке, мада су знатно ређе у питању тачке у којима је нека величина, односно параметри који се одређују, познати. Код ове врсте аутоматске калибрације, свака контролна тачка представља параметар који треба да се одреди (Doherty, 2003). Концепт мерних тачака се не ограничава само на вредности коефицијента филтрације већ на све врсте параметара који се одређују током еталонирања модела.

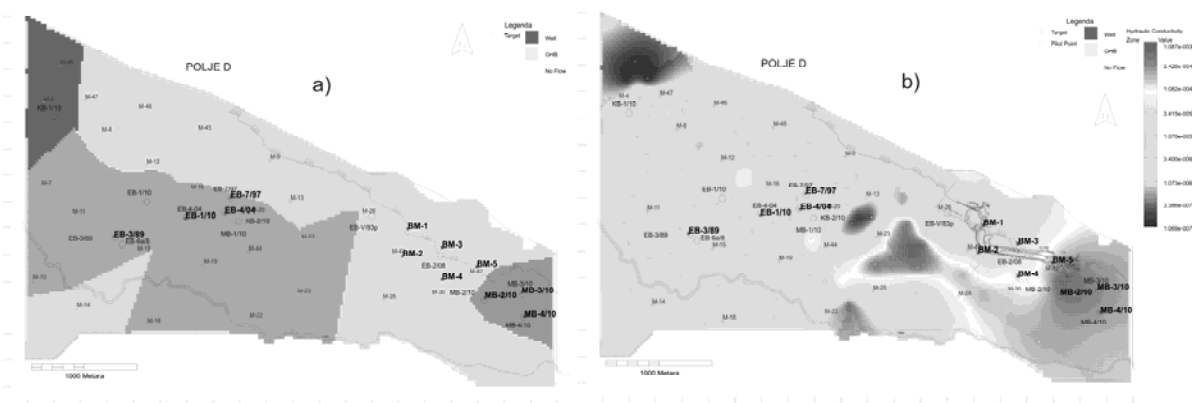
Постоји више начина задавања контролних тачака на моделу: преко регуларне мреже (слика 2-а) и преко триангулације (слика 2-б).



Сл. 2. а) Пилот поинтс – регуларна мрежа б) Пилот поинтс – мрежа добијена триангулацијом

Fig. 2. a) Pilot points - regular network b) Pilot points – network resulting from triangulation

Сам процес примене контролних тачака у PEST-у назива се регуларизација односно регулисање дистрибуције одређених параметара и/или вредности граничних услова. Резултат оваквог приступа у PEST калибрацији даје резултат који укључује хетерогеност средине (струјне области) по питању дистрибуције параметара средине који су одређивани Doherty et al, 2010). На слици 3 испод су приказани традиционални резултат калибрације модела (мануелна или PEST) и PEST са регуларизацијом.



Сл. 3. Зоне хомогених вредности параметара модела – а) класичан PEST б) хетерогеност средине као резултат примене PEST-а са регуларизацијом

Fig. 3. Zones of homogeneous values of model parameters - a) classic PEST b) heterogeneity of the porous media as a result of the implementation of PEST with regularization

Време трајања аутоматске калибрације са регуларизацијом или са контролним тачкама

Један од највећих проблема са употребом Pilot pointa у PEST-у је потребно време за прорачуне. Повећање броја контролних тачака које се уносе у модел условљава продужење временског периода који је потребан за еталонирање модела. Сваку контролну тачку PEST посматра као параметар чија се оптимална вредност одређује. Сваки параметар у PEST-у захтева једну или две симулације по итерацији. Ако нпр. у моделу има 1000 пилот тачака, онда је потребно 1000 симулација (или 2000, ако се примењује централно диференцирање) само по једној итерацији. Већи број итерација је често потребан за сваки пролаз прорачуна током калибрације модела. Ако сваки пролаз прорачуна траје неколико минута, онда је потребно много времена да се цео процес заврши. Време трајања калибрације на овај начин мери се часовима, некада и данима.

PEST омогућава два начина убрзања прорачуна: Parallel PEST и BeoPEST. Оба начина се заснивају на симултаном калибрацији повезивањем више рачунара са основном идејом повећања процесорске снаге која се користи у прорачунима. Основни рачунар назива се „Master“, а сви други „Slave“. На сваком „Slave“ рачунару мора постојати директоријум у коме су смештене идентичне полазне датотеке за прорачуне као што су на „Master“ рачунару. Подржано је до 100 директоријума, односно „Slave“ рачунара. Последица оваквог приступа прорачунима је редукција времена потребног за еталонирање модела.

Parallel PEST омогућава повезивање више рачунара кроз мрежу, док је BeoPEST специјална верзија PEST-а која обезбеђује да се прорачуни обављају симултано користећи интернет и удаљене повезане рачунаре (тзв. Cloud Computing) (Hunt et al, 2010). „Slave“ рачунари са којима се преко интернета повезује „Master“ рачунар могу бити на познатим локацијама или се могу изнајмити рачунари са цену од 30\$, па навише по дану коришћења.

Други начин употребе BeoPEST-а је везан само за „Master“ рачунар, при чему се додавањем „Slave“ директоријума на њему повећава ангажованост процесора.

Друга могућност скраћења времена аутоматске калибрације са применом PEST-а са регуларизацијом је употреба тзв. модула SVD.

PEST + SVD (Singular Value Decomposition) анализа

Појединачно разлагање параметара модела (Singular Value Decomposition - SVD) је процес кроз који је одређивање параметара знатно олакшано. Применом SVD реформулише се калибрација модела тако што се обавља процена сета „супер-параметара“, који су линеарна комбинација „правих“ параметара (Tonkin and Doherty 2005). Скуп „супер-параметара“ је обично много мањи, нпр. 10-20% од укупног броја параметара, што за последицу има знатно краће време потребно за прорачуне. На пример, ако је било 500 пилот тачака и за сличај дефинисања броја супер-параметара на 15% укупног броја параметара, онда је број супер-параметара 75. Практично то значи да ће иместо 500 прорачуна по једној итерацији, из примену SVD бити извођено 75 прорачуна по итерацији. Увек се поставља питање колики број „супер-параметара“ треба изабрати. Нови утилиту SUPCALC у оквиру PEST-а се користи за превазилажење овог проблема. Њиме се одређује минимално потребни и

максимално допуштени број „супер-параметара“ за калибрацију модела. На аутору модела остаје да изабере једну од вредности из допуштеног дијапазона вредности.

У односу на класичан PEST, SVD има још једну предност која се односи на знатно квалитетнију обраду одређених типова параметара модела који имају различите осетљивости на промене њихових вредности током аутоматске калибрације. Међутим, један од недостатака SVD-а је и знатно сложенија процедура подешавања и извршавања овог модула.

Међутим, и SVD може постати спор ако је потребно оптимизовати преко 2000 параметара. Тада је упутније користити алгоритам LSQR (Sparse Equations and Least Squares).

Null Space Monte Carlo (NSMC)

Резултати добијени применом SVD нису једнозначни и нужно захтевају оцену кроз анализу неодређености модела. Један од најбољих начина је примена Null Space Monte Carlo (NSMC) технике која је направљена за рад на калибрисаним моделима применом SVD-а са много непознатих параметара (Doherty et al, 2011; Tonkin and Doherty, 2009).

NSMC техника омогућује аутору модела генерисање више калибрисаних модела с различитим скуповима (сетовима) међусобно усклађених параметара. Добијени резултати, односно модели ће се међусобно разликовати. Дакле, основно решење NSMC није егзактно решење, већ дијапазон могућих решења (Herckenrath et al, 2011). За случај да је постављен критеријум од 10 генерисаних модела и да се анализира само вредност коефицијента филтрације, као резултат ће се добити 10 модела са различитим матрицама коефицијента филтрације. Такви модели (10 њих) могу се користити за прогнозне прорачуне.

Овај концепт има оправдања ако се врати на почетак и дефиницију модела да је то приближно представљање реалног (хидрогеолошког) система. С друге стране, ово је доминантно амерички начин давања решења неког задатка где се покушава избећи одговорност, односно процес одлучивања пребацивати на другог (друге службе).

Закључак

PEST је модуларан програм за оптимизацију параметара током процеса аутоматске калибрације модела, односно решавања инверзних проблема. Програм има широку примену која превазилази моделирање подземних вода. Развојем програма PEST уведен је концепт Пилот или контролних тачака којим се превазилази концепт хомогених зона са вредностима параметара средине или зона са задатим граничним условима. Последица оваквог приступа је знатно време потребно за извођење прорачуна. Временом су у оквиру програма PEST уграђена два модула којим се може скратити време прорачуна (Parallel PEST и ВеоPEST) преко међусобног повезивања више рачунара (мрежним кабловима и преко интернета) односно, искоришћавањем већег процента ангажоване процесорске снаге рачунара. Други начин скраћења времена је примена модула SVD у PEST-у којим се уводе „супер-параметри“ који су у линеарној зависности са оригиналним параметрима модела. За калибрисане модела уз помоћ модула SVD мора се извршити анализа неодређености. У оквиру PEST-а развијен је нов модел назван NSMC (Null Space Monte Carlo) којим се добија дијапазон могућих решења.

Захвалница

Аутори се захваљују Министарству просвете, науке и технолошког развоја Републике Србије за финансирање пројекта 176022 и ИИИ-43004.

AUTOMATIC CALIBRATION OF HYDRODYNAMIC MODELS USING PEST PROGRAM

Abstract: In hydrogeology, in hydrodynamic modeling of groundwater aquifer's regime, model calibration represents the alignment of the simulation results with the data registered in the field. In practice, usually it was done "by hand" (manual) – by "testing" with different values of hydrogeological parameters and hydraulic characteristics of the boundary conditions. By using the PEST program, automatic calibration of models has been introduced, and it has proved to significantly reduce the subjective influence of the model creator on results. With the new approach of PEST, i.e. with the introduction of so-called control points, the concept of homogeneous zones with parameter values of porous media or zones with the given boundary conditions has been outdated. However, the consequence of this kind of automatic calibration is that a significant amount of time is required to perform the calculation. PEST contains two modules for the shortening of that process - Parallel PEST and BeoPEST. Another way to overcome the duration of the calculation is the use of SVD analysis with the mandatory assessment of model uncertainty with module NSMC.

Key words: pilot points with regularization, Parallel Pest, Beopest, Singular value decomposition, Null Space Monte Carlo

Introduction

The process of model calibration is the most delicate phase of the model creating. It involves alignment the simulation results (e.g. aquifer regime) with the data registered in nature (state of piezometric level and elements of the groundwater balance). It is a delicate and difficult process that requires a significant amount of time for successful realization.

Model calibration can be performed manually and automatically, through programs used for this purpose. In domestic practice much more common is a process of manual calibration, which is based on a series of testing with various combinations of values of model parameters. This method requires a great deal of experience of the model author, while it remains an open question whether a solution with the optimal set of parameters is obtained.

In order to avoid a participation of subjective judgment of researcher and as well because of a large number of possible combinations of model parameters, the need arose for a more efficient approach to the process of model calibration. Automatic calibration ensures these requirements. It optimizes the values of selected parameters of the hydrogeological system, where the default criterion, i.e. objective function, is minimizing the difference between the registered values and results of the model simulation.

The industry standard for automatic calibration is PEST program (Model-Independent Parameter Estimation and Uncertainty Analysis). It is the world's most advanced software package for parameter optimization of any simulation model and analysis of prognostic calculations uncertainty. PEST is successfully used in many scientific fields: hydrogeology, hydrology, hydraulics, geophysics, geomechanics, chemistry, aeronautics, mechanical engineering, biology, etc. (Doherty, 2013).

Based on given criteria, PEST changes the model parameters until the output results of the simulation approximately coincides with the measured data (levels, flows) (Doherty, 2010). The optimization method is based on the Gauss-Markar-Levenberg's algorithm of finding the minimum sum of the residual squares (the difference between the measured data and the results of model simulation). In the case of hydrodynamic models, optimization of parameters is an iterative process.

A number of parameters can be set, i.e. their optimal numerical values within the set criteria can be determined: for an unlimited number of zones of hydrogeological parameters (hydraulic conductivity - horizontal and vertical values), specific yield, specific storage, and for an unlimited number of types, features and sections of the boundary conditions (intensity of infiltration and

evapotranspiration, evapotranspiration depth, flow rates or piezometric levels, conductivities of Head-dependent boundary conditions, etc.)

Although there is a possibility of simultaneous determination of a number of parameters, a result which is obtained using the conventional PEST is within the framework set by the author of the model. This applies to dimensions, number and shape of the zones of some parameter which don't change during the calculation. The values of parameters (e.g. hydraulic conductivity) within a predefined zones are the constant values, and so are the predefined values of the homogeneous field with given parameters, which is not the case in nature (Figure 1).

Pest with regularization

One of the ways that this drawback of classical PEST can be overcome is made by introducing a completely new concept that involves geostatistical data processing and application of Kriging method in interpolation between control points (Tonkin and Doherty, 2005). These points are called Pilot points, and do not necessarily indicate the fictitious points, although they are much less frequently points where some value and parameters to be determined are known (Doherty, 2003). In this type of automatic calibration, each control point represents a parameter to be determined. The concept of measuring points is not limited to the values of the hydraulic conductivity, it applies to all types of parameters that are determined during the calibration of the model.

There are several ways of assigning control points on the model: through regular network (Figure 2-a) and through triangulation (Figure 2-b).

The process of implementation of the control points in the PEST is called regularization, i.e. regulation of distribution of certain parameters and / or values of the boundary conditions. This approach in PEST calibration gives a result that includes heterogeneity of the porous media (groundwater flow area) regarding the distribution of parameters that were determined (Doherty et al, 2010). Figure 3 below shows the traditional result of model calibration (manual or PEST) and PEST with regularization.

The duration of the automatic calibration with regularization or control points

One of the biggest problems with the use of Pilot Point in PEST is the time needed for the calculations. Increase of the number of control points that are entered in the model causes the extension of the time period required for calibration of the model. Each control point PEST sees as a parameter whose optimal value is being determined. Each parameter in the PEST requires one or two simulations per iteration. If, for example, the model has 1000 pilot points, then a 1000 simulations are needed (or 2000, if central differentiation is applied) just per one iteration. A larger number of iterations is often required for each model run during the model calibration. If each model run takes a few minutes, then it takes a long time to complete the entire process. Time duration of this type of calibration is measured by hours, sometimes days.

PEST provides two methods of calculation acceleration: Parallel PEST and BeoPEST. Both methods are based on the simultaneous calibration by connecting multiple computers with a basic idea of increasing processing power that is used in the calculations. The main computer is called the "Master" and all others are called "Slave". Each "Slave" computer must have the directory which contains the identical baseline files for calculations that "Master" computer has. Network supports up to 100 directories, i.e. "Slave" computers. The consequence of this approach to calculation is the reduction of time required for the model calibration.

Parallel PEST allows connecting multiple computers through a network, while BeoPEST is a special version of PEST which ensures that the calculations are carried out simultaneously using the Internet and remote connected computers (so called Cloud Computing) (Hunt et al, 2010). "Slave" computers which are connected to the "Master" computer via Internet may be at known locations or they can be rented computers with the price of \$ 30 and up per day of use.

Another way to use BeoPEST applies only to the "Master" computer, and by adding a "Slave" directory on it, the involvement of the CPU increases.

Another possibility for shortening the time needed for automatic calibration by using PEST with regularization is the use of so-called SVD module.

PEST + SVD (Singular value decomposition) analysis

Singular value decomposition (SVD) is the process which makes the determining of the parameters much easier. Applying SVD reformulates the model calibration by assaying the "super-parameters" set, which are a linear combination of the "real" parameters (Tonkin and Doherty 2005). A set of "super-parameters" is typically much smaller, for example 10-20% of the total number of parameters, which results in a significantly shorter time required for the calculations. For example, if there was 500 pilot points and if the number of super-parameters is defined as 15% of the total number of parameters, then the number of super-parameters is 75. This means that 75 calculations per iteration will be performed by using SVD, instead of 500 calculations per iteration. It's always the question how many "super-parameters" should be selected. A new utility within the PEST - the SUPCALC is used for overcoming this problem. It determines the minimum required and maximum allowed number of "super-parameters" for model calibration. It remains on the authors of the model to choose one of the values of the allowable range of values.

In relation to the classic PEST, SVD has another advantage, which relates to a considerably higher quality of processing certain types of model parameters that have different sensitivities to changes in their values during automatic calibration. However, one of the drawbacks of SVD is a considerably more complex adjustment procedure and the performance of this module.

However, SVD also can become slow if it is necessary to optimize more than 2000 parameters. Then, it is more advisable to use LSQR algorithm (Sparse Equations and Least Squares).

Null Space Monte Carlo (NSMC)

Results obtained by SVD are not unambiguous and necessarily require assessment through the analysis of models uncertainty. One of the best ways is the use of Null Space Monte Carlo (NSMC) technique, which is designed to operate on calibrated models using the SVD with many unknown parameters (Doherty et al, 2001; Tonkin and Doherty, 2009).

NSMC technique allows model author to generate more calibrated models with different sets of mutually adjusted parameters. The results obtained, i.e. the models will differ from each other. Therefore, the basic solution of NSMC is not exactly a solution, but the range of possible solutions (Herckenrath et al, 2011). In case that a criterion of 10 generated models is set, and that only the value of hydraulic conductivity is to be analyzed, the result will be 10 models with different matrix of hydraulic conductivity. Such models (10 of them) can be used for forecasting calculations.

This concept is justified if we go back to the beginning and to the definition of the model as an approximate representation of the real (hydrogeological) system. On the other hand, this is the

dominantly American way of giving solutions of a problem, which is trying to avoid responsibility, and to switch decision-making process to someone else (other services).

Conclusion

PEST is a modular program for optimization of parameters during the automatic calibration of the model, i.e. solving inverse problems. The program has a wide range of applications that surpasses the groundwater modeling. Development of the program PEST introduced the Pilot concept, i.e. the concept of control points that surpasses the concept of homogeneous zones with values of the porous media parameters, or zones with the given boundary conditions. The consequence of this approach is the considerable amount of time required for the execution of the calculation. Over time, within the PEST program, two modules which can shorten the time needed for calculation (Parallel PEST and BeoPEST) have been installed through the interconnection of computers (network cables and via the Internet), i.e. by using a higher percentage of engaged processing power of computers. Another way of shortening the time is the application of the SVD module in PEST, which introduces a "super-parameters", that are in a linear dependence with the original parameters of the model. For the calibrated models, the analysis of uncertainty must be completed, by using the SVD module. A new model called the NSMC (Null Space Monte Carlo), which receives a range of possible solutions, has been developed within the PEST.

Acknowledgemet

Our gratitude goes to the Ministry of Education, Science and Technological Development of the Republic of Serbia for financing projects OI-176022 and III-43004.

Литература - References

- Doherty J. (2001). A methodology for handling desaturated conditions in MODFLOW. Watermark Numerical Computing, Brisbane, Australia.
- Doherty J. (2003). Groundwater model calibration using pilot points and regularization. *Ground Water* 41(2), pp. 170–177.
- Doherty J. (2010). PEST - Model-Independent Parameter Estimation, User Manual: 5th Edition, Watermark Numerical Computing, pp. 1-336
- Doherty J. (2010). PEST - Model-Independent Parameter Estimation, User Manual: 5th Edition, Watermark Numerical Computing, pp. 1-336
- Doherty J. (2013). Addendum to the PEST Manual, Watermark Numerical Computing, pp. 1-294.
- Doherty E.J., Hunt J.R. (2010). Approaches to Highly Parameterized Inversion: A Guide to Using PEST for Groundwater-Model Calibration, Scientific Investigations Report 2010–5169, U.S. Geological Survey, Reston, Virginia.
- Doherty E.J., Hunt J.R. Fienen N.M. (2010). Approaches to Highly Parameterized Inversion: Pilot-Point Theory, Guidelines, and Research Directions. Scientific Investigations Report 2010–5168, U.S. Geological Survey, Reston, Virginia.
- Doherty E.J., Hunt J.R. and Tonkin J.M. (2011). Approaches to Highly Parameterized Inversion: A Guide to Using PEST for Model-Parameter and Predictive-Uncertainty Analysis. Scientific Investigations Report 2010–5211. U.S. Geological Survey, Reston, Virginia.
- Tonkin, M.J., and J. Doherty (2005). A hybrid regularized inversion methodology for highly parameterized environmental models. *Water Resources Research* 41, W10412, DOI: 10.1029/2005WR003995.
- Tonkin J.M., Doherty E.J. (2009). Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. *Ground Water* 48, pp. 1–17. doi:10.1029/2007WR006678