УДК [622.276 : 622.031] : 004.9(045)

An effective field scale reservoir model for history matching and reservoir prediction
Part 2. Parametrical upscaling and calibration

A.G. Madatov

Polytechnic Faculty of MSTU, Higher Mathematics Chair

Abstract. The history matching problem in connection to an effective reservoir model concept is defined as a multi-well data inversion problem for a class of process-driven deterministic reservoir models. A two-stage strategy of the production data inversion has been introduced and considered in detail. The strategy includes parametrical upscaling as a pre-processing part and model calibration as a final part. The computer aided approach has been recommended for the parametrical upscaling. It is based on combination of standard step-wise local optimization techniques and active expert-user interface. The fully automatic approach is suggested for the model calibration process. The relevant data inversion algorithm is developed, which combine multi-start strategies (Evolution and Simulated Annealing) with the multi-start mode of a gradient method. The pilot computer programs are created and tested on 2D testbed and on multidimensional synthetic prototypes of a real history matching problem. The results are analyzed based on set of parametrical upscaling and calibration sessions.

Аннотация. Подбор динамической модели эксплуатации нефтяного месторождения представленной в рамках эффективного подхода, рассматривается, как обратная задача для промысловых данных, накопленных по системе добывающих скважин. Предлагается двухэтапная стратегия решения этой задачи, которая включает параметрическое шкалирование (рационализацию признакового пространства модели) в качестве обязательного элемента предобработки и собственно калибровку на финальном этапе подбора. Интерактивный подход к инверсии рекомендуется использовать на этапе шкалирования. Он основан на комбинировании стандартных пошаговых алгоритмов локальной оптимизации при активном участии эксперта в процессе расчетов. Полученная в результате рационализации упрощенная модель затем подбирается в автоматическом режиме. Для этой цели разработан и реализован алгоритм поиска множества локальных минимумов функции цели, комбинирующий метод симуляции обжига, эволюционные алгоритмы с градиентным методом в мульти-стартном режиме. Соответствующая программа минимизации протестирована на стенде, соответствующем двумерной обратной задаче, а также на примере многомерной обратной задачи подбора динамической модели эксплуатации нефтяного месторождения. Результаты анализируются и обсуждаются.

Key words: history matching, production well data, inversion, upscaling, calibration, model specification, relative permeability
Ключевые слова: Подбор динамической модели, промысловые данные, инверсия, шкалирование, калибровка

1. Introduction

Multi-well reservoir engineering implies having clear ideas about subsurface structure and texture of oil-gas field as well as good understanding of the complex dynamics of the fluid flow process, which occurs during production time (production history). Thus, the successful management and in time control over production is impossible without proper reservoir model identification (*Talukdar, Brusdal*, 2005; *Sylvester et al.*, 2005). According to a general definition (*Montgomery*, 2001) the model identification implies setting up of control model parameters plus defining their variation range and most likely values.

The problem of reservoir model identification is inherently associated in reservoir engineering with Production History Matching (*Bentley*, 1998; *Rahon*, 1996; *Haverl et al.*, 2005). Roughly it consists in adjusting of the relevant control model parameters until the difference between simulated outputs and measured data (data residual[1]) is acceptably small.

Strictly speaking, this fitting process should be formulated as an inverse problem (*Menke*, 1984) in relation to the observed data during production history and production indicators. The methods of the relevant inverse problem solutions consequently depend on the class of reservoir model used for model parameters attribution – population from one side and quality of available data from the other side.

In the first part of this article we have presented the concept of effective reservoir models primarily aimed at quick and interpretable Reservoir Model (RM) identification. This concept aims for reduction of the dimension of an inverse problem associated with history matching by maximising use of *a priori* available information about

---

[1] We will use also term "misfit" as a qualitative measure of a data residual.

reservoir on stages of relevant earth model design, attributing and specification of control model parameters. In particular, use of rather universal functions for describing absolute and relative permeability allows accommodation of geological and litho-facies information about corresponding sedimentation processes in an effective reservoir model. The formal sensitivity analysis and informal RM constraining against local geology settings create prerequisites for practical solutions of history matching problems against available production data. This solution includes two successive stages: parametrical upscaling and calibration of an effective reservoir model.

In this part of the article both problems have been considered in more detail.

2. History matching problem. Review of approaches and their classification vs. scale and stage of reservoir engineering

A history matching can be defined as a reservoir model identification process aimed at getting a given reservoir description which yields a correct simulation of its past performance (*Rahon*, 1996; *Alkan et al.*, 2003). The results of such identification then could be applicable for:

- better understanding of production dynamics at field scale;
- fast forecasting of reservoir performance at field block scale;
- full scale quantitative prediction of production parameters at well-by-well scale.

There are several strategies developed and implemented in reservoir engineering for these purposes (*Portella, Prais*, 1996; *Bentley*, 1998; *Christie et al.*, 2002; *Gosselin et al.*, 2003; *Haverl et al.*, 2005; *Lygren et al.*, 2005; *Mouret et al.*, 2003; *Staples et al.*, 2005; *Stephen et al.*, 2005; *Zhang et al.*, 2003; *Thomas et al.*, 2005). The common point for all of the developed approaches is fitting of production data or production indicators to its synthetic prototype by adjusting relevant model parameters while the misfit between simulated outputs and measured data (data residual) is acceptably small.

Thus, this problem can be treated as a classical ill-posed task in *Hadamard* (1932) sense. Depending on scale, data type and data quality the choice of its solution method lies between fully automatic optimization process (*Al-Deeb et al.*, 2002; *Zhang et al.*, 2003) and fully manual model tuning (*Zeltov*, 1998; *Kanevsaya*, 2002).

Regardless of the model class and strategy of data inversion it is necessary to define the set of model attributes/parameters sensitive to the production indicators/data; their range of variations; their most likely values and the way to adjust them. Converging speed and expected accuracy of the fitting process depends a lot on the type of reservoir simulator. The full scale history matching of cellular 3D reservoir model could take several months, whereas use of production indicators as a rough production history data set and focusing on key reservoir sectors allows reduction in time needed to a few weeks or even days (*Tipping et al.*, 2006).

Which class of reservoir model is suitable for history matching?

What kind of reservoir model specification is optimal?

What are the best mismatch criteria and the ways to achieve them?

In order to get reasonably correct answers on these and related questions it is necessary to redefine the relevant inverse problem in more detail. Namely:

- Formulate the purpose for the history matching at every given stage of reservoir engineering.
- Define the type and amount of production data suitable for the given scale of data fitting.
- Define the consequence of fitting stages.
- Define the type of reservoir model for every stage.

Regardless of the type of reservoir model it is necessary to undertake the following steps in order to prepare it for identification:

1. fix the seismically visible subsurface geometry elements in the form of structural framework, which is supposed to be known with sufficient accuracy for data fitting;
2. initialize as many seismically invisible elements as possible, such as initial water-oil contact position, geo-pressure regime, flow unit architecture and porosity distribution. This could be done based on geology settings, available well measurements and wire-log data analysis;
3. assign the rest of necessary simulation model input by using stochastic or deterministic functions[2], which can serve for controlling of the spatial distribution of reservoir transport properties. Define the set of required coefficients considered as model parameter vector;
4. leave the set of parameters that are most sensitive to the simulation results model as a Control Model Parameter sub-set (CMP vector), assumed to be adjusted during the fitting session. Set up the ranges of their variations and most likely positions inside of them;
5. define the goal function of data fitting, stop criteria and strategy of its achievement;
6. define the method of verification for tuned reservoir models.

---

[2] This could be, for example, generalized functions (geo-pseudos) which control populating of absolute permeability and/or relative permeability values over a cellular reservoir model (see part 1 of this paper).

More rigorous definition of the inverse problem in connection to history matching is given in Appendix 1. Below we consider the relevant data fitting problem in contest of a production life of an oil/gas field and data collected as a production history record. This will allow entry of the effective reservoir model concept into the existing sequence of history matching stages.

2.1. The natural stage sequence of history matching

We suggest distinguishing between three possible stages and their associated purposes of history matching depending on 3 integrity levels of reservoir engineering: 1) entire field; 2) field block – near well locality; 3) well interval.

It is possible also to consider them as a natural sequence of history matching stages as amount and quality of collected data increases during production history. These stages can be specified as following:

1. Identification of the observed production dynamics phenomena for better understanding of reservoir performance at field scale (*White et al.*, 2001; *Salhi et al.*, 2005; *Sarma et al.*, 2005).
2. Generation of quickly updatable prediction of integrated indicators of reservoir performance and draft analysis of different production scenarios at the block/field/near-well location level (*Mouret et al.*, 2003).
3. Getting updatable well-by-well reservoir prediction and risk assessment with 4D seismic data taken into account at the single well interval (*Stephen et al.*, 2005; *Kalkomey*, 1996).

It follows from this list that the details of reservoir model descriptions and quality of related predictions increase with the number of stages. Moreover, the listed tasks must be resolved in its natural order since each of achieved inverse problem solutions will reduce the non-uniqueness of the following one. It is possible of course to start history matching session directly from the third stage, but the success of result in this case becomes questionable due to highly non-unique solution of relevant inverse problem. Apart from "Closed Loop" strategy (*Sarma et al.*, 2005) in practice, this consequence is often just ignored and the history matching process considered only in the context of the final stage (*Gosselin et al.*, 2003; *Alkan et al.*, 2003). Alternatively the seismic time-lapse data are using at the first stage to let relevant surrogate model be fitted (*Calvert*, 2005). Such desequencing inevitably leads to failure in the unique and stable inverse problem solution or/and to unjustified expense in data acquisition.

The first stage is associated with creation of low level ("black box") field scale reservoir model based on Experimental Design approach to identification of the corresponding Response Surface Model (*Myers, Montgomery*, 1995). Such models appeared to be practically helpful at well planning and field scale production control (*Salhi et al.*, 2005; *Carrears, Turner*, 2006).

The final stage is normally based on coupled geo-mechanical & fluid-dynamical reservoir models (*Chin et al.*, 2002), which require both 4D seismic data and multi-well static & dynamic data to be identified. The inversion of such kind of multi-scale data normally requires inversion-on-inversion strategy (*Bornard et al.*, 2005) and huge computer resources.

The second stage, i.e. the generation of quickly updatable reservoir prediction, is not considered at all as an independent problem. Evidently the second stage is perfectly fitted to the development of an effective reservoir model concept. The closest prototype to this definition could be found in Computer Aided History Matching combine approach (CAHM) (*Pedersen et al.*, 2005; *Mouret et al.*, 2003). Here the General Theory of Bayesian modelling is applied for data fitting (*Glimm et al.*, 2001; *Alkan et al.*, 2003).

In the context of the reviewed stage list it is important to stress that the result of low level reservoir model identification could be considered as an input to the next stage, where effective level reservoir model could be defined and identified. Then the output of the relevant inverse problem solution could be appropriate for entering into the next stage, where the most detailed level of reservoir models is investigated.

2.2. The type and specification of real data to be fitted at every stages of history matching

The amounts and types of the optimal dataset are obviously specific at each of the considered above stages. In agreement with introduced data fitting purpose the data specification can be listed as follows (see also *Kanevskaya*, 2002):

1. Entire field: field scale geology setting; 3D seismic data; wire log data and production indicators such as oil/gas in place, produced oil/gas, water flooding; OWC/GWC/OGC positions, pressure regime, direct seismic wave attributes etc. integrated on field block and over entire production time scale.
2. Field block- near well locality: field block scale geology & stratigraphy setting; 3D seismic data; petrophysics static data; wire log data and production data such as oil/gas/water production rates, pressure regimes, injection fluid flow rate etc. integrated over entire borehole of production and injection wells and represented in year-by-year form along full interval of a production history.
3. Well interval: detailed scale geology & stratigraphy setting; 3-4D seismic data; petrophysics static data (framework – wire log data) and production data to be included: well-by-well production rates and formation pressure differentiated along production units; well-by-well injecting fluid flow rates; seismic time-lapse data differentiated along production horizons.

According to this list the integrity level of static and dynamic data decreases as the stage number increases.

2.3. Type of reservoir models to be tuned at every stage of history matching

The correct sequence of history matching stages accounts for logic hierarchy of the model levels from low to high level since it supports the natural data assimilation in an updatable model. Indeed, using production indicators instead of full scale production history data set and focusing on key reservoir sectors instead of considering every production well allow saving a lot of time and money. Therefore validated factors of low level reservoir model (like reservoir framework and texture type, generalized pressure regime, initial oil reserves and WOC/WGC position) could be fixed at later stages or used there for initialization of higher level reservoir model (*Furre et al.*, 2003).

We suggest the following sequence of reservoir models to be used at history matching as the scale level scan through three following stages: field, well vicinity, well interval:

1. Low level speculative reservoir model based on formal incorporation of most sensitive production factors like Experimental Design approach with its different modifications: Surface Model Response, Black Box model, Factors' model etc. (*Joreskog et al.*, 1976; *Cheong, Gupta*, 2005; *Carrears, Turner*, 2006).
2. Effective level process-driven reservoir models based on a few highly sensitive and clearly interpretable control model parameters (*Alkan et al.*, 2003).
3. High level and coupled (geo-mechanical & fluid-dynamical) reservoir models based on large set of multidisciplinary model parameters, which control multiple populations of transport properties according to Bayesian theory (*Chin et al.*, 2002; *Sarma et al.*, 2005).

Below each type of the model is considered in more detail.

2.3.1. Low level speculative reservoir model

Since the relevant initial stage of production history matching could be started right after the start of a production process, the main objectives of the model identification are only general setting of geology and petrophysics. The following features of the field scale reservoir model should be evaluated at this level:

1. Subsurface geometry of the field including fault tracing and tectonic blocking in form of structure framework.
2. Litho-facies type of the main reservoir and its sedimentation environment.
3. Flow unit subdivision of the main reservoir in form of its texture type.
4. Oil-gas in place reserves and initial position of relevant fluid contacts with draft estimation of transition zone thickness.
5. 3D porosity $Po(x,y,z)$ and geo-pressure $P(x,y,z)$ distributions applicable for purposes of initialization of higher level reservoir model.
6. Well skin factor estimations averaged over field sectors.
7. Type of generalized functions to be used for specification of absolute and relative permeability: $K_A(Po;\underline{x})$ and $K_R(S;\underline{x})$ together with enter pressure function $P_E(S;\underline{x})$, where S and Po are saturation and porosity respectively; $\underline{x}$ is the CMP vector.

2.3.2. Effective level reservoir model

This level of details in a reservoir model description assumes quick inversion of the relevant scale data: oil/gas/water production rates; pressure regimes; injecting fluid flow rate. This is prerequisite of generation of real time updatable prediction scenarios. The following features of the field block scale reservoir model should be identified at this level:

1. Set of independent control model parameters (CMP vector) together with attached range of variations and most likely values for each of them.
2. Intervals of noticeable sensitivity for each component.
3. Results of parametrical upscaling in terms of a texture scenario and relevant CMP vector with optimized dimension.
4. Set of possible calibration results (local solutions of inverse problem) ranked according to their consistency, stability and fitting quality.
5. Validated set of relevant prediction scenarios, which ensure forecasting of future reservoir performance within the range of the given history matching scale.

2.3.3. High level - coupled reservoir models

This level of detail is supposed to be maximal since it corresponds to the last stage of history matching, where maximum data becomes available for data inversion. It is also involves obligatory use of 4D seismic data. Since the details of model description is very high the dimension of relevant inverse problem is also very high. Normally this requires changing the type of model from deterministic to stochastic. Conformally, the inverse problem strategy should be changed at this stage from local to global optimization (*Dueck, Scheuer*, 1990; *Lygren et al.*, 2005).

The following results should be achieved in the high level reservoir model identification:

1. Stochastic model parameterization in elasto-plastic and fluid transport scales.
2. Optimally simple multi-dimensional search window within the chosen model parameter space.

3. Updatable global solution of relevant inverse problem in terms of maximal likelihood of a reservoir model and relevant production risk assessment.

2.4. Review of available history matching strategies

The rigorous definition of inverse problem associated with production history matching is considered in Appendix 1. Here we review available approaches within the range of the above-introduced stage sequence and corresponding reservoir model types. The natural stage sequence is preserved in this review and the type of RM to be identified at the given stage is used for sub-titling.

2.4.1. History matching at low level speculative reservoir model

The scale and stage of this level require combination of quite different RM factors related to both framework and texture parts. They are selected by only one important sign: high level of sensitivity to the production result at field scale. For example, averaged thickness of reservoir layer, OWC position and transition zone could be incorporated with residual water saturation and/or anisotropy of absolute permeability field (*White et al.*, 2001; *Tipping et al.*, 2006).

Normally, the upper rank of the sensitivity includes up to 5-10 key factors attributed as control model parameters. Therefore, the dimension N of relevant inverse problem posed with regards to integral production indicators is adequate, i.e. $N \approx 5\text{-}10$ (*Cheong, Gupta*, 2005; *Friedmann et al.*, 2003).

Normally the modifications of normalized L_2 misfit measure are using here as a generalized data fitting stop criterion. The fitting with consequent fixing of adjusted model parameters is implemented as a common approach. For example, an averaged reservoir thickness and initial WOC position are tuned first versus oil in place at the given averaged porosity and 3D structure model. Then these framework model parameters are fixed and the boundary flow conditions computed based on estimation of geo-pressure regime in field scale. Finally residual water/oil saturation is adjusted based on sector averaged total production indicators (*Cheong, Gupta*, 2005).

The manual expert-based fit could be used at first stages for tuning of averaged field scale framework parameters such as effective reservoir thickness; sealed fault positions; initial OWC positions etc (*Kanevskaya*, 2002). The sort of multiple regression methods like Experimental Design (Model Surface Response) and/or sort of Factor analysis (*Joreskog et al.*, 1976) could be implemented on later fitting stages.

Historically the modifications of an Expert System approach were the first at identifying of field scale models. Strictly speaking, they cannot be considered as history matching processes, since no quantitative misfit estimations were applied. Some semi-empirical criteria like single sweep draining effect were used for tuning of expert-based key model parameter. In most of the cases such approaches were exclusively related to the given expert and were focused on specific field or even on some problem areas within the specific field (*Reynolds et al.*, 1996; *Bentley*, 1998). As a result they suffer from locally restricted validity.

The more advanced are Experimental Design and Model Surface Response methods (*Box, Draper*, 1987; *Montgomery*, 2001). Here the multiple regression and/or polynomial fitting tools (*Wong, Boerner*, 2004) are used to get the simple mathematical link between low level model factors predefined by expert and relevant synthetic data computed by 3D reservoir simulator. Ideal are the factors which ensure close to linear link with related production indicators. Then a so-called *surrogate equation* is used to find formal link between model factors and production indicators in form of simple linear or quadratic functions of normalised model parameters:

$$Y = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n \qquad \text{(Plackett-Burman design)}$$
$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_{11}(x_1)^2 + b_{12}(x_1 x_2) \ldots + b_n x_n \qquad \text{(D-optimal design)}$$

Series of simulations should then be performed in agreement with predefined empirical design to produce a combinatorial matrix, which is then used to fit coefficients in relevant surrogate equation by using for example least square technique. Three most popular empirical designs predefined for 3D orthonormalized reservoir model are displayed in fig. 1.

Not all of the identified model factors or regressors are equally sensitive to the production indicators. The multiple sensitivity analysis allows fixing of less important factors and removing them from further prediction.

The evident benefit of Experimental Design approach is its computational simplicity and ability to incorporated physically incomparable model parameters in common formal model surface response.
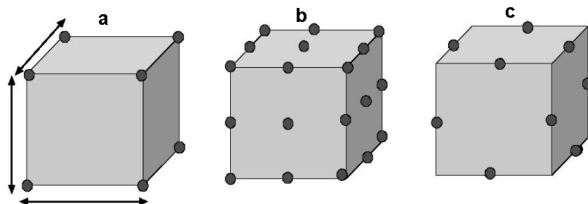

Fig. 1. Most popular empirical designs implemented for 3D orthonormalized model
a – Minimax design (requires 2^3 simulation runs);
b – Box-Behnken design (requires 3^3 simulation runs);
c – Median design (requires 12 simulation runs)

The drawbacks of this approach are typical for any "black box" system analysis. Namely, the fitted coefficients of a surrogate equation are not physically interpretable. Hence, there is no way for posterior ranking of calibrated models. In addition, the accuracy of linear or quadratic approximation is very low and obviously is not enough to match more differential production history data like well production rates.

2.4.2. History matching at effective level reservoir model

As mentioned above, this level of reservoir model representation and identification is intermediate between pure formal low level and pure stochastic high level. There are no industry-available tools, for generation of such kind of reservoir models, which simultaneously can describe production history at multi well scale and are suitable for quick production data inversion and real time (few days) updatable prediction at the relevant scale.

So, in connection to the considered topic we will refer here on the closest prototype of the effective reservoir model concept – Computer Aided History Matching approach (*Pedersen et al.*, 2005; *Mouret et al.*, 2003; *Alkan et al.*, 2003).

The main purpose of history matching at this level of analysis is not only to match the recent production data of the field but also to create a model parameter set which can be globally representative for effective production forecasts. Relevant dynamic data set is represented here in form of multi-well production / injection rates. The standard reservoir simulators are normally used for its prototyping. This allows defining a goal function at data fitting in form of a model constrained misfit (see Appendix 1). The specific of history matching at this stage consists in plural approach to the matched model rather than getting single universal match. This imposes strict limitations on dimension of corresponding inverse problem.

Area compartmentalization (*Pedersen et al.*, 2005), or area clusterization (*Mouret et al.*, 2003; *Thomas et al.*, 2005) based on geological or/and seismic information are key procedures of a combine strategy of Computer Aided History Matching.

The relevant work flow generally includes three main steps:
1. Allocation of field compartment /cluster boundaries based on sensitivity analysis.
2. Extraction of few most sensitive model attributes as key model parameters for each field compartment / cluster.
3. Implementation of local or global optimization technique for auto adjustment of above-defined key model parameters.

Although most of field compartment /cluster boundaries can normally be clearly defined as they are coinciding with high sensitivity contours, a complete blocking requires geology interpretation. In particular, the faults (*Rahon et al.*, 1996) or the facial maps (*Ringrose et al.*, 2003; *Thomas et al.*, 2005) could be used as a natural lateral zone boundary. The Computer Aided History Matching strategy seems to be correct in terms of selection and application principles. All minima are accepted as new parameter sets defining "optimal" matches of the global search thus creating "multiple starting points" for local optimization. The leak on the selectivity of the global "parameter sets" seems to be the essential problem while applying evolutionary strategies to the global optimization.

Still, this strategy misses a routine aimed for minimizing dimension of posed inverse problem. Also it could lead to poorly interpretable results of matching when the first step yields quite formal subdivision of a field on sensitive zones and results in mixing of framework and texture features of the reservoir model.

The discussed concept is free from these drawbacks because it distinguishes between framework and textural parts of reservoir model. This allows significant reduction of the model space dimension.

The most important difference of this approach from other strategies reviewed here is the deterministic formulation of a reservoir model. It allows explicit accounting for all available geology – stratigraphy – petrophysics knowledge in the form of framework – textural parts of relevant earth model and implementing of the generalized function concept for populating /defining of absolute/relative permeability (see part 1 of this paper).

2.4.3. History matching at detailed level reservoir model

The detailed level of reservoir model identification requires combination of multi-well production data and 4D seismic observations (time-laps seismic). Therefore a reservoir model is generally considered as hierarchically coupled elastic, geo-mechanic and fluid-dynamic parts, where model parameters are closely related. As a result the dimension, computer complexity and uniqueness of relevant multi-step inversion problem increase dramatically (*Eide et al.*, 2002; *Furre et al.*, 2003; *Zhang et al.*, 2003).

The consequence of models vs. data identifications could be illustrated by the upper level workflow diagram displayed in fig. 2. Here acoustic or poro-elastic model and geo-mechanical model are considered as an external, which accommodate the standard fluid-dynamic model suitable for reservoir simulation (*Gunning, Glinsky*, 2003). Therefore seismic inversion here is treated as an upper level model identification process since it confronts directly the problem of integrating earth model and seismic model with seismic data. The integration is further enhanced when well logs and well production data are incorporated, since it will capture both regional trends in acoustic property variations and it will point constraints at the well location. The wireline log and production well data inversion at the next stage are properly constrained by upper level acoustic model. The 3D

mode of pressure-saturation distribution computed based on internal inversion is used as a feedback for updating of seismically driven external acoustic or poro-elastic model.

The typical strategy of such inversion on inversion process could be roughly shown as follows:

1. generation of random population of acoustic /poro-elastic model parameters;
2. computing of synthetic seismic data prototypes and matching them with real seismic data by tuning of sensitive acoustic/poro-elastic model parameters;
3. producing of geo-mechanical reservoir model based on tuned acoustic/poro-elastic properties;
4. generation of random population of reservoir model parameters whose variations are constrained by upper level geo-mechanical reservoir model;
5. computing of synthetic production data prototype and matching them with real production data by tuning sensitive reservoir model parameters;
6. producing fluid dynamical reservoir model based on tuned flow transport properties;
7. getting updated acoustic/poro-elastic model parameters and go to step 2.

Fig. 2. Global loops' strategy in the upper level workflow form (after *Gunning, Glinsky*, 2003)

The seismic data have good spatial coverage but it is associated with considerable observation errors and inherent restriction in resolution (*Madatov*, 2005). At this the well observations are available only along well traces but have usually much higher precision than seismic data. So, multiple repeating of inversion loops and feeding back of inversion results allows regularization of generally non-stable inverse problem solution. Normally, a Bayesian definition of goal function is addressed to both external and internal inverse problems (*Christie et al.*, 2002).

In Bayesian inversion, the model parameters have to be assigned in the form of a prior Probability Density Function (PDF). The relevant real data have to be linked to these variables through likelihood models involving forward models associated with these observations. The prior and likelihood PDF models uniquely define the posterior PDF which constitutes the solution of inverse problem (see Appendix 1). Since posterior PDFs are not analytically tractable they may be only explored by multiple model sampling and data simulation (multiple realization). Since the dimension of model parameter space at this level of data inversion is pretty high stochastic-based strategy is normally used. In particular Markov Chain Monte Carlo methods are used to determine posterior probability distributions (*Sambridge*, 1999). Such global optimization methods are successfully applicable for maximization of likelihood goal function: Neighbourhood Algorithms (*Christie et al.*, 2002); Simulated Annealing Algorithm (*Goffe et al.*, 1994); Genetic Algorithm (*Goldberg*, 1989; *Mallick*, 1999).

3. Aligning of effective reservoir model complexity to the history matching purposes

Apart from the above-reviewed low level type of reservoir model all other RM types imply multi-parameter and multi-cellular nature regardless of the deterministic or stochastic type of the model specification used. Due to high non-linearity of forward problem operators, which are practically implemented in reservoir engineering, there are two main reasons why a reservoir model built for history matching purposes must be preliminarily simplified – non-uniqueness of data inversion solution and low converging speed of relevant iterative computation at significant time consuming single iteration.

We claim that the first reason is most important. Indeed, the production data inversion has inherently non-unique solution because of finite data accuracy. Practically it means that a set of ε-equivalent solutions of history matching problem should potentially be considered as fitted (calibrated against data) reservoir models (see Appendix 1). When the amount of ε–equivalent solutions is too big and some/all of them are poorly interpretable, reservoir management becomes problematic. The only way for resolving this sub-problem is to reduce the dimension of the inverse problem by restricting the adjustable model parameter components.

The second sub-problem – converging speed – arises because the model-to-data fitting process implies multiple data simulation. In this context it is crucially important to optimize a grid representation of reservoir model, which is going to be fitted.

So, the simplification of reservoir model implies reducing the number of adjustable model parameter components and optimization of RM grid. The prerequisite of such simplification is that the number of model parameters used for specification of some reference reservoir model and number of cells in its fine grid representation are neither controlled nor restricted at the RM initialization stage. As a result they are usually
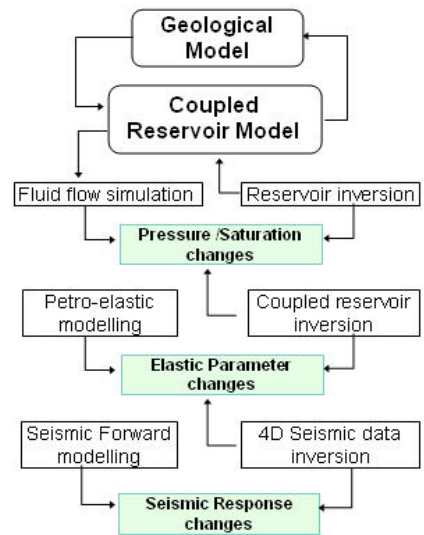
much too large then necessary for accurate simulations. The problem is that the logic "the more detailed reservoir model the better its management" remains most popular. Still the model details and ways of its specification used for data simulation and data inversion are quite different regardless of the type of models and data. Thus a reservoir model complexity must be adjusted depending first of all on simulation purposes. So far, this process in reservoir engineering is normally focused on simulation purposes associated with the terms "homogenisation" or "upscaling" of the initial fine scale reservoir model.

There are two groups of approaches available for a reservoir model upscaling: a grid optimisation group and model parameter reduction group. The related sub-problems are addressed to effective homogenisation of the established fine model in the way which ensures preserving of key heterogeneity features of the analysed reservoir. The difference between existing solutions is in the stage of implementation of the relevant computer procedures. Grid optimisation is normally related to cellular models that are fully ready for simulation whereas the model parameter reduction should be started on one step before, i.e. during RM specification.

Since both approaches are addressed to simulation purposes, the reference sample to be prototyped after upscaling is not real but synthetic reservoir response. From this point of view we will treat these approaches as simulation driven upscaling strategies as opposed to data (inversion) driven upscaling strategy, or parametrical upscaling. Nevertheless it is important to start a discussion on parametrical upscaling with a review of existing simulation driven upscaling strategies, since their main elements are preserved in it.

3.1. Review of available simulation driven upscaling strategies

Here we consider grid optimization and usage of pseudo functions as main representatives of practical upscaling approaches aimed for faster reservoir simulation.

Grid optimization can be understood as formal or process based non-uniform coarsening of an initially uniform and dense RM grid or mesh and redistribution of relevant input values in a new grid. In other words the fine grid porosity, permeability, clay content etc. are recomputed into an effectively modified grid without changing their unit scales.

In contrast, model parameter reduction by using pseudo functions aims to reorganise the specification rules. This could be stochastic RM parameterisation rules based on possible statistical scenarios (*King*, 1989; *Christie et al.*, 1995; *Hristopulos, Christakos*, 1999) when stochastic approach to simulation is implemented. This could be effective control model parameters of generalised pseudo-functions (*Kyte, Berry*, 1975; *Jakupsstovu et al.*, 2001; *Ringrose et al.*, 2003) when deterministic approach is implemented.

The upscaling of RM transport properties which control simulation of single phase fluid flow speed (porosity, permeability and clay content) could be based on both approaches, whereas a relative permeability model, which controls multi-phase fluid flow simulation, is normally specified based on generalised pseudo functions (pseudos) (*Kyte, Berry*, 1975). Apart from simple averaging used for porosity redistribution, both approaches are related to an effective medium theory (*Saez et al.*, 1989; *Pickup et al.*, 1995; *Fokker*, 2001).

3.1.1. Grid optimisation

Grid coarsening implies a wide range of techniques from simple averaging or filtering up to detailed fine-grid numerical flow simulation depending on the petrophysical property to be populated. Two groups of techniques can be considered in this context:

<u>Permeability Guided techniques</u> such as those of (*Verdière, Guérillot*, 1996; *Qi et al.*, 2001) preserve the variation of permeability within the coarse grid resulting in finer gridding around regions of extreme permeability.

<u>Flow Unit Guided techniques</u> such as those of (*Durlofsky et al.*, 1996; *Mahani, Muggeridge*, 2005) refine the grid in areas of high flow-rate.

The key point of the Permeability Guided techniques can be reduced to generation of locally uniform single coarse blocks, where a fine-scaled permeability field can be substituted with a single effective permeability value $\widetilde{K}_A$. In the local upscaling techniques such as Dual mesh methods (*Verdière, Guérillot*, 1996; *Audigane, Blunt*, 2003); Nested mesh method (*Gautier, Blunt*, 1999). In Space Renormalization method (*King*, 1989; *Christie et al.*, 1995) the solution is formulated as follows:

$$\widetilde{K}_A = \frac{\oint_L \underline{Q}\,\underline{n}dl}{L\Delta P}\ , \tag{1}$$

where L and ΔP indicate representative respectively the coarse scale size in the relevant space coordinate direction and cumulative pressure gradient achieved on this distance from the fine scale pressure solution. The contour integral in the numerator of (1) represents the sum of elementary flow components through bounded block border. Clearly, the effective permeability of the heterogeneous medium $\widetilde{K}_A$ is equal to the permeability of the equivalent homogeneous medium which reveals equivalent fluxes for the same boundary conditions.

The main disadvantages of the local grid optimization approach are dependence of the result on local boundary conditions and independence from geology input.

In the Global upscaling approaches (*Cullick, Lake*, 1995; *Holden, Nielsen*, 1998) the entire permeability field is subject to grid optimization. Here the cell shape and coarse-grid architecture depend on features of background earth model. Formally it means that the fine-scale heterogeneities in terms of spatial correlation and variance can be best retained in the coarser grid approximating permeability field. The relevant routine for getting globally upscaled effective permeability field is defined in such way, that the corresponding coarse scale pressure P^{GS} and flux $\underline{Q}^{GS}$ are good approximations of relevant fine scale (reference) pressure $<P^{FS}>$ in and the flux $<Q^{FS}>$ averaged over each *i*-th coarse grid block volume Ω_i and interface Γ_i . At this the coarse grid block is bounded in agreement with prior given flow unit – barrier morphology.

The advantages of Global upscaling vs. Local upscaling methods can by summarized as follows:
- The coarse block boundaries more closely coincide with abrupt changes in permeability.
- The statistical information about permeability inside every geologically bounded coarse block is better preserved (see *Cullick, Lake*, 1995 for more details).

The Flow Unit Guided optimisation approach accommodates predefined earth model architecture as much as possible (see fig. 3). Here the density of cells purposely increases for more accurate flow unit (regions of higher permeability) approximation and decreases within the poorly permeable or self-isolated areas with potentially low volumetric fluid flux. Common approach consists in usage of theoretical sweep efficiency as quantitative indicator of macro level flow unit connectivity for arbitrary heterogeneity and well configuration (*Datta-Gupta*, 2003). The streamline technique is used here first for single-phase forward problem solution. This yields a quick ranking of RM elements based on connectivity flow unit quality. Then the initial fine scale grid can be optimised as follows:
- By using non-Cartesian flow-based system where densification-coarsening of a grid follows to the relevant streamline behaviour (*Durlofsky et al.*, 1996).
- By merging of earth model elements, which have low sensitivity to the simulation results (*Madatov, Sereda*, 2003) or by combining of layers, which have similar sweep efficiency (*Ates et al.*, 2003).

Note that the second solution is especially effective when Finite Elements concept is implemented for simulation, since is does not destroy the whole mesh, but just reorganises the relevant global conductivity matrix (*Sogerlind*, 1976). The level of simplification in this solution can be remarkable. In particularly, it was shown that 93 layers RM can be replaced by 30 without any loss in accuracy (*Al-Deeb et al.*, 2002).

It is important to stress that the appropriate level of a grid optimisation for all approaches should be constrained in agreement with the main upscaling criterion. In particular by critical heterogeneity details, which are still captured (*Moulton et al.*, 1998).
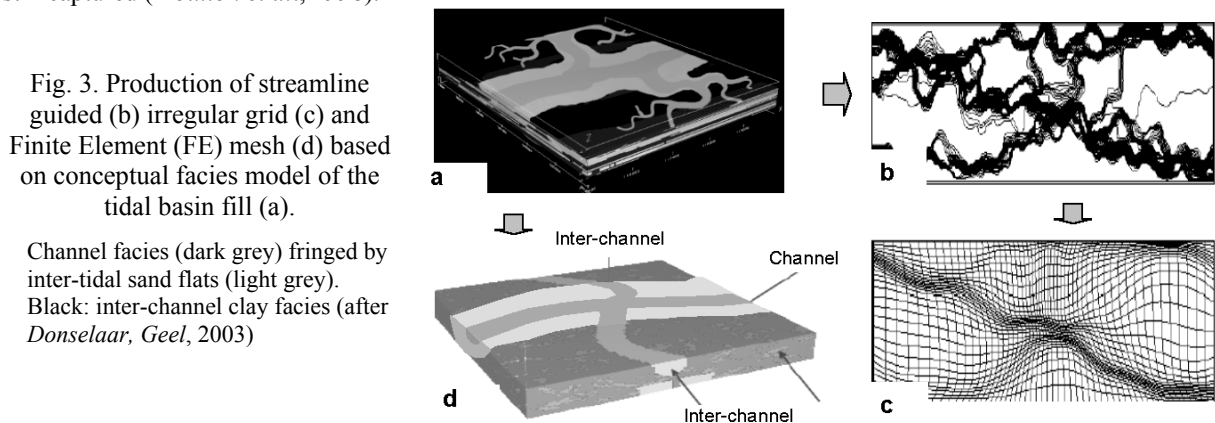

Fig. 3. Production of streamline guided (b) irregular grid (c) and Finite Element (FE) mesh (d) based on conceptual facies model of the tidal basin fill (a).

Channel facies (dark grey) fringed by inter-tidal sand flats (light grey). Black: inter-channel clay facies (after *Donselaar, Geel*, 2003)

3.1.2. Reduction of model parameters by using effective pseudo functions

The key difference of this strategy from grid optimisation methods is the implementation of lithology/facies controlled functional links between cell position (*x,y,z*) and model property to be specified. This leads to substitution of model parameters by arguments of some model functions. These arguments can be treated as control model parameters (components of CMP vector) used for adjusting against dynamic data.

Usage of effective functions (geo-pseudos) with reduced number of coefficients becomes obligatory when multi-phase forward problem is simulated and the relevant RM is matter of calibration. This is because relative permeability and entry pressure parameters are given only in the form of continuous functions of saturation (*Calhoun Jr.*, 1982).

The idea of the pseudo function approach can be clarified with reference on Todd-Longstaff (T-L) method (*Jakupsstovu et al.*, 2001).

Here effective properties are defined as functions of adjustable effective parameters α and ω, where α represents miscibility of the solvent and oil when contacted; ω reflects heterogeneity of porous media at mixing (for example a highly permeable channel between injector and producer would result in poor mixing and low yield ω). Relative permeability is given by a linear combination of straight line function for pure miscible (Kr_m) and pure immiscible (Kr_{im}) relative permeability:

$$Kr_e = (1-\alpha)Kr_{im} + \alpha \, Kr_m. \tag{2}$$

The effective viscosity depends on both α and ω in the form:

$$\mu_{le} = \mu_l^{(1-\alpha\omega)} \mu_m^{(\alpha\omega)}, \tag{3}$$

where l denotes oil, solvent. The mixture viscosity is given by the quarter power law:

$$\mu_m^{(-1/4)} = c \, \mu_s^{(-1/4)} + (1-c) \, \mu_o^{(-1/4)}. \tag{4}$$

In more advanced approaches to the pseudo function construction Bi-spline approximation is proposed instead of linear form (2) (*Okano et al.*, 2005).

In this connection an effective reservoir model concept introduced in part 1 of this paper is entirely consistent with this upscaling strategy. In order to avoid confusion we introduce the term data driven parametrical upscaling (or just parametrical upscaling) in contrast to the above-reviewed upscaling using pseudo functions.
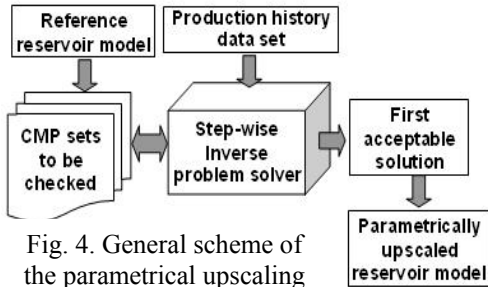
3.2. Data driven parametrical upscaling approach

Fig. 4. General scheme of the parametrical upscaling

Formally, the only difference between the simulation driven and the data driven upscaling strategies is in the treatment of a reference reservoir model to be reproduced with its simplified prototype. Namely: a specific synthetic output of the fine-scale reservoir simulation is the reference for the first group of strategies and a real production history data set should be the reference for the second group. Evidently both reviewed strategies are applicable here. Still, informally, the aim of the data driven upscaling strategy is to reduce as much as possible the relevant inverse problem dimension with preservation of some predefined fitting accuracy level. Thus use of effective functions for RM specification and reduction of their coefficients (CMP vector components) become of primary importance. Note that simulation driven upscaling could be considered as a desired phase of simulation, whereas data driven parametrical upscaling should be considered as an obligatory pre-processing phase at history matching.

The chart in fig. 4 illustrates the main elements of the work flow associated with parametrical upscaling. Here "Reference reservoir model" block denotes framework RM defined on fine scale together with the set of attached texture models. Each set in turn is associated with relevant CMP vector set according to the ERM concept (see part 1 of this paper). The dimension of relevant inverse problem N entirely depends on number of CMP vector components released for adjustment. If the routine starts from minimal N (simplest possible reservoir model) and proceeds toward model complicating up to reference model complexity, then obviously the first acceptable fit will define an optimal solution (see Appendix 1 for more rigorous definitions). At this we assume that the reference reservoir model is detailed enough to reproduce a production history.

The inverse problem solver is based on step-wise local optimization strategy since the first acceptable local solution achieved during the data fitting session yields the desired result. A brief general description of the local minimization algorithms implemented in the IP solver is available in Appendix 2. Some important features of the minimization process related to the particular goal function shape are discussed in section 6. Here the parametrical upscaling approach is described in general.

The design of an effective reservoir model opens two possibilities for its simplification:

1. Simplification of texture model description in terms of reduction of number of necessary flow units and flow barrier involved in specification.
2. Minimisation of number of tuneable control parameters assigned for current texture (1) by blocking of variation range for the least sensitive of them.

First procedure implies merging of several low sensitive elements (flow units or flow barriers) of the initial texture model into one associated with the nearest more sensitive element. From this point of view it could be called *Architecture Upscaling*.

Second procedure implies fixing of the low sensitive components of the model parameter vector $\underline{x} \in \mathbf{X}$ from available after the previous phase. They are normally fixed at most likely positions. From this point of view it could be called *Component-by-Component Freezing* of CMP components.

Below we give more formal formulations of these procedures in agreement with terminology and abbreviation introduced in Appendix 1.

Let us assume the most detailed texture associated with some reference framework model F (Ω) be specified on (x,y,z) grid Ω^0 and Core Model Parameters be defined in agreement with chosen generalized permeability functions. Then the total amount of CMP components is given by $N_0 = \sum_{i=1}^{I} \sum_{k=1}^{K^i} n_k^i$, where "$i$" is the index of specification type which are different for flow unit and flow barriers (see part 1 of this paper), I is the total amount of CPM types ($I=2$); K^i is the total amount of elements related to the same specification type, which are included in the texture; n_k^i is the number of CMP components per element. See Appendix 3 to get an example.

As soon as the standard sorting and normalisation procedure is introduced, every physically possible CPM combination is presentable in the form of dimensionless core model parameter vector as follows (see also Appendix 1):

$$\underline{\mathbf{x}} = \{(X_1,X_2,..,X_n)_1,(X_1,X_2,..,X_n)_2,..,(X_1,X_2,..,X_n)_{K1},\ldots,(X_1,X_2,..,X_m)_1,(X_1,X_2,..,X_m)_2,\ldots,(X_1,X_2,..,X_m)_{K2}\}^T. \quad (5)$$

where $K1$, $K2$ are the total amount of flow units and flow barriers in the given RM architecture; n and m are the number of CMP components required for specification for each of the flow units and flow barriers respectively.

Then the Architecture Upscaling implies reduction of full scale dimension N_0 related to the reference texture down to N_U by reducing of $K1$ and $K2$ conditioned as follows:

$$\underline{\mathbf{x}}_0 \Rightarrow \underline{\mathbf{x}}_u: [\mathbf{M}[\underline{\mathbf{x}}_u] - \underline{\mathbf{p}}^*)] \leq \varepsilon, \; \exists \underline{\mathbf{x}}_U \in \mathbf{X}_U \subset \mathbf{X}_0 \; (N_U < N_0); \; \underline{\mathbf{p}}^* \in \mathbf{P}, \quad (6)$$

where $\mathbf{M}[\underline{\mathbf{x}}]$ is the forward modelling operator defined in the full model parameter space $\mathbf{X}_0$ and all enclosed subspaces $\mathbf{X}_U$ inside of it; $\underline{\mathbf{p}}^* \in \mathbf{P}$ is the data vector defined with the finite accuracy ε (see Appendix 1 and *Madatov, Sereda*, 2005 for more details).

The Component-by-Component Freezing of CMP components does not imply further reduction of dimension N_U but just exclusion of some low sensitive components of the vector $\underline{\mathbf{x}}_u$ from optimisation process without losing condition (6), that can be represented as follows:

$$\underline{\nabla}\underline{\mathbf{x}}_u \approx [0,0,a_1,0,a_2,\ldots.a_m,0,0]^T : [\mathbf{M}[\underline{\mathbf{x}}_u] - \underline{\mathbf{p}}^*)] \leq \varepsilon \text{ and } m < N_U, \quad (7)$$

where m is the number of non-zero components in gradient of the vector $\underline{\mathbf{x}}_u \in \mathbf{X}_U$.

There is a variety of ways for two-level parametrical upscaling process (5-7). For example, in application of a multi-well data inversion to 1D pore pressure data inversion we have developed consequent layer merging routine starting from a reference (most detailed) basin model with obligatory check for satisfying conditions (6-7) at every intermediate phase (*Madatov, Sereda*, 2003). The sensitivity analysis was used here as a guide line at both stages of two-level parametrical upscaling.

In contrast with set of 1D models to be used for data inversion in those calibration routines, we here face a multi-run of full scale 3D reservoir simulations to be used at the production data inversion. From this point of view, it is not considered practically reasonable to begin the Architecture Upscaling from a reference (most detailed) effective RM. Instead we suggest starting the parametrical upscaling procedure from the simplest possible architecture with the minimal possible dimension for the attached model parameter space $\mathbf{X}_U$. Then extension of current architecture with new textural elements or/and involving some more component of CMP vector $\underline{\mathbf{x}}_U \in \mathbf{X}_U$ in model adjustment entirely depends on satisfying the fitting quality conditions (6-7). Evidently, the first effective RM suitable from this view point will be the simplest one. Sensitivity analysis is assumed to be used as a guide line at every steps of this consequent process. Moreover, the result achieved at the inverse problem solution for this model could be treated as one of the possible calibration results.

The following basic elements of the parametrical upscaling algorithm should be highlighted in connection to its practical implementation:

- It should be based on the most detailed reservoir model architecture, in which all available wire log data and geology setting are involved in generation of reference texture.
- It should be started with the simplest possible reservoir model texture.
- It should be finished at the first acceptable IP solution.
- It should not change the reservoir model framework.

The two-stage parametrical upscaling routine satisfying these requirements can be represented by the following steps:

Step 1. Set up a start reservoir model on a fine grid with maximal dimension of the model parameter space $\mathbf{X}_0$ specified in agreement with chosen textural type of the given framework.

Step 2. Perform sensitivity analysis for all available components of $\mathbf{X}_0$.

Step 3. Get simplest architecture $\mathbf{X}_U$ for the given model with the fewest CMP components to be free for adjustment based on minimal number of sensitive components (2).

Step 4. Run fitting session in agreement with conditions (6-7) for the current dimension of the relevant inverse problem and follow one of the following three conditions:

- if the fitting criteria are not satisfied while set of temporarily blocked components is <u>not</u> empty then unfreeze more components of the model parameter space $\mathbf{X}_U$ following condition (7) at the given number of flow elements K_U and repeat step 4;
- if the fitting criteria are not satisfied while set of temporarily blocked components <u>is</u> empty then increase K_U by adding more flow elements following condition (6) and repeat step 4;
- if the fitting criteria are satisfied then go to step 5.

Step 5. Finish two-stage upscaling process and get result as one of possible local solutions of the data inverse problem.

It follows from this algorithm the crucial part of it is sensitivity analysis. The definition of sensitivity vector, description of its computation and specific in application of sensitivity analysis to an effective reservoir model are discussed in part 1 of this paper (section 7).

Fig. 5 illustrates the Architecture Upscaling stage. Here the six-layered texture of reservoir model is considered as a reference one (fig. 5a). Then some variants of its simplification are presented started from simplest (single-layer texture).

Below we discuss the implementation of this approach to synthetic model example.
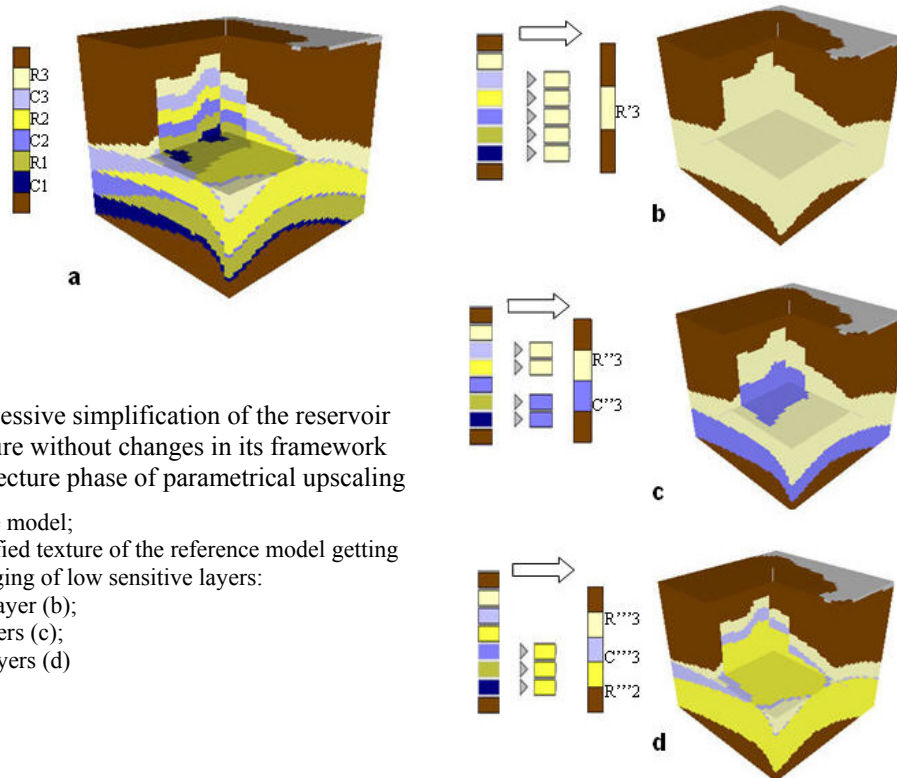

Fig. 5. Successive simplification of the reservoir model texture without changes in its framework during architecture phase of parametrical upscaling

   a – reference model;
   b-d – simplified texture of the reference model getting
      by merging of low sensitive layers:
        single layer (b);
        two layers (c);
        three layers (d)

4. Calibration of an effective reservoir model against production history data

Based on specification rules introduced in part 1 of this paper and on terminology introduced in Appendix 1 we can now define calibration as a data inversion procedure aimed for fitting one or more control model parameter sets (**x** vectors) to the production history data (**p** vector) with acceptable accuracy (ε) and conditioned by all available *a priori* information about the reservoir model. Obviously, such a set of tuned reservoir models can serve for producing of well-founded prediction of reservoir performance with risk assessment (*Christie et al.*, 2002). The philosophy behind an effective reservoir model design ensures clear interpretability for each of the calibrated models. So, the risk assessment here could accommodate some expert evaluations in addition to the statistical summary (*Madatov, Sereda*, 2006).

Since the calibration routine aims for multiple solutions of the specific inverse problem, we suppose that the crucial pre-processing phase – parametrical upscaling introduced above – must be completed before starting the calibration phase.

There are two possibilities for getting a multiple solution of the specific inverse problem:

1. step-wise strategy of the goal function minimisation;
2. multi-start strategy of the goal function minimisation.

The general scheme of the calibration routine shown in fig. 6 illustrates these two possibilities in application to a parametrically upscaled effective RM and a production history data set. Note that the data set used for calibration and upscaling is supposed to be the same.

It follows from this scheme that the critical point which defines choice between the IP solvers type is the method for setting of a start model, or more formally, of position for the vector **x** within the predefined multidimensional search window $\mathbf{X}^A$ (see Appendix 1, section A1.3).

Fig. 6 General scheme of a calibration routine

If start models could be defined by expert, then multiple implementation of step-wise IP solver ensures the desired set of local IP solutions. Such mode in calibration routine can be associated with a computer aided history matching approach (*Pedersen et al.*, 2005; *Mouret et al.*, 2003; *Alkan et al.*, 2003).

If the start models cannot be defined by expert, then only implementation of a multi-start IP solver can guarantee obtaining a multiple solutions' set. Such a processing mode in turn can be associated with fully auto mode of history matching (*Al-Deeb et al.*, 2002; *Zhang et al.*, 2003).

The advantages of the first strategy could come to light as soon as the set of start CMP vectors specified by expert can regularly lead to the desired results. In contrast, the multi-start strategy does not require well-proven expertise for successful implementation, however the price for this is much higher CPU time requirements. Nevertheless, the choice in practice is more often in favour of the fully auto mode.

Brief general description of the step-wise and the multi-start minimization algorithms used in the IP solvers is available in Appendix 2. Some important features for both minimization strategies regarding specifics of the particular goal function shape are discussed in section 6. The formal definition of a calibration problem and the ways of resolving it are available in Appendix 1. Here the calibration algorithm is described generally.

The following four elements of a calibration routine could be highlighted as basics:

- It implements IP solvers for in advance simplified multi-dimensional space of the control model parameters. In other words the parametrical upscaling result(s) in form of an optimally simple effective reservoir model assumed to be prepared in advance.
- The predefined multidimensional search window must be explored during calibration as much as possible in order to avoid missing possible local IP solutions in the final set of calibrated models.
- The calibration results in form of CMP vectors should be comparable in order to get them tested and ranked[3].
- The calibrated model with the top ranking after verification shall be used for generation of a most likely prediction scenario.

The step-by-step algorithm for computer aided mode of calibration routine (left branch on the scheme) can be represented as following:

Step 1. Get sensitivity analysis results for upscaled model at different positions of the start CMP vector.

Step 2. Choose variants of calibration based on Step 1, where different sets of sensitive components of CMP vector are taken for adjustments.

Step 3. Set up the start point at one of the predefined variants while the list of variants is not empty.
- Run the IP solver.
- Exclude tested variant from the list of calibrations and go to step 3.

Step 4. Collect the successful termination results as calibrated models.

Note, that the start positions for the CMP vector at the beginning of adjustment must be prepared in advance based on expert estimations or/and based on sensitivity analysis. Regardless to the choice, it is important to ensure evaluation of the IP solutions coming from different local minima of the goal functions. To a certain degree the distance between achieved IP solutions represented in the metric of multidimensional model vector space as well as difference in their sensitivity signatures could be treated as a measure of the IP solution diversity (see Appendix 2, s/section A2.2.3 for more detail).
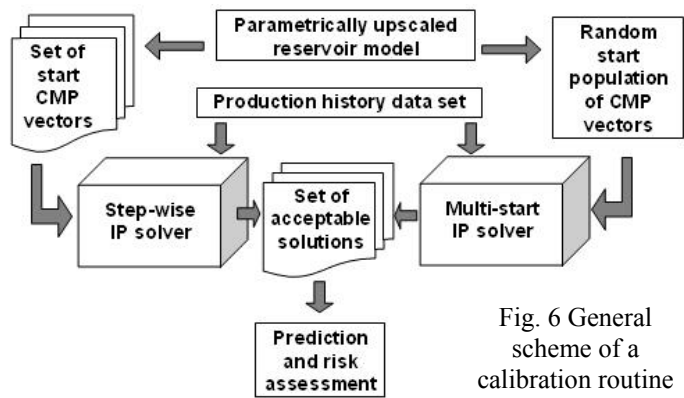
---

[3] This condition can be satisfied by mapping of every of CMP vector-solutions into the data comparison space P in form of synthetic data prototypes, i.e. by implementing a one-by-one transformation $\underline{\mathbf{p}}^{\wedge} = M[\underline{\mathbf{x}}]$.

The step-by-step algorithm for the auto mode which implements multi-start strategy can be represent as following routine:

<u>Step 1</u>. Run the IP solver in multi-start mode.

<u>Step 2</u>. Collect all local solutions achieved within the search window and analyze the diversity of relevant terminal populations.

<u>Step 3</u>. Recognize local solutions, which reveal different sensitivity signatures as a set of calibrated models belonging to the different local minima.

Clearly this three-step algorithm essentially does not require significant user interface during running of the calibration routine. Therefore, the successful termination now entirely depends on multi-start minimization approach adjusted for specific of the multidimensional landscape of the goal function.

Below this specific is considered in 2D approximation in close connection with the signatures of the step-wise and the multi-start minimization strategies.

5. Combining of data inversion algorithms and their tuning by using 2D testbed

The success of data inversion techniques in practice is normally a result of proper choice of a goal function and optimal combination of the well proven algorithms adapted to its signature (*Orlyanskaya*, 2002).

Since the choice of appropriate optimization methods and tuning of their work parameters are very dependent on multi-dimensional landscape of a goal function, each of these algorithms must be first adapted to the goal function peculiarities before being combined in common algorithm. The simplest way for resolving of this problem is testing of involved optimization methods on few-dimensional landscape of a goal function. Note that reduction of relevant model parameter space to two-dimensional is not nonsense for practical applications. Our calibration experience in basin model applications proves that the 2D definition of a data inverse problem could be the case on condition if the elements of the relevant framework are highly sensitive to the data prototypes[4] and the parametrical upscaling went successfully (*Madatov, Sereda*, 2006; *Williams, Madatov*, 2007).

The misfit goal functions $R(\underline{x},\underline{p})$ with soft and/or hard constraints of the model in connection to a history matching problem are introduced in Appendix 1. The standard local and global optimization techniques applicable for the given inverse problem are summarized in Appendix 2. Below we will consider specification of the particular inverse problem and its definition for 2D modification of goal function $R(\underline{x},\underline{p})$. Some peculiarities of the relevant 2D testbed and related features of the standard minimization routines are analyzed as well. Finally we will describe the combine minimization algorithms applicable for the multi-dimensional IP case in computer aided and fully auto modes.

5.1. The 2D testbed for testing and refinement of optimization algorithms

The 2D testbed applicable for the given testing purposes is constructed as goal function array $R(x_1, x_2; \underline{p})$ represented in matrix form and computed with regular steps Δx_1; Δx_2 inside of certain discrete 2D search window $[x_1^{min}-x_1^{max}; x_2^{min}-x_2^{max}]$. Here the pair (x_1, x_2) represents components of CMP vector $\underline{x}$, which are realised for adjustments. Evidently, a grid density of the 2D testbed should be controlled in order to ensure sufficiently accurate interpolation.

Let the size $L \times M$ of the 2D test bed ensure necessary interpolation accuracy. Then it is necessary to run $L \times M$ simulations for getting the array $R(x_1, x_2; \underline{p})$ completed.

For particular test purposes we have used the simplest single-layer texture (see fig. 5b) of reference reservoir model introduced in Appendix 3 together with multi-well production system and schedule. The specification rules introduced for effective reservoir models were implemented (see part 1 of this paper) with linear function for clay content trend approximation. Based on the representative sensitivity analysis we have chosen two most sensitive control model parameters to be released for adjustment (red font digits in Table 1).

Table 1. The CMP component specification for reference model

Specification type	Index	CMP components				
		A0	A1	A2	Log(Ss)	Log(Sm)
Flow units	R1	**0.6**	0.81	0.03	-2.36	**-1.30**
	Index	Uncertain Clay Content Fraction				
Flow barriers	TF1-5	0.2	-0.5	0.2	-0.7	0.0

The intervals opened for adjustments were [0.005-0.5] for component x_1 (Sm) and [0.5-0.7] for component x_2 (A0). The regular steps Δx_1; Δx_2 were defined based on condition MAX(ξ) < 0.1ε, where ε is expected data fitting errors equal to 5 % and ξ interpolation accuracy. Such high demands to the interpolation

---

[4] As for sedimentation rate vs. overpressure (*Magara*, 1978).

accuracy are supposed to ensure substituting of a single forward modelling run $M[x_1,x_2]$ by interpolation into a point with coordinates x_1,x_2 anywhere inside the search window [0.005-0.5; 0.5-0.7]. This in turn accelerates the converging speed for relevant 2D minimization problem by orders of magnitude.

Following inter-cell interpolation was implemented (*Zienkiewicz, Taylor*, 2000):

$$R(x_1,x_2) = R(1,1)Fs1 + R(2,2)Fs2 + R(3,3)Fs3 + R(4,4)Fs4, \qquad (8)$$

where x_1,x_2 is **x** position inside some grid cell 1,2,3,4 (see fig. 7); $R(x_1,x_2)$ is desired quantity value of interpolated function; $R(1,1\text{-}4,4)$ are the reference values of R-function computed for cell nodes (1-4) by using reservoir simulator; $Fs1\text{-}4$ are shape functions given by:

$$\begin{cases} Fs1 = 0.25*(1+xxr)*(1+yyr); & Fs2 = 0.25*(1-xxr)*(1+yyr); \\ Fs3 = 0.25*(1-xxr)*(1-yyr); & Fs4 = 0.25*(1+xxr)*(1-yyr); \end{cases} \qquad (9)$$

where $xxr = (2.*xx - x1-x0)/(x1-x0)$; $yyr = (2.*yy - y1-y0)/(y1-y0)$. The node numbering is given in agreement with consequence of the cell bypass (see fig. 7a).

As it is illustrated by fig. 7b the logarithmic scale for sampling of component x_1 (Sm) appeared to be fitted to the interpolation accuracy purposes much better than the linear one.

Finally the required density of the 2D testbed grid [0.005-0.5; 0.5-0.7] was defined according to the following discrete axis representation:

$$\begin{cases} x_1: \text{Log}(Sm_i) = -2.3 + 0.1667i \,, i = 0,1\ldots12 \\ x_2: A0_j = 0.5 + 0.01j \,, j = 0,1\ldots21. \end{cases} \qquad (10)$$

The quality of inter-cell interpolation was tested for all nodes of this grid by comparison of WOPR curves computed with industry standard reservoir simulator and WOPR curves interpolated according to formulas (8-9) (see example in fig. 7b). The single (reference) and double (test) densities of the grid were used for this accuracy control. The maximal misfit between simulation and interpolation did not exceed 0.5 %.

Each particular *ij*-th WOPR curve was considered as data prototype vector **p** computed in agreement with *i,j* position of model vector **x** within the search window. Any of vectors **p** was given by its components $WOPR_k$ along discrete time grid according to index *k* (see fig. 8b). All together 273 synthetic data prototype vectors $\mathbf{p}_{i,j} = M[x_j,x_j]$ ($WOPR_k$ curves) were computed to fill up the testbed grid.

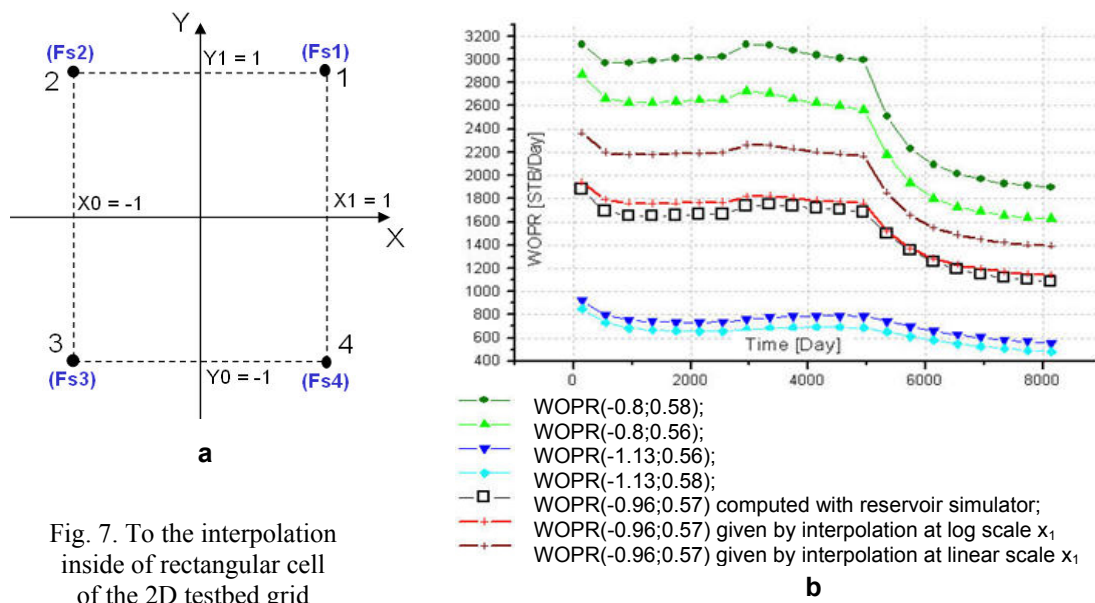Fig. 8 represents the testbed grid and $\mathbf{p}_{i,j}$ set in form of plots and maps.

Fig. 7. To the interpolation inside of rectangular cell of the 2D testbed grid

a – single cell with the nodes (circle) and attached shape function (blue) marked; b – comparison of synthetic data curves (WOPR) computed with standard simulator (black markers) and interpolated WOPR curves (brown and red)
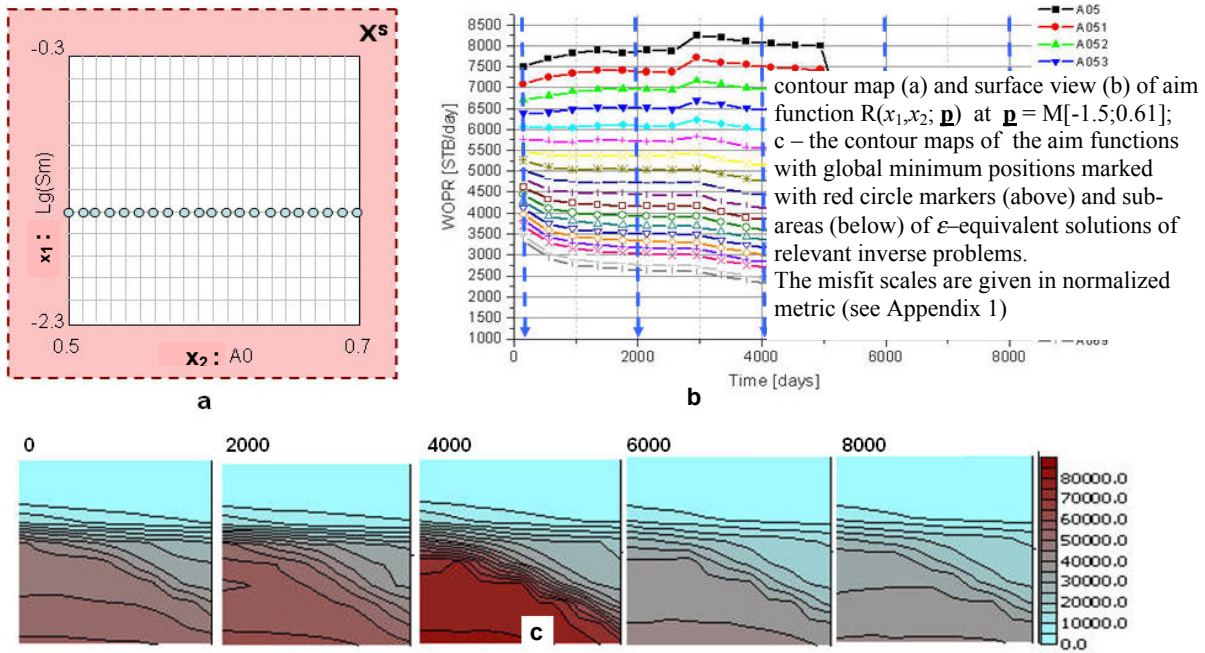
contour map (a) and surface view (b) of aim function R(x_1,x_2; **p**) at **p** = M[-1.5;0.61]; c – the contour maps of the aim functions with global minimum positions marked with red circle markers (above) and sub-areas (below) of ε–equivalent solutions of relevant inverse problems.
The misfit scales are given in normalized metric (see Appendix 1)

Fig. 8. Simulation of the synthetic data set {WOPR$_{ijk}$} for 2D testbed
a – regular testbed grid with row i=5 /x_1 = -1.47/ highlighted;
b – Set of WOPR curves in evolvement along x_2 range (curve code) and x_1 = -1.47;
c – {WOPR$_{ijk}$} slices at five fixed time levels (k=constant in days) indicated as codes of contour maps

a – contour map and b – surface view of aim function R(x_1,x_2; **p**) at **p** = M[-1.5;0.61]; c – the contour maps of the aim functions with global minimum positions marked with red circle markers (above) and sub-areas (below) of ε–equivalent solutions of relevant inverse problems.
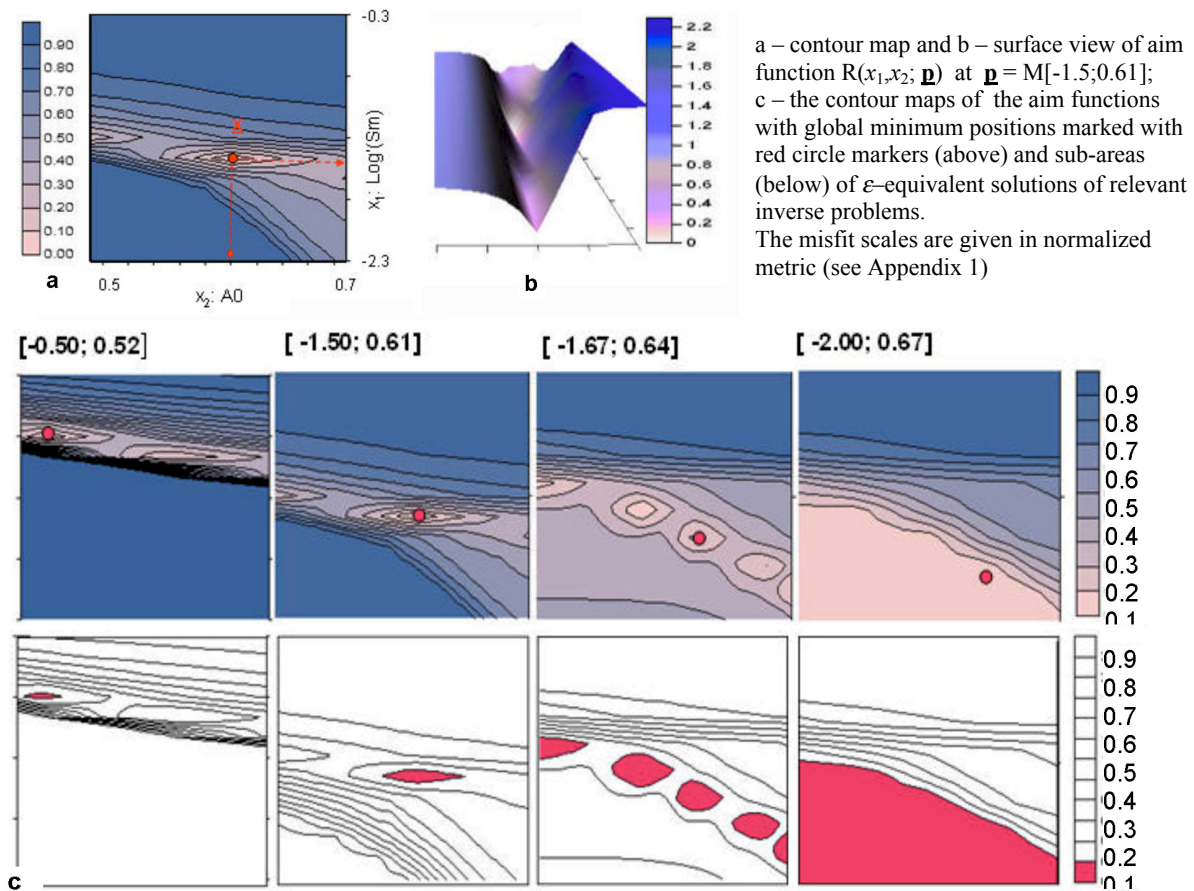The misfit scales are given in normalized metric (see Appendix 1)

Fig. 9. Contour maps of the 2D aim functions computed at 4 reference positions of model vector (x_1,x_2) (code of maps)

510

The slices $WOPR_k(x_1,x_2)$ displayed in fig. 8c reveal significant non-linearity in the synthetic data variation within the range of control model parameters.

Fig. 9 illustrates variety of goal function shapes at four arbitrary positions of the global minimum, which correspond to four positions of reference model vector (x_1,x_2) used for computation of real data prototype **p** (code of maps). As it is clearly visible, the goal function is multi-modal with the minima following along the bottom of narrow gully in $R_{i,j}$ (**p**) relief. Note that these features hold similarity at changing of (x_1,x_2) positions and regardless to the choice of two sensitive components of model parameter vector **x**. Evidently all 273 of possible vector **x** positions inside the 2D testbed represent full range of variants of $R_{i,j}$ (**p**) function shapes. All of them were investigated by using available optimization routines.

Note that the non-uniqueness in solution of minimization problem for Ri,j (**p**) develops as the reference model vector coordinates move toward the edges of the searching window: x_1^{min} and x_2^{max}. For example, the sub-areas of conditionally unique inverse problem solution at fitting quality ε=0.1 include more and more local minima and finally all of them are merged into a single non-resolvable sub-area of ε–equivalent solutions (low right map in fig. 9c). Evidently, it is necessary to include additional components of model vector into a data fitting process to increase resolution of the inverse problem solution in this case.

5.2. Testing and analysis of the standard optimization methods

The 2D testbed described above was used for testing and refinement of the different optimization methods. The standard step-wise optimization methods that were tested are component-by-component descent, gradient related methods and variable metric or quasi-Newton methods (*Eigenstat, Walker*, 1994), as well as multi-start methods – random search with self-learning and Evolution Algorithms (EAs). See Appendix 2 for more details.

Note that the specific of using the 2D testbed allows controlling of entire trajectories for any of the step-wise descent strategies as well as a dynamic of any multi-start session implemented for the optimization. This appeared to be extremely helpful for analyzing of successful and unsuccessful process terminations. From another side this creates the temptation to set up a start point for step-wise routines in the favourable position regarding to the local minima configuration. Of course such input in form of the local minima configuration is only possible during the testbed exercise. The random generation of the start points was used at the testing of step-wise procedures in order to exclude a subjective component. The multi-start methods originally use random generation of start points/populations, so a subjective component here is excluded by definition.

5.2.1. Step-wise strategies

The unimodal configuration of the goal function $R_{i,j}$ (**p**) within the search window represents a most favourable situation for local minimization process. This occurrence happened at about 25 % cases of global minimum positions, which are entirely defined by i,j positions of the reference model. Still, even at such a best configuration of the goal function $R_{i,j}$ (**p**) about 35 % of gradient minimization attempts and more then 50 % component-by-component minimization attempts fail due to the small step increment and/or capturing of the descent trajectory by the search window edge (see fig. 10a). The situation becomes even worse when the goal function $R_{i,j}$ (**p**) is multi-modal with the set of local minima in its shape stretched along the narrow gully (about 75 % cases). We believe that this situation is quite general. In other words the multi-modal shape of the goal function $R_{i,j}$ (**p**) is going to be a typical regardless of the choice of two components of the **x** vector required for generation of a similar 2D testbed. Some examples of unsuccessful termination are shown in fig. 10.
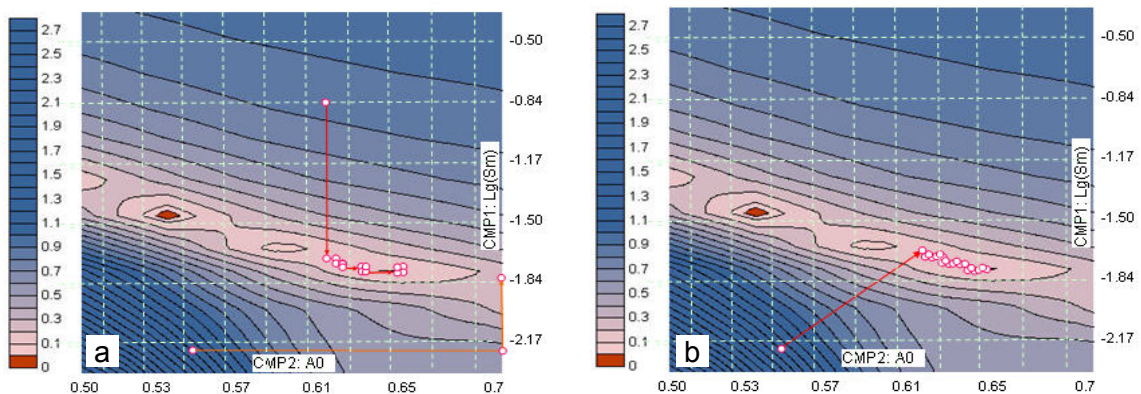

Fig. 10. Examples of unsuccessful termination of step-wise descents on testbed at the global minimum setting (-1.45; 0.54) from different start points.
a – component-by-component method: attraction by secondary local minimum (red arrows); slicing by the edge at wrong start component (orange arrows); b – conjugated gradient method: attraction by secondary local minimum (red arrows).

The variable metric routines commonly reveal extremely high sensitivity to the precision of the $R_{i,j}$ (**p**) gradient computation and rather low converging speed (see Appendix 2). So, the results of their testing on the 2D testbed were estimated as negative.

A common feature of all tested step-wise procedures is a quite sharp approach to the area where the goal function $R_{i,j}$ (**p**) has low values, i.e. to the *R*-gully proximity. At this the gradient related methods have incontestable advantage. At the next descent steps both component-by-component and conjugated gradient methods reveal approximately equal low effectiveness. Taking into account that the component-by-component descent is part of conjugated gradient methods, we made choice in its favour during scanning at the *R*-gully proximity. This is mainly because of better CPU time requirements at other equal characteristics.

In connection with the component-by-component descent strategy we suggest using *DSK Powell* method (*Press et al.*, 1992) with the linear acceleration of the step length $q(k)$ and parabolic approximation of a solution locality when a goal bracket is found (see Appendix 2 for more details).

The random descent strategy with self-learning (in particular, Kelly-Wheeling's algorithm *Kelly, Wheeling*, 1962) was also tested in step-wise modification. Its implementation to the particular shape of the goal function $R_{i,j}$ (**p**) was estimated as unlikely successful.

The plot in fig. 11 represents a converging speed summary of the step-wise strategy applications to the minimization of the 2D goal function $R_{i,j}$ (**p**) at different testbed configuration depending on current choice of the global minimum coordinates (x_1,x_2): **p** = M(x_1,x_2).

We have used a misfit level equal to 0.10 as the data stop criterion for fixing of the total amount of required goal function calculations during minimization. This quantity represents the argument axis on the plot. The number of achieved successful terminations was analyzed against this argument. The column bars on the plot are spaced inside intervals only for demonstration. The last interval [200-250] on the horizontal axes includes also the cases when amount of required goal function calls exceeds an upper limit (250 calls) and causes termination of the further calculations.

The conjugated gradient descent reveals the best converging speed. This ensured about 30 % of minimization session being successfully terminated inside the first interval (goal function calls is less then 50). The component-by-component descent reveals slightly worse, but comparable converging quality.

The random descent reveals approximately equal distribution of successfully terminated runs against required goal function calls.

The simple gradient descent reveals the worst converging speed. Here most of the runs were unsuccessfully terminated due to exceeding of the goal function call's limit.

Evidently the best result in terms of converging speed could be found in a combination of gradient steps with component-by-component steps on condition that the latter are made along the currently most sensitive direction. Thus, the following triple-step strategy of using step-wise routine can be recommended:

1. Single gradient /conjugated gradient step.
2. Sensitivity analysis.
3. Single component-by-component step along most sensitive direction.
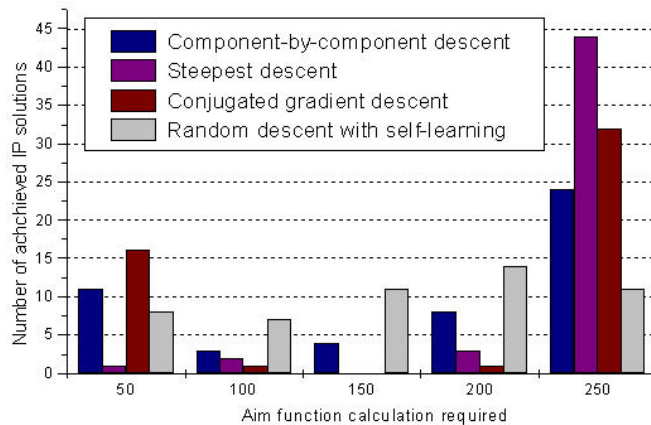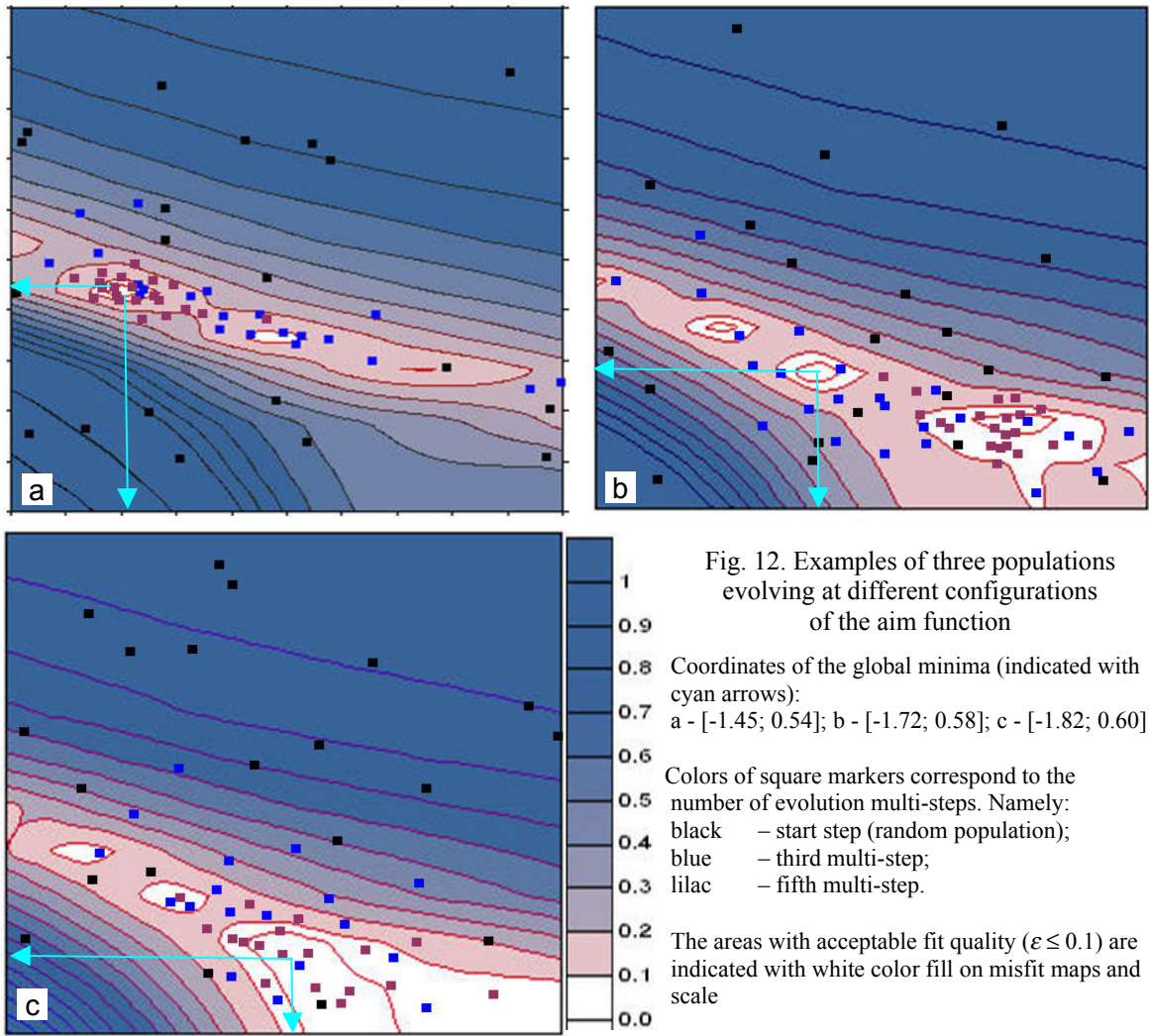
Obviously the application of this

Fig. 11. Converging speed summary based on 50 minimization sessions on the 2D testbed at the different global minimum positions and start point settings

strategy to the real multi-dimensional inverse problem could require some corrections and some user experience so we suggest using it just in computer-aided modification of a data fitting process at the parametrical upscaling and at the interactive calibration (see below).

5.2.2. Multi-start strategy

The Evolution Algorithms (EAs) were tested for multiple random testbed realizations. The set of standard EA tools includes following routines (see Appendix 2 and review (*Bäck*, 1996) for more details):

- <u>Generation</u> of the random multi-start set $\{\underline{\mathbf{x}}^0\}$;
- <u>Selection</u> of the best individuals $\{\underline{\mathbf{x}}^I\} \subset \{\underline{\mathbf{x}}^0\}$;
- <u>Crossover</u> for getting the offspring set $\{\underline{\mathbf{x}}^{I+1}\}$ with adjustable number of parents;
- <u>Mutation</u> of the offspring set $\{\underline{\mathbf{x}}^{I+1}\}$ with adjustable dispersion.

Fig. 12. Examples of three populations evolving at different configurations of the aim function

Coordinates of the global minima (indicated with cyan arrows):
a - [-1.45; 0.54]; b - [-1.72; 0.58]; c - [-1.82; 0.60]

Colors of square markers correspond to the number of evolution multi-steps. Namely:
black   – start step (random population);
blue   – third multi-step;
lilac   – fifth multi-step.

The areas with acceptable fit quality ($\varepsilon \le 0.1$) are indicated with white color fill on misfit maps and scale

The dynamics in changes of a fitness distribution against stage of evolution process was considered as a main workability criteria of the relevant tested strategy by analogy with converging speed in the step-wise strategies. Additionally, the amount and quality of cached local minima within the testbed search window, the total amount of the aim function calls and the diversity of evolving populations were analysed.

The scatter plots in fig. 12 illustrates a typical distribution of the successively evolving populations (21 members in each population).

The start population randomly distributed within the search 2D window was tightened around the best achieved solutions at each new EA step. These solutions were normally attracted by the first captured local minima. The intensity of the selection after each k-th stage of the evolution was estimated based on standard deviation measure (τ_k) for log-normal fit of raw fitness distributions (*Press et al.*, 1992).

The test summary is presented in fig. 13. The typical result consists of rather high selecting intensity during two first EA steps (Is ~ 1.5-3.0) and then very slow progress in the averaged fitness of a population.

The misfit level $\varepsilon = 0.10$ was used as the data stop criterion. The relevant sub-spaces of ε-equivalent IP solutions are represented by white
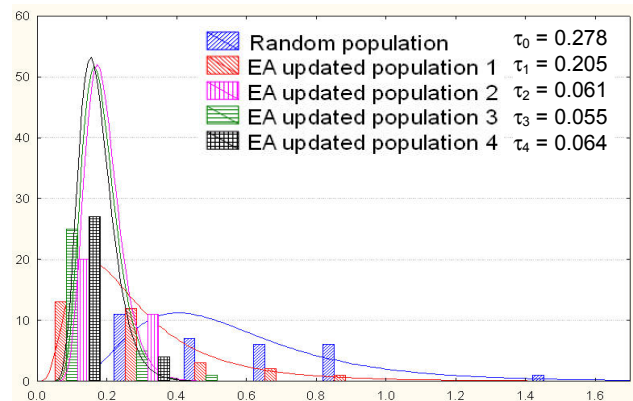

Fig. 13. The log-normal fits (line plots) of raw fitness distributions (columns plots) evaluated for start random population and four evolving populations during using standard EA routine

areas on the aim function contour maps (see fig. 12). Since the fitting results with misfit values less or equal to this limit represent ε-equivalent IP solutions (see Appendix 1) the fitness ranks for relevant members of population $\{\underline{x}^k\}$ were considered to be coincided. It appeared that it is not the deepest but the most area extensive local minimum that normally serves as an attractor of evolution process. This phenomenon is repeated for each of the three presented cases (see fig. 12).

Note that standard EA routine at considered testbed conditions is normally terminated by allocation of a single but not obviously a global minimum.

5.3. Combined minimization algorithm applicable for auto calibration mode

As shown on the 2D testbed tests, no single tested minimization routine can be treated as a perfectly fitted for the purposes of fully auto data fitting process associated with a reservoir model calibration.

Indeed, the RM calibration implies exploring as many local minima of the goal function as possible at the single data fitting session (see s/section 4). All step-wise routines are focused on a single local solution of an inverse problem by their definition (see Appendix 2). They could be implemented at multiple step-wise session, which has different start points, if each new search is well conditioned and allocated within the new search sub-space belonging to the same model parameter space. This Computer Aided History Matching strategy (*Furre et al.*, 2003; *Mouret et al.*, 2003; *Haverl et al.*, 2005; *Staples et al.*, 2005) implies active interface from user and cannot be recommended for auto mode.

From another side, a multi-start EA routine cannot guarantee the desired solution in form of many local minima captured during single session. Additionally, the shrinkage of an EA process toward the local attractor is quite slow. This implies time consuming operation for the multi-dimensional IP solution.

So, there is a need for development of a new IP solution algorithm, which should be specifically adjusted for auto history matching purposes. In particular this algorithm must be capable of allocating not a single but a multiple IP solution in terms of number of local minima explored while the single optimization session. At this, the converging speed should be acceptable.

Such a new algorithm was developed based on combining of the multi-start and the step-wise strategies, where the EA strategy was used as a basis one.

There are two key ideas behind the developed combined algorithm:
- implementing of the multiple gradient steps for acceleration of shrinkage of a start population;
- implementing of the stop criteria scheduling similar to the temperature scheduling in Simulated Annealing method (see Appendix 2) for controlling several populations that are simultaneously attracted by several local minima.

The details of the combined algorithm could be understood from the work-flow diagram in fig. 14.
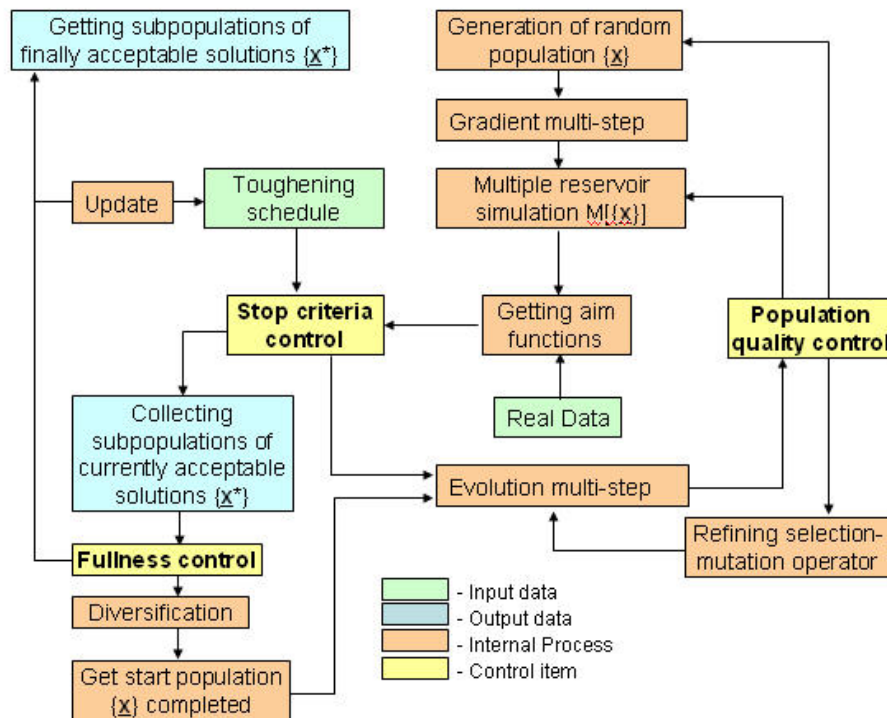

Fig. 14. Upper level work-flow diagram of the combined multi-start inversion algorithm
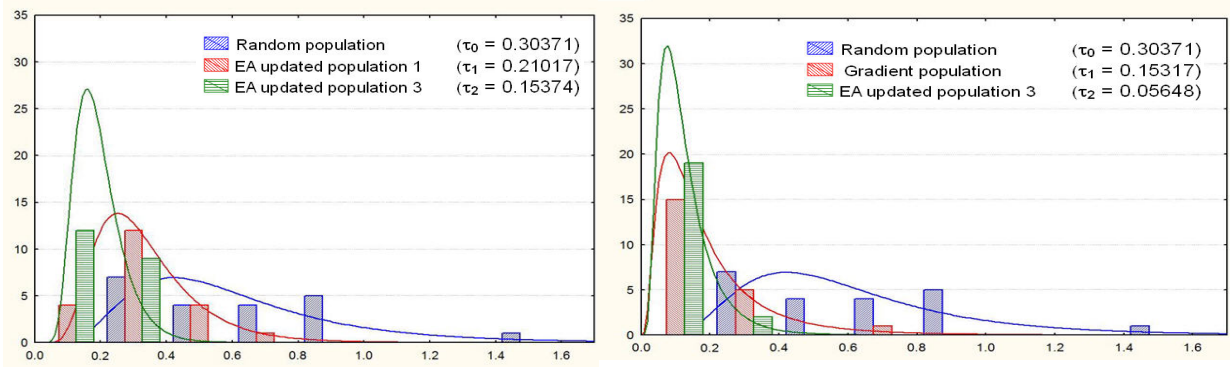
Fig. 15. The log-normal fits (line plots) of raw fitness distributions (columns plots)
evaluated for 31 member's populations produced by standard EA routine (above)
and newly developed combined EA routine (bellow) during first three multi-steps

According to this general scheme the minimization routine begins from generation of start population (subset $\{\underline{x}^0\}$) distributed randomly within the search window limits. Next step consists in applying a single gradient step to every member of the start population, i.e. in performing "Gradient multi-step" for the subset $\{\underline{x}^0\}$. The average fitness of the start population after this stage is normally increased much faster than after the first standard EA step (see fig. 15-16).

The result is then entered into the main internal loop of the process, which includes following sequence of operation:
- multiple reservoir simulation;
- multiple misfit estimation together with relevant setting of the fitness rank for each $\underline{x}$ belonging to $\{\underline{x}\}$;
- passing by stop criteria control where the currently acceptable solutions are collected and excluded from the further refinement;
- entering into the next evolution step.

Every evolution step includes a triple operation according to the EA method: selection parents – crossover – mutation of offspring population (see Appendix 2 for more details). This operation aims at improvement of an averaged fitness of output population $\{\underline{x}^{k+1}\}$. Some side effects are also possible, like compression of population around a best representative, which does not obviously deliver solution. This phenomenon is associated with collapsing of population diversity. Thus population quality control, which includes control of averaged population fitness and diversity, is an obligatory step before proceeding toward the next evolution step. The relevant control item ("Population quality control") allows elaborating of certain feedback correction for the subsequent evolution step when it is necessary. Namely, we suggest following feedback link between current population diversity $Ds(k)$ and dispersion level σ_{k+1} for the mutation operator, which is coming next in agreement with EA strategy:

$$\sigma_{k+1} = \frac{\sigma_k}{\left[\dfrac{Ds(k)}{Ds(k+1)} + \xi\right]} ,$$
(11)

where ξ is the small regularization constant.

Evidently the relative decreasing of the population diversity, which accompanies collapsing tendency, will be immediately compensated by increasing of the dispersion constant in the mutation operator, which is implemented to the next following offspring population.

Also the increment in averaged fitness of the current population, the number of individuals (vector $\underline{x}$) belonging to the currently acceptable solution set $\{\underline{x}^*\}$ are controlled by "Fullness control" and "Stop Criteria control" respectively (see fig. 14). If both controls indicate too slow evolution process, then global loop could be restarted again from the generation of random population.

The next stage of evolution process – exploring potential minima – can be started as soon as the "Fullness control" reports about exceeding the size of the currently acceptable solution set $\{\underline{x}^*\}$. At the beginning of this stage some diversification of solution population $\{\underline{x}^*\}$ should be performed.

The essence of this procedure is detecting of all potential local minima that have manifested themselves in the population $\{\underline{x}^*\}$. The right hand column in fig. 16 illustrates the case where three potential minima have attracted three sub-populations. This phenomenon is accompanied by generation of three sub-populations inside
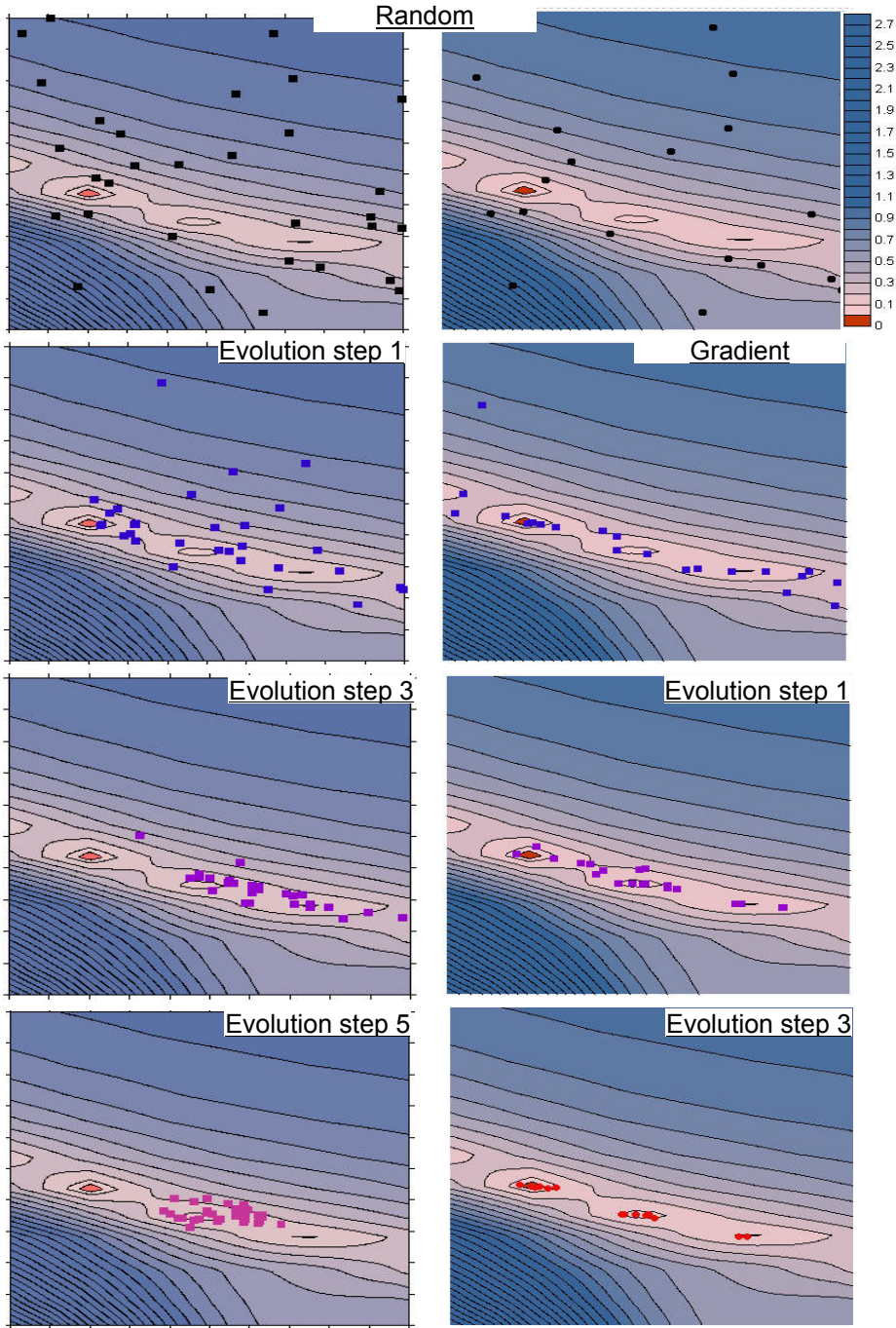
515

Fig. 16. Examples of testbed exploring by using standard (left) and combine (right) evolution algorithms

Population size is 31 and 21 for standard and for combined EA algorithms respectively. The total amount of the aim function calls are: 155 for standard and 189 for combined EA algorithms respectively.
The final data stop criterion is indicated by red color on contour maps and misfit scale

acceptable solution set {$\underline{\mathbf{x}}$*} in form of fuzzy clouds separated from each other by empty space intervals. Thus the diversification in this context means detecting of uniform accumulations {$\underline{\mathbf{x}}$*$_m$} which all together constitute population {$\underline{\mathbf{x}}$*} as follows: {$\underline{\mathbf{x}}$*}$_k$ = [{$\underline{\mathbf{x}}$*$_1$} $\cup$ {$\underline{\mathbf{x}}$*$_2$} $\cup$...$\cup${ $\underline{\mathbf{x}}$*$_M$}]$_k$. Here *M* is number of potential local minima. The membership of any given individual $\underline{\mathbf{x}}$* to the given sub-population {$\underline{\mathbf{x}}$*$_m$} is based on criterion:

$$\left(\frac{\left\| \underline{\mathbf{x}} * - \overline{\overline{\underline{\mathbf{x}}}}_m \right\|}{\left\| \overline{\overline{\underline{\mathbf{x}}}}_m \right\|} \approx Ds_m \quad \text{and} \quad \frac{\left\| \underline{\mathbf{x}} * - \overline{\overline{\underline{\mathbf{x}}}}_m \right\|}{\left\| \overline{\overline{\underline{\mathbf{x}}}}_m \right\|} \rangle\rangle Ds_i, \forall i \neq m \right) \tag{12}$$

As soon as the diversification is made the process of further evolution can be started in the form of *M* parallel EA processes for each of the detected sub-populations $\{\underline{\mathbf{x}}^*_m\}$.

This new stage is associated with updating of data stop criterion in the direction of its decrease (see "Updating" of the "Toughening schedule" in the general scheme in fig. 14). This action is quite analogous to decreasing of a control temperature in the Simulated Annealing method (see Appendix 2). Indeed, it reduces the probability that each subsequent evolution step will deviate from local minimum vicinity. The combination of selecting, crossover and mutation normally ensures quite a quick clutch of offspring population $\{\underline{\mathbf{x}}^*_m\}$ toward relevant local minimum (compare upper and lower histograms in fig. 15).

The more separated the sub-populations $\{\underline{\mathbf{x}}^*_m\}$ inside the acceptable solution set $\{\underline{\mathbf{x}}^*\}$ that are collected by the end of the first stage of evolution process (before diversification) the more local minima will be explored at the next parallel stages of evolution process. Which of the potential targets will be recognized as a local IP solution? The answer entirely depends on the "Toughening schedule".

Note that both stages of evolutionary processes can include intermediate loops with collecting population $\{\underline{\mathbf{x}}^*\}$ and the further diversification depending on number of subdivisions of initial stop criterion, i.e. depending on the "Toughening schedule".

In the example shown in fig. 16 the "Toughening schedule" has included two successive levels of data stop criteria: 0.1 and 0.05. This allowed capturing three minima (including two local and one global) by the end of the third evolving population.

So, the key new elements of the combined algorithm which distinguish it from the standard one, are:
1. implementing single gradient multi-step as the first stage of evolution process,
2. diversification of rough solution population and exploring all detected potential minima.

The first element ensures much quicker convergence of the evolution process to the set of potential solutions (fig. 15). The second element allows escaping from trap situation and exploring many potential IP solutions (minima of the aim function) simultaneously (fig. 16).

6. The model examples of history matching in computer aided and fully auto modes

The reference reservoir model with the structural framework in form of blocked anticline and six-layer texture has been used for simulation of the synthetic production history observed in three production wells (see Appendix 3 for more details). The relevant Well Oil/Water Production Rates (WOPR/WWPR curves) were considered as a prototype of the real data vector **p** in connection with obtaining the relevant effective reservoir model parametrically upscaled and calibrated.

6.1. Parametrical upscaling

The parametrical upscaling session according to the general strategy (see s/section 4.2) was started from the single-layer texture. Then the two-layer and three-layer textures have been tested in different layer-merging combinations. The triple-step strategy of local optimisation introduced in subsection 4.3.1 has been used in computer aided mode. The results of these tests are summarized in Table 2. Here the successfully simplified models, which yield final misfit less than 5 % are marked out with the bold font. The best result is additionally indicated with red text.

Some of the test results are also displayed in fig. 17.

The application of the parametrical upscaling routine to the given plausible reference RM proved its workability and revealed some important features, including:

Table 2. The parametrical upscaling summary

No	Texture (in number of layers)	Misfits achieved (%)				Computing details		Comments	Ref. figure
		Total	Wells			Iterations spent	CPU time elapsed, min		
			P1	P2	P3				
1	1	10.4	4.2	24.1	3.1	44	633	10 CMP components adjusted.	17a
2	2	9.8	7.3	19.1	3.2	38	420	6 CMP components adjusted. FT multipliers not sensitive to the result	17b
3	2	5.1	8.3	3.9	3.0	28	272	7 CMP components adjusted.	17c
4	2	6.4	10.1	4.4	4.8	40	489	11 CMP components adjusted.	
5	**2**	**4.0**	7.5	1.9	2.7	12	140	4 CMP components adjusted. FT multipliers not sensitive to the result	17d
6	3	5.6	5.1	6.6	5.3	29	419	9 CMP components adjusted.	17e
7	3	6.4	9.6	4.7	4.8	35	490	7 CMP components adjusted. FT multipliers not sensitive to the result	
8	3	7.0	8.6	6.6	5.9	44	659	12 CMP components adjusted.	
9	3	**4.8**	4.7	5.6	4.0	21	205	6 CMP components adjusted. FT multipliers not sensitive to the result	17f
10	3	**4.3**	6.9	3.4	2.6	19	189	5 CMP components adjusted.	

- A model oversimplified down to the single-layer texture cannot be found systematically with practical data fitting accuracy (~5 %).
- Almost three-fold reduction of the initial model parameter space dimension (from 19 to 7) is achievable based on the two/three-layer texture. At this computer aided mode during parametrical upscaling reveals much higher converging speed then fully auto mode.
- Merging of the layers with the high contrast in permeability systematically leads to worsening of the upscaling results and *vice versa*.
- The best upscaling result can be reached based on the preserved reference texture with repeated layers in it (see fig. 17d). Almost five-fold, from 19 to 4, reduction in the initial IP dimension was achieved in the considered test.
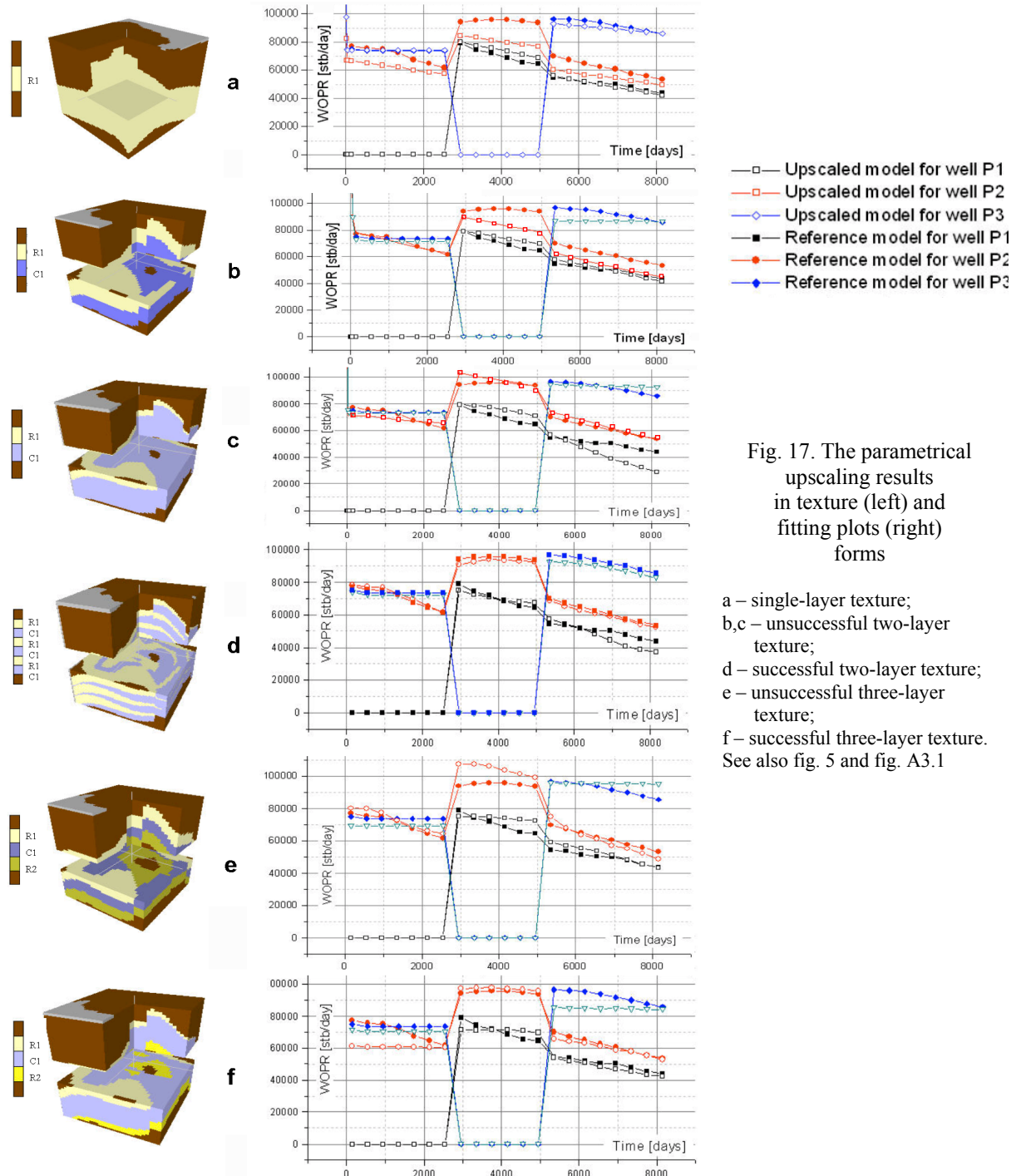

Fig. 17. The parametrical upscaling results in texture (left) and fitting plots (right) forms

a – single-layer texture;
b,c – unsuccessful two-layer texture;
d – successful two-layer texture;
e – unsuccessful three-layer texture;
f – successful three-layer texture.
See also fig. 5 and fig. A3.1

6.2. Calibration

The best result of parametrical upscaling in terms of dimension of the specific inverse problem was achieved for the repeated two-layer texture (see fig. 17d and Table 2). In agreement with this the 11-dimensional search window was used for exploration of the goal misfit function landscape and getting local solutions of the specific inverse problem, i.e. for RM calibration. The reference synthetic production history (see Appendix 3) was used as a real data prototype. Both "computer aided" and "fully auto" strategies were implemented for the IP solutions. The results in the form of CMP vector components, the sensitivity signatures of the relevant solution and the final misfit are represented in fig. 18.

Both computer-aided and auto modes of the calibration allowed recognition of the three local minima within the range of the explored search window at the level of fitting quality $\varepsilon \leq 0.055$. These results were interpreted as prototypes of three prediction scenarios, based on different sensitivity signatures and on distinction in the RM specifications (see fig. 17d and Appendix 3). Namely:

- Scenario A − Highly sensitive flow unit (layer) "R1", low sensitive flow unit (layer) "C1" and non sensitive flow barriers (fault segments). The clay content is given by a common trend.
- Scenario B − Highly sensitive flow barriers (fault segments); low sensitive flow units (layers) "C1" and "R1". The clay content is given by layer-by-layer trends.
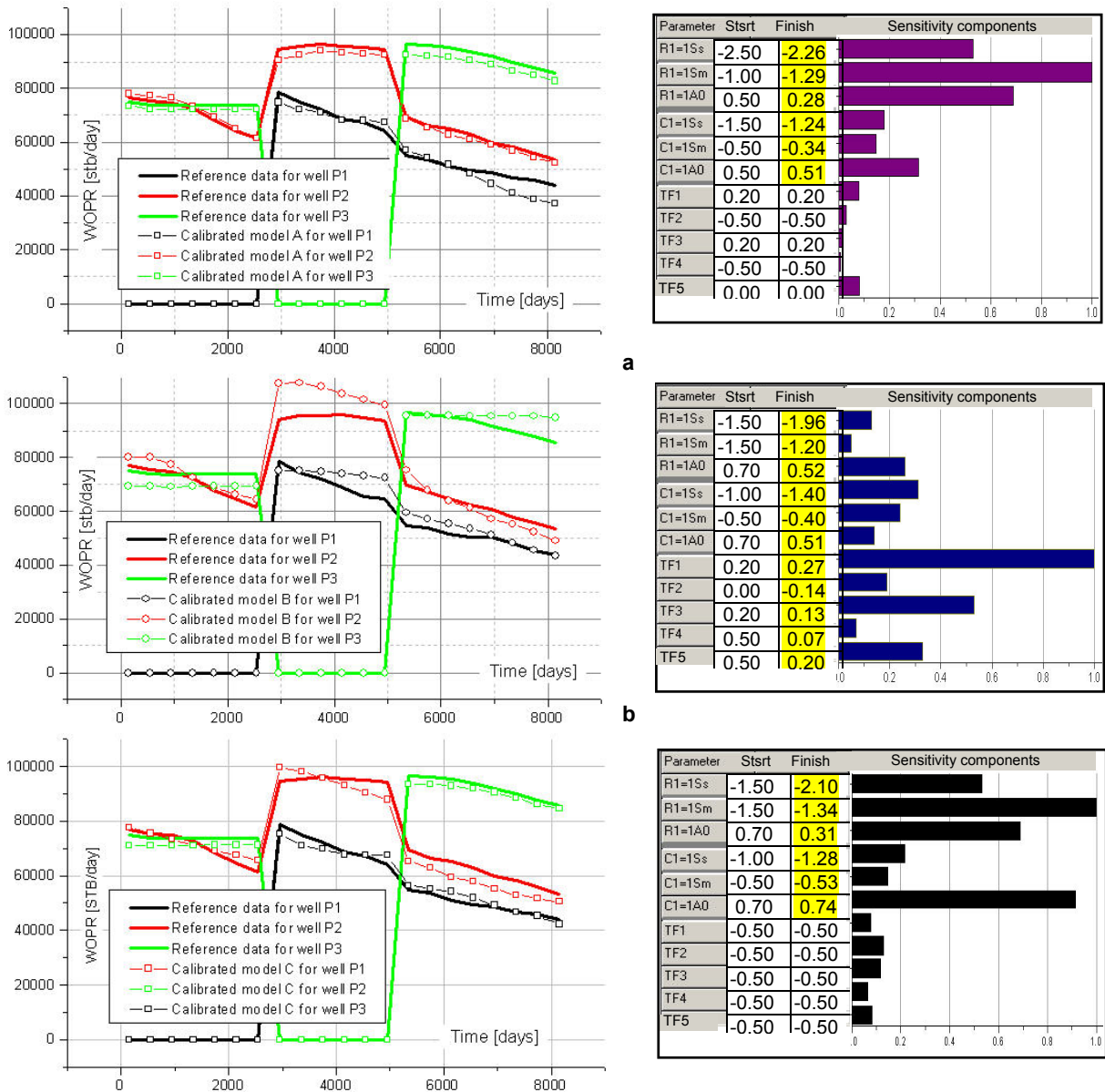

Fig. 18. The calibration results in fitting plot (left) sensitivity distribution (right) forms
a − Scenario A; b − Scenario B; c − Scenario C. The adjusted model vector components are highlighted
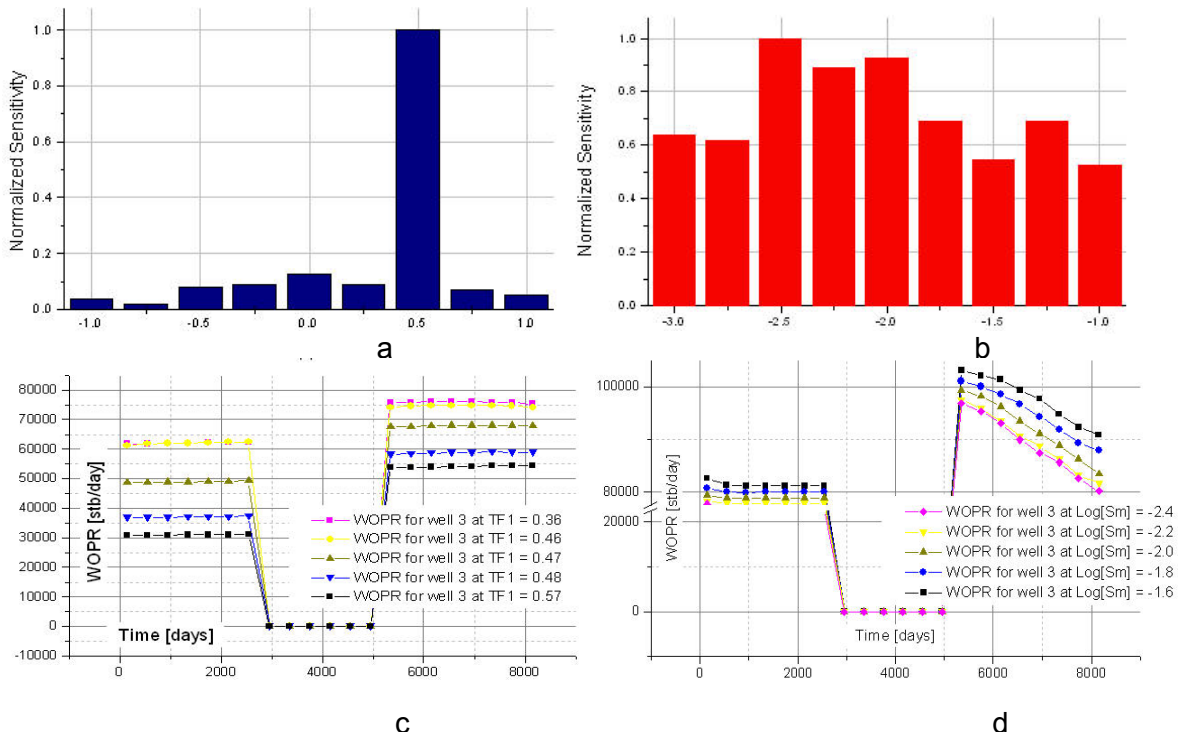
Fig. 19. The results of stability test for prediction scenarios: A (right) and B (left)

a,b – sensitivity distribution along full variation range for the CMP component indicated in the legend of plots c,d respectively
c,d – synthetic WOPR curves (well 3) computed for certain value of CMP component (curve code) from its variation range

Table 3. The fitting quality summary for the calibrated model

Model	Misfit				Calibration Rank
	well P1	well P1	well P1	Total	
A	0.07481	0.01944	0.02697	**0.0471**	**2**
B	0.06821	0.05913	0.04066	**0.0537**	**3**
C	0.04546	0.04748	0.0447	**0.0459**	**1**

Table 4. The validation summary for the calibrated model

Model	Prediction error				Prediction rank
	well P1	well P1	well P1	Total	
A	0.11097	0.20123	0.37390	**0.2283**	**2**
B	0.13339	0.32878	0.30689	**0.2564**	**3**
C	0.17482	0.09291	0.07406	**0.1139**	**1**

- Scenario C – Highly sensitive flow units (layers) "R1" and "C1" and low sensitive flow barriers (fault segments). The clay content is given by layer-by-layer trends.

    The summary of the fitting sessions for all considered models is presented in Table 3.

    The quality ranks of the fitted model were estimated in agreement with the total final misfit.

    Further validation tests performed for the calibrated models A-C have confirmed this rank sequence. The validation tests consisted of checking of the mismatch between the synthetic prototypes of Well Water Production Rate data computed by using of a standard reservoir simulator for the reference RM (see Appendix 3) and the prediction scenarios A-C obtained for the relevant well positions based on the achieved calibration results. The summary of the validation tests is presented in Table 4.

    Both summaries show that model B reveals worst quality based on both data fitting tests and prediction tests, whereas model C has the best ranking for both tests. The final decision about the ranking distribution among the calibrated models should be matter of comprehensive analysis, where model flexibility, consistency and stability should play an important role (*Madatov, Sereda*, 2006).

    Evidently, there are many other options for validation of the calibrated models based on the prediction tests. One can use production history data for just some of the available wells and/or for only part of the time

520

interval. Then the prediction tests could be based on checking of the prediction errors for the temporarily hidden data. We have done this by excluding one of the production well data from the list of real data to be fitted and also by shortening the real history time interval by hiding the last few points. The results did not change the ranking distribution shown in Tables 3-4.

In addition to the prediction tests it is important also to check stability of the calibrated models to the inherent uncertainty of the IP solution. The origin of this uncertainty is connected with the finite accuracy of the data fitting as well as with the tolerant errors in an RM framework, which is supposed to be fixed in advance. We presume that the shape of sensitivity distribution along the full range of variations for the corresponding model vector component is a good measure of the calibrated model stability (*Madatov, Sereda*, 2006).

The plots in fig. 19 illustrate this insight.

It is clearly visible that the sensitivity of a calibrated model for the component "TF1" (transmissibility multiplier for the fault segment 1) is non-uniformly distributed along its variation range, whereas it is quite uniform for component "R1-SM" (specific surface for pure clay component of the flow unit R1). There is a narrow interval with the "TF1" variation range (around TF1 = 0.5), where the sensitivity reveals peculiarity in form of abnormally high value (fig. 19a,c), whereas the sensitivity to the perturbation of "R1-SM" is uniformly high for whole range of its variations (fig. 19b).

We claim that the ideal for prediction model in terms of stability should reveal a high and close to uniform sensitivity level for few components of CMP vector anywhere inside their range of variation. At this point the rest of CMP vector components can be blocked without damaging the calibration quality.

The above-described tests show that both step-wise and multi-start strategies of the goal function minimization can be successfully implemented for the calibration purpose. At this the first option (step-wise techniques) implies an expert-based input for the relevant computer routine with active user interface. The second option (multi-start techniques) ensures auto mode of calibration. The price for the auto mode is naturally related with increased CPU time.

It is important to stress that a calibration process is not finished with several history matched models ranked in agreement with the final misfit levels. The set of calibrated models should also be tested against real production data in order to get them validated and check their stability against tolerant RM errors.

7. Conclusion

This part of the paper concludes the introduction and detailed description of an effective reservoir model concept aimed for the generation, upscaling and calibration of a special kind of reservoir model specifically created for quick identification and updating while real time reservoir engineering.

The rules of specification introduced in part 1 of this paper allows clear interpretation and facile tuning of the relevant effective RM based on a few highly sensitive and clearly interpretable control model parameters. The history matching problem was defined for this class of deterministic ("process driven") reservoir models in terms of multi-well data inversion problem. The relevant fitting strategy in this context implies successive solutions of parametrical upscaling sub-problem and RM calibration sub-problem. We have recognized these tasks to be part of the suite of model identification sub-problems, which forms a consequent history matching process, where an output from early stage of the history matching could serve as an input for later stage. The relevant effective reservoir model built for a field scale in its turn has proper place within the hierarchy of reservoir model descriptions, which naturally become more and more detailed with the increasing of a production life. The suitable niche for the developed class of reservoir models could be found between low level ("black box") and detailed level of reservoir description. Here the reservoir model integrity allows block-by-block (effective) description of the reservoir properties required for simulation. In particular, a framework part (structure level) could be fixed, a data driven petrophysical part could be initialized and the most uncertain and highly sensitive sub-seismic part (flow /barrier unit's properties) could be fitted against production data.

Such special design of an effective RM allows easy optimisation of its complexity against the required matching accuracy by using newly developed parametrical upscaling approach. The main assumption behind this approach is that a finely gridded reference (source) reservoir model is usually excessively detailed in the context of history matching. The core of the relevant strategy consists in consequent simplification of reference thin-layered texture and blocking some of the low sensitive components of the reference model parameter vector space. The relevant computer routine obviously implies an active user interface so parametrical upscaling can be considered as a computer-aided data fitting process.

The optimally reduced dimension of the specific inverse problem opens the practical possibility for exploring multidimensional search window and detecting on it acceptably deep local minima of a goal function. Each of detected minima is then considered as a history matched model. A full set of the appropriate IP solutions forms the range of calibrated against production data effective RMs. Each of these models after validation and stability tests can be ranked and used to produce a prediction scenario.

We believe that the developed approach to the calibration with all associated precautions taken into account creates a solid prerequisite for a well grounded prediction of reservoir performance and risk assessment at the effective phase of history matching.

We have distinguished two approaches to the practical solution of introduced specific inverse problem: computer aided and fully automatic.

At this, the computer aided approach is based on combination of the standard step-wise local optimization techniques. It requires an active expert-user interface. This approach is recommended for parametrical upscaling, which is considered by us as an obligatory pre-processing part of the full scale history matching.

The fully automatic approach can be based on the newly developed combined multi-start inversion technique, which is based on Evolutionary strategy. This approach is adopted for the calibration phase. However it does not exclude implementing of elements of the computer aided approach.

Both approaches were developed in the form of pilot computer programs. The workability of each particular optimization method has been first tested on the simplified 2D testbed. The best methods were then implemented for the parametrical upscaling and calibration of plausible reservoir models. In both cases the synthetic production data were used as real production data prototypes. This implementation was considered as close to real data test, where a multidimensional data inversion problem is supposed to be resolved.

The achieved results could be considered only as a first step in the further development of a new generation of reservoir models suitable for quick history matching, model updating and prediction of a production dynamics on well-by-well level.

Acknowledgements

List of references

Al-Deeb M., Bloch H., Salem El-Abd, Sharfeddine M. Fully integrated 3D reservoir characterization and flow simulation study: A field case example. *Paper presented at the 10 Abu Dhabi International Petroleum Exhibition and Conference*, 12-13 October, 2002.

Alkan H., Pusch G., Junker H. A global search supported cahm case study improved by geoscientific expertise. *Paper D-46 presented at the EAGE 65th Conference & Exhibition, Stavanger, Norway*, 2-5 June, 2003.

Ates H., Bahar A., El-Abd S., Charfeddine M., Kelkar M., Datta-Gupta A. Ranking and upscaling of geostatistical reservoir models using streamline simulation: A field case study. *Paper SPE 81497 presented at the SPE 13th Middle East Oil Show & Conference held in Bahrain*, 9-12 June, 2003.

Audigane P., Blunt M. Dual mesh method in upscaling. *Paper presented on SPE Reservoir Simulation Symposium held in Houston, Texas*, 3-5 February, 2003.

Bäck T. Evolutionary algorithms in theory and practice – evolution strategies, evolutionary programming, genetic algorithms. *New York – Oxford, Oxford University Press*, 385 p., 1996.

Bentley L.R. On incorporating time-lapse seismic survey data into history matching of reservoir simulations. *CREWES Research Report*, 12 p., 1998.

Bornard R., Allo F., Coleou T., Freudenreich Y. Petrophysical seismic inversion to determine more accurate and precise reservoir properties. *Paper 94144 SPE presented at the SPE Europec/EAGE Annual Conference held in Madrid, Spain*, 13-16 June, 2005.

Box G.E.P., Draper N.R. Empirical model-building and response surfaces. *NY, Wiley Press*, 231 p., 1987.

Bridges J.W., Shaker S.S., Madatov A.G. Integration of geophysical data into a basin geopressure model results in improved pore pressure prediction. *Paper presented at the 2001 Offshore Technology Conference held in Houston, Texas*, 30 April-3 May, 2001.

Bunday B.D. Basic optimization methods. *M., Radio-Svyaz'*, 128 p., 1984 (in Russian).

Calhoun Jr. J.C. Fundamentals of reservoir engineering. *University of Oklahoma Press, Norman*, 426 p., 1982.

Calvert R. Insights and methods for 4D reservoir monitoring and characterization. *SEG and EAGE issue, Tulsa, Okla.*, 219 p., 2005.

Carrears P.E., Turner S.E. Tahiti: Development strategy assessment using design of experiments and response surface model. *Paper SPE 100656 presented at Western Regional/AAPG Pacific Section/ GSA Cordilleran Section Joint Meeting held in Anchorage, Alaska, USA*, 08-10 May, 2006.

Cavicchio D.J. Adaptive search using simulated evolution. *Unpublished doctoral dissertation, University of Michigan, Ann Arbor*, 187 p., 1970.

Cheong Y.P., Gupta R. Experimental design and analysis methods for assessing volumetric uncertainties. *SPE Journal*, September, p.324-335, 2005.

Chin L.Y., Thomas L.K., Sylte J.E., Pierson R.G. Iterative coupled analysis of geomechanics and fluid flow for rock compaction in reservoir simulation. *Oil & Gas Science and Technology – Rev. IFP*, v.57, N 5, p.485-497, 2002.

Christie M., Subbey S., Sambridge M. Prediction under uncertainty in reservoir modelling. *Paper presented on 8th European Conference on the Mathematics of Oil Recovery, Freiberg, Germany*, 3-6 September, 2002.

Christie M.A., Mansfield M., King P.R., Barker J.W., Culverwell I.D. A renormalization-based upscaling technique for WAG floods in heterogeneous reservoirs. *Expanded Abstracts (Society of Petroleum Engineers)*, p.353-361, 1995.

Cullick A.S., Lake L.W. Global scale-up of reservoir model permeability with local grid refinement. *Journal of Petroleum Science and Engineering*, v.14, p.1-13, 1995.

Datta-Gupta A. Ranking and upscaling of geo-statistical reservoir models using streamline simulation: A field case study. *Paper SPE 81497 presented at the SPE 13th Middle East Oil Show & Conference held in Bahrain*, 9-12 June, 2003.

Denesh A. PVT and phase behaviour of petroleum reservoir fluids. *Amsterdam-Tokyo, Elsevier*, 400 p., 1998.

Donselaar M.E., Geel C.R. Reservoir architecture model for heterolithic tidal deposits. *Paper P034 presented at the EAGE 65th Conference & Exhibition, Stavanger, Norway*, 2-5 June, 2003.

Dueck G., Scheuer T. Threshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing. *J. Comp. Phys.*, v.90, p.161-175, 1990.

Durlofsky L.J., Behren R.A., Jones R.C. Scale up of heterogeneous three dimensional reservoir description. *SPE Journal*, v.1, p.313-326, 1996.

Eide A.L., Omre H., Ursin B. Prediction of reservoir variables based on seismic data and well observations. *Jour. Am. Stat. Assoc.*, v.97, N 457, p.18-28, 2002.

Eigenstat S.C., Walker H.F. Globally convergent inexact Newton method. *SIAM J. Science. Optimisation*, v.4, p.393-422, 1994.

Fokker P.A. General anisotropic effective media theory for the effective media permeability of heterogeneous reservoirs. *Transport in porous media*, v.44, p.205-218, 2001.

Friedmann F., Chawathe A., Larue D.K. Assessing uncertainty of channelized reservoir using experimental design. *SPEREE*, August, p.264, 2003.

Furre A-K., Munkovold F.R., Nordby L.H. Improving reservoir understanding using time-lapse seismic at the Heidrun Field. *Paper A-20 presented at the EAGE 65th Conference & Exhibition*, Stavanger, Norway, 2-5 June, 2003.

Gautier Y., Blunt M.J. Nested gridding and streamline-based simulation for fast performance prediction. *Computational Geosciences*, v.3, p.295-320, 1999.

Gilks W.R., Richardson S., Spiegelhalter D.J. The Markov chain Monte Carlo in practice. *London, Chapman & Hall*, 256 p., 1996.

Glimm J., Hou S., Kim H., Sharp D.H. A probability model for errors in the numerical solutions of a partial differential equation. *Computational Fluid Dynamics Journal*, v.9, p.485-493, 2001.

Goffe W.L., Ferrier G.D., Rogers J. Global optimisation of statistical functions with simulated annealing. *J. Econometrics*, v.60 (1/2), p.65-100, 1994.

Goldberg D.E. Genetic algorithms in search, optimization, and machine learning. *Addison-Wesley: Reading, MA.*, p.74-88, 1989.

Gosselin O., Aanonsen S.I., Aavestmark I., Cominelli A., Gorand R., Kolasinski N., Ferdinandi F., Kovalcic L., Neylon K. History matching using time-lapse seismic (HUTS). *SPE 84464 paper presented on SPE Annual Technical Conference and Exhibition, Denver, Colorado, USA*, 5-8 October, 2003.

Grefenstette J.J., Baker J.E. How genetic algorithms work: A critical look at implicit parallelism. [ICGA3], p.20-27, 1989.

Grefenstette J.J. Optimization of control parameters for genetic algorithms. *In IEEE Transactions on Systems, Man and Cybernetics*, v.16, N 1, p.122-128, 1986.

Gunning J., Glinsky M. Bayesian seismic inversion delivers integrated sub-surface models. *Paper D-35 presented at the EAGE 65th Conference & Exhibition held in Stavanger, Norway*, 2-5 June, 2003.

Hadamard I. Le probleme de Cauchy et les equations aus derivees partielles lineaires hyperboliques. *Hermann edition*, 71 p., 1932.

Haverl M.C., Aga M., Reiso E. Integrated workflow for quantitative use of time-lapse seismic data in history matching: A North Sea field case. *Paper SPE 94453 presented at the SPE Europec/EAGE Annual Conference held in Madrid, Spain*, 13-16 June, 2005.

Himmelblau D.M. Applied non-linear programming. *McGrow Hill Book Company*, 540 p., 1972.

Holden L., Nielsen B.F. Global upscaling of permeability in heterogeneous reservoirs; The output least squares (OLS) method. *Tech. rep., Norwegian Computing Center, Oslo, Norway*, 1998.

Holland J.H. Adaptation in natural and artificial systems. *Univ. of Michigan Press, Ann Arbor. Reprinted in 1992 by MIT Press, Cambridge MA*, 90 p., 1975.

Hristopulos D.T., Christakos G. Renormalization group analysis of permeability upscaling. *Stochastic Environmental Research and Risk Assessment*, v.13, p.131-160, 1999.

Ingber L. Simulated annealing: Practice versus theory. *Math. Comput. Modelling*, v.18, p.29-57, 1993.

Jakupsstovu S., Zhou D., Kamath J., Dourlofsky L., Stenby E. Upscaling of miscible displacement processes. *Proceedings of the 6th Nordic Symposium of Petrophtsics, Trondheim, Norway*, 15-16 May, 2001.

Jong K. An analysis of the behaviour of a class of genetic adaptive systems. *Doctoral dissertation, University of Michigan, Dissertation Abstracts International, 36(10),* 5140B, University Microfilms No. 76-9381, 1975.

Joreskog K.G., Klovan J.E., Reyment R.A. Geological factor analysis. *New York, Elsevier SPC*, 223 p., 1976.

Kalkomey C.T. Use of seismic attributes as predictors of reservoir properties: Potential risks. *66th Ann. Internat. Mtg., SEG, Expanded Abstract*s, 1996.

Kanevskaya R.D. Math modelling of reservoir engineering. *Moscow-Igevsk*, 140 p., 2002 (in Russian).

Kelly R.J., Wheeling R.F. Digital computer program for optimizing non-linear functions. *Mobil Oil Corp., Research Dept. Pricenton*, July, 1962.

King P.R. Use of renormalization for calculating effective permeability. *Transport in porous media*, v.4, p.37-58, 1989.

Kirkpatrick S., Gelatt C.D., Vecchi M.P. Optimization by simulated annealing. *Science*, v.220, p.671-680, 1983.

Kyte J.R., Berry D.W. Pseudo functions to control numerical dispersion. *SPE Journal,* August, p.269-275, 1975.

Lépine O.J. Uncertainty analysis in predictive reservoir simulation using gradient information. *SPE Journal*, v.4, N 3, p.187-199, 1999.

Luenberger D.G. Linear and non-linear programming. *Addison Wesley press*, 507 p., 1984.

Lygren M., O. Husby O., Osdal B., El Quair Y., Springer M. History matching by using 4D seismic and pressure data on the Norne Field. *Paper C003 presented at the 67th EAGE Annual Conference, Madrid, Spain*, 13-16 June, 2005.

Madatov A.G. Optimization of the multi dimensional basis of the model space at the seismic processing applications. *Soviet Geophys. J.*, v.13, N 1, p.45-56, 1991 (in Russian).

Madatov A.G. The overpressure driven seismic velocity response. Review of standard models and methods for extraction in the context of basin modelling approach to overpressure prediction. *Proceedings of the Murmansk State Techn. University*, v.8, N 1, p.84-119, 2005a.

Madatov A.G., Sereda A.-V.I. A multi-well data inversion purpose-built for calibration of an effective basin model at overpressure prediction. *Proceedings of the Murmansk State Techn. Univ.*, v.8, N 1, p.44-83, 2005b.

Madatov A.G., Sereda A.-V.I. The pre-drill and real time while drilling overpressure prediction based on Effective Basin Model concept. *Proceeding of the Murmansk State Techn. University*, v.9. p.361-388, 2006.

Madatov A.G., Sereda A.-V.I. Upscaling of a basin model required for pore pressure prediction before and during drilling. *Proceedings of the Murmansk State Techn. Univ.*, v.6, N 1, p.119-144, 2003.

Madatov A.G., Williams K.E. An effective basin model for pore pressure prediction. *Marine and Petroleum Geology.* In press. 2010.

Magara K. Compaction and fluid migration. *Elsevier Scientific Publishing Company*, p.319, 1978.

Mahani H., Muggeridge A.H. Improved coarse grid generation using vorticity. *Paper SPE 94319 prepared at the SPE Europec/EAGE Annual Conference held in Madrid, Spain*, June, 2005.

Mallick S. Some practical aspects of prestack waveform inversion using a genetic algorithm: An example from the east Texas Woodbine gas sand. *Geophysics*, v.64, N 2, p.326-336, 1999.

Menke W. Geophysical data analysis. Discrete inverse theory. *Academic press, New York*, 312 p., 1984.

Metropolis N., Rosenbluth A.W., Rosenbluth M., Teller A.H., Teller E. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, v.21, p.1087-1092, 1953.

Montgomery D.C. Design and analysis of experiments. *John Wiley and Sons, New York City*, 2001.

Moulton J.D., Dendy J.E., Hyman J.M. The black box multigrid numerical homogenization algorithm. *J. Comput. Phys.*, v.141, p.1-29, 1998.

Mouret C., Coste J.F., Valois J.P. Optimized inversion of production data for reservoir characterization. *Paper D-44 presented at the EAGE 65th Conference & Exhibition, Stavanger, Norway*, 2-5 June, 2003.

Myers R.H., Montgomery D.C. Response surface methodology: Process and product optimization using designed experiments. *New York City, Wiley*, 381 p., 1995.

Okano H., Pickup G.E., Christie M.A., Subbey S., Sambridge M., Monfared H. Quantification of uncertainty in relative permeability for coarse-scale reservoir simulation. *Paper SPE 94168 presented at the SPE Europec/EAGE Annual Conference held in Madrid, Spain*, 13-16 June, 2005.

Orlyanskaya I.V. Modern approaches to global optimization. 2002 (in Russian). URL: http://zhurnal.ape.relarn.ru/articles/2002/189.pdf .

Pedersen S., Tennebo P., Sonneland L., Carrillat A. Real time update of a reservoir property model in geosteering applications. *Paper D014 presented at the 67th EAGE Annual Conference held in Madrid, Spain*, 13-16 June, 2005.

Pickup G.E., Ringrose P.S., Corbett P.W., Jensen J.L., Sorbie K.S. Geology, geometry and effective flow. *SPE 94018. Petroleum geoscience*, N 1, p.37-42, 1995.

Pohlheim H., Pawletta S., Westphal A. Parallel evolutionary optimization under Matlab on standard computing networks. *Evolutionary Computation and Parallel Processing Workshop*, p.174-176, 1999.

Portella R.C.M., Prais F. Use of automatic history matching and geostatistical simulation to improve production forecast. *SPE Annual Tech. Conference, Denver, SPE 53976*, October, 1996.

Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P. Numerical recipes in C. The art of scientific computing. *Cambridge University Press*, 960 p., 1992.

Qi D.S., Wong P.M., Liu K.Y. An improved global upscaling approach for reservoir simulation. *Petrol. Sci. Technol.*, v.19, p.779-795, 2001.

Rahon D. Gradients method constrained by geological bodies for HM. *SPE Annual Tech. Conference, Denver, SPE–36568*, October, 1996.

Reynolds A.C., He N., Chu L., Oliver D.S. Reparameterization techniques for generating reservoir descriptions conditioned to variograms and well-test pressure data. *SPE Journal*, December, p.413-426, 1996.

Ringrose P.S., Skjetne E., Elfenbein C. Permeability estimation functions based on forward modeling of sedimentary heterogeneity. *SPE 84275 presented at the SPE Annual Techn. Conference and Exhibition, Denver*, CO, 5-8 October, 2003.

Saez A.E., Otero C.J., Rusinek I. The effective homogeneous behaviour of heterogeneous porous media. *Transport in Porous Media*, v4, p.213-238, 1989.

Salhi M.A., Van-Rijen M., Wei L., Dijk H., Alias Z., Upadhyaya A., Lee H. Structured uncertainty assessment for a mature field through the application of experimental design and response surface methods. *Paper SPE 93529 presented at 14th SPE Middle East Oil-Gas Show Conference, held in Bahrain International Exhibition Center*, 12-15 March, 2005.

Sambridge M. Geophysical inversion with a neighbourhood algorithm-I: Searching a parameter space. *Geophysical J. International*, v.138, p.479, 1999.

Sarma P., Durlofsky L.J., Aziz K. Efficient closed-loop production optimization under uncertainty. *Paper SPE 94241 presented at the SPE Europec/EAGE Annual Conference held in Madrid, Spain*, 13-16 June, 2005.

Sogerlind L.J. Applied finite element analysis. *Wiley, New York*, 422 p., 1976.

Staples R., Stevens T., Leoutre E., Jolley S., Marshall J. A 4D seismic history matching – the reality. *Paper D009 presented at the EAGE Annual Conference held in Madrid, Spain*, 13-16 June, 2005.

Stephen K.D., Soldo J., MacBeth C., Christie M. Multiple model seismic and production history matching: A case study. *Paper SPE 94173 presented at the SPE Europec/EAGE Annual Conference held in Madrid, Spain*, 13-16 June, 2005.

Sylvester I.F., Cook R., Swift R., Pritchard T., McKeever J. Integrated reservoir modelling enhances the understanding of reservoir performance of the Dolphin Gas Field, Trinidad and Tobago. *Paper SPE 94343 presented at the SPE Europec/EAGE Annual Conference held in Madrid, Spain*, 13-16 June, 2005.

Syswerda G. Uniform crossover in genetic algorithms. [ICGA3], p.2-9, 1989, URL: http://www.geatbx.com/docu/algindex-10.html

Talukdar S., Brusdal L.H. Reservoir management of the Njord Field. *Paper SPE 92419 presented at the SPE Europec/EAGE Annual Conference held in Madrid, Spain*, 13-16 June, 2005.

Tarantola A. Inverse problem theory: Methods for data fitting and model parameter estimation. *Elsevier, Netherlands*, p.386, 1987.

Thomas P., Le Ravale-Dupin M., Roggero F. History matching reservoir models simulated from the Pluri-Gaussian method. *Paper SPE 94168 presented at the SPE Europec/EAGE Annual Conference held in Madrid, Spain*, 13-16 June, 2005.

Tikhonov A.N. About stability of inverse problem. *Proceeding of USSR Academy of Science*, v.39, N 5, p.195-198, 1943 (in Russian).

Tikhonov A.N., Arsenin V.Ya. The methods of solution of ill-posed problems. *M., Science*, 310 p., 1979 (in Russian).

Tipping D.E., Deschenya M., Deimbacher F., Kovyazin D. Increasing of production efficiency for a HC field in Siberia based on quantitative production risk estimations, SPE 101808. *Paper presented on SPE Conference held in Moscow, Russia*, 3-6 October, 2006 (in Russian).

Verdière S., Guérillot D. Dual mesh method in heterogeneous porous media. *Paper presented at 5th European Conference on the Mathematics of Oil Recovery, Leoben*, September, 1996.

Vidal S., Longuemare P., Huguet F., Mechler P. Reservoir parameters quantification from seismic monitoring integrating geomechanics. *Oil & Gas Science and Technology Rev. IFP*, v.57, N 5, p.555-568, 2002.

Voigt H.M., Anheyer T. Modal mutations in evolutionary algorithms. *ICEC94*, v.1, p.88-92, 1994.

White C.D., Willis B.J., Narayanan K., Dutton Sh.P. Identifying and estimating significant geologic parameters with experimental design. *SPE Journal*, September, p.311-324, 2001.

Whitley D. Evolutionary Computation. *Journal, Cambridge, Massachusetts, MIT Press*, p.726-733, 1999.

Williams K.E., Madatov A.G. An effective basin model for pore pressure evaluation. *Paper presented at American Association of Petroleum Geologists Hedberg Conference, The Netherlands*, May 6-9, 2007.

Wong P.M., Boerner S.T. Quantifying uncertainty in statistical regression techniques for reservoir mapping. *Paper C042 presented at the EAGE 66th Conference & Exhibition, Paris, France*, 7-10 June, 2004.

Zeltov Y.P. Reservoir Engineering. *M., Nedra*, 365 p., 1998 (in Russian).

Zhang F., Skjervheim J.A., Reyngolds A.C. An automatic history matching example. *Paper D-45 presented at the EAGE 65th Conference & Exhibition held in Stavanger, Norway*, 2-5 June, 2003.

Zienkiewicz O.C., Taylor R.L. The finite element method for fluid dynamics. *Elsevier*, 689 p., 2000.

APPENDIX 1: Formal definition of parametrical upscaling and calibration problems in context of production data inversion

Here we formally define parametrical upscaling and calibration of an effective reservoir model in order to be consistent with general theory (*Tikhonov*, 1943; *Tikhonov, Arsenin*, 1979; *Menke*, 1984; *Tarantola*, 1987). The specifics of the reservoir engineering problem will be preserved to keep link with practical solutions of a history matching problem.

Let us first introduce a <u>Production History Data set</u> as a set of physical measurements observable during production history time interval [0,T] on sparse well system $W(x,y,z)$ associated with some real reservoir geo-fluid system. This set could be digitally represented by production log data, well test descriptions, multi-rate tests, well deliverability etc., which all form some multi-well production history record. Let also well control data such as borehole pressure, injection rates, switching on and off time interval etc. be available. We will associate it with the given production data file in the form and amount suitable for the standard reservoir simulation (see Appendix 3 to get example). Then the production data for each *j*-th of the HC phases observed at *k*-th history moment within the *i*-th production well belonging to the multi-well system $W(x,y,z)$ can be treated as a f_{ijk} component of the real data vector **f** defined in Euclidean <u>Real Data vector space</u> **F** , which accommodates all possible data vectors.

Generally, a synthetic data prototype vector produced by the standard reservoir simulator requires some scale and/or grid transformations to be comparable with the real data vector **f**. For example, the borehole fluid pressure measured in [Pa] can be used for data fitting only after recalculation of synthetic pressure from a reservoir scale grid on a near well scale grid (*Calhoun Jr.*, 1982), or a cumulative oil production rate [bopd] can be comparable with its synthetic prototype only after converting into the standard surface PVT conditions (*Denesh*, 1998).

For generality reasons we have to introduce also a <u>Data Comparison vector space</u> **P** as a universal Euclidean space, which accommodates all potentially possible synthetic data prototype vectors $\underline{p}^{\wedge} \in$ **P**. Then we have to introduce also a data transforming operator ensuring link between **F** and **P** vector spaces.

<u>A Data Transforming operator</u> $T[\underline{f}] = \underline{p} \in$ **P** is defined on the Real Data space **F** for continuous mapping of any real dynamic data set given in form $\underline{f} \in$ **F** into the Data Comparison space **P**, where comparison and fitting with its synthetic prototype $\underline{p}^{\wedge}$ are possible.

Generally the data transformation operator $T[\underline{f}]$ can be understood as pre-processing of a raw production data set before its use for model fitting. Depending on the production data type, $T[\underline{f}]$ operator can be based on a rather simple one-by-one scale or grid conversion routine and/or on much more complicated routines up to data inversion. For example seismically driven production indicators are not directly convertible into the Data Comparison space **P** since the relevant data transform operator $T[\underline{f}]$ for them generally belongs to the seismic data inverting class (*Bornard et al.*, 2005; *Gosselin et al.*, 2003), which is far more complicated than the considered history matching problem.

The raw production data $\underline{f} \in$ **F** can be characterized with some inherent finite data error value $\varepsilon_{\underline{f}}$. Naturally the estimation of the final real data accuracy $\varepsilon_{\underline{p}}$ on the Data Comparison space **P** should be considered as a composition of raw data uncertainty with some processing errors associated with the raw data transformation. As soon as the data transformation operator $T[\underline{f}]$ performs continuous mapping $\underline{f} \Rightarrow \underline{p}$, the condition $\varepsilon_{\underline{p}} \geq \varepsilon_{\underline{f}}$ is generally guaranteed (see fig. A1.1). Further in the text we will associate real production data with the vector $\underline{p}=T[\underline{f}]$ and a definitive real data accuracy $\varepsilon_{\underline{p}}$ – with the ε-locality attached to its most likely position on the Data Comparison space **P** in the form of closed simply-connected subset $\{\varepsilon\underline{p}\} \subset$P (see Fig. A1.1). The notifications $\underline{p} \in$ **P** and relevant feature of the end-vector $\underline{p}=T[\underline{f}]$ in fig. A1.1 must be understood in the same sense.

Fig. A1.1 The operator scheme of the real data transformation

An effective model and a data simulation process associated with the above introduced reservoir geo-fluid system should now be described formally.

In agreement with the effective model concept we have to distinguish between the fixed part of the reservoir model in form of a framework and its tenable part in form of Control Model Parameter set, by adjustment of which the data inversion problem (the model fitting problem) is going to be resolved (see also *Madatov, Sereda*, 2005). Note that such consideration is general enough for accommodating all types of effective RM architectures considered in part 1 of this paper.

Let the framework part $F(\Omega)$ be fixed, its textural part, which controls the reservoir transport properties, is also chosen and the rest of the petrophysical input necessary for data simulation is initialized. So, an effective RM is prepared on the fine grid $\Omega(x,y,z)$ and ready for fitting. Note that the grid $\Omega(x,y,z)$ must accommodate the multi-well production system $W(x,y,z)$ in snapped form.

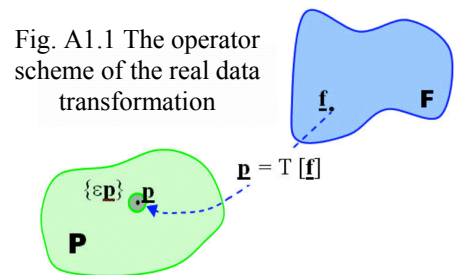Remember that an effective RM prepared for history matching has the following prerequisites:

1. The geology setting and seismic data coverage for the given reservoir fluid-system are detailed enough to construct and to fix the structure framework $F(\Omega)$, which includes both continuous and discontinuous subsurface features.
2. The prior litho-stratigraphy knowledge is detailed enough to reconstruct a facies architecture of the reservoir and hence to establish a textural model for it.
3. Both framework and textural parts of an effective RM are represented in form of unique 3D lithology specification defined on the grid $\Omega\,(x,y,z)$ with details sufficient for simulating of the data prototype $\mathbf{p}\in\mathbf{P}$.
4. The available petrophysical data allows initialization of 3D porosity distribution with acceptable for simulation accuracy on the specified above grid $\Omega\,(x,y,z)$.
5. Generalized functions required for specification of the absolute and relative permeability in agreement with (3-4) are chosen together with default values for the relevant control model parameters (CMP vector components) and their variation ranges

   Note, that these assumptions can normally be guaranteed at the relevant stage of a history matching.

The total amount of components of CMP vector released for adjustment is given by $N = \sum_{k=1}^{K} n_k + \sum_{i=1}^{I} m_m$,

where: *K,I* are the total amounts of litho-facies units and barriers constituting the chosen RM texture respectively; n_k, m_i are the numbers of CMP components per *k-th* flow unit, *i-th* flow barriers respectively. Note, that in general case $n_k \neq n_{k-1}$ and $m_i \neq m_{i-1}$ (see part 1 of the article for more details).

Since the CMP variation range is assumed be predefined for each of the CMP components it is convenient to get it normalized by simple transformation: $\underline{\mathbf{x}} = (\boldsymbol{\chi} - \boldsymbol{\chi}_{min}) / (\boldsymbol{\chi}_{max} - \boldsymbol{\chi}_{min})$, where $\boldsymbol{\chi}$ and $\underline{\mathbf{x}}$ denote CMP vector before and after normalization respectively.

As soon as standard sorting and normalisation procedures are defined, any of physically possible CMP combinations is presentable in the form of dimensionless vector $\underline{\mathbf{x}}$ as following:

$$\underline{\mathbf{x}} = \{(X_1, X_2, \ldots, X_{n1}), (X_1, X_2, \ldots, X_{n2}), \ldots, (X_1, X_2, \ldots, X_{nk}), \ldots, (X_1, X_2, \ldots, X_{m1}), (X_1, X_2, \ldots, X_{m2}), \ldots, (X_1, X_2, \ldots, X_{mi})\}. \quad \text{(A1.1)}$$

Evidently, all possible $\underline{\mathbf{x}}$ components' combinations can be accommodated inside some Euclidean Model Parameter space $\mathbf{X}$ having dimension *N*.

Now, the <u>Control Model Parameter vector</u> $\underline{\mathbf{x}}$ to be fitted against the real data vector $\mathbf{p}\in\mathbf{P}$ is introduced in connection with model specification rules for an effective RM as dimensionless vector $\underline{\mathbf{x}}$ (A1.1) defined on the multidimensional Model Parameter space $\mathbf{X}$.

The range of variations of the vector $\underline{\mathbf{x}}\in\mathbf{X}$ for each of Litho-facies types included in a reservoir texture can normally be predefined in agreement with the relevant forward model assumptions and restrictions in form of *a priori* compact subset $\mathbf{X}^A \subset \mathbf{X}$. We will associate $\mathbf{X}^A$ also with the multidimensional "search window" in context of the CMP tuning.

Regardless of the numerical approach to the forward problem solution, which delivers the synthetic data prototype $\mathbf{p}^{\wedge}\in\mathbf{P}$ it is possible to introduce a reservoir simulator as a forward modelling operator defined on space $\mathbf{X}$.

The <u>Forward Modelling Operator</u> $M[\underline{\mathbf{x}}]$ is defined on the Model Parameter space $\mathbf{X}$ in connection with a forward model of phenomenon observable on the real data space $\mathbf{F}$ and transformable into the Data Comparison space $\mathbf{P}$ as a combination of unique routines for an effective RM specification, methods of its representation in cellular form and a relevant numerical solver of a forward problem.

As the three above mentioned procedures are well posed and the framework $F(\Omega)$ is fixed, $M[\underline{\mathbf{x}}]$ provides unique, stable and certain mapping of the arbitrary vector $\underline{\mathbf{x}}$ from the Model Parameter space $\mathbf{X}$ to the Data Comparison space $\mathbf{P}$ where the relevant synthetic data vector $\mathbf{p}^{\wedge}\in\mathbf{P}$ can be compared with its production data prototype $\mathbf{p}\in\mathbf{P}$.

It is also convenient to use dimensionless metrics on Data Comparison space $\mathbf{P}$ for measuring distance (misfit) between two arbitrary vectors $\mathbf{p}_0, \mathbf{p}_1$ given on it in the form:

$$d(\mathbf{p}_0, \mathbf{p}_1) = ||\mathbf{p}_0 - \mathbf{p}_1|| / ||\mathbf{p}_0||, \; \forall \{\mathbf{p}_0, \mathbf{p}_1\} \in \mathbf{P} , \quad \text{(A1.2)}$$

where $||\mathbf{p}||$ denotes a norm of the vector $\mathbf{p}$, which could be represented in geometric or harmonic forms.

Fig. A1.2. The operator schemes of the resolving of a conditionally correct inverse problem

Now we can define a data fitting problem associated with history matching in terms of a conditionally correct inverse problem (*Tikhonov, Arsenin*, 1979) (see fig. A1.2).
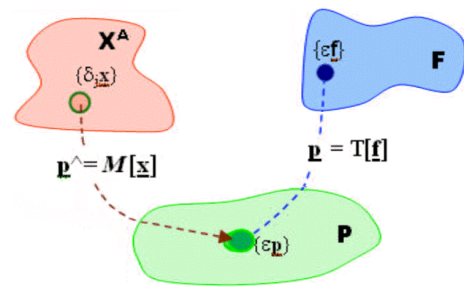
Let the *N*-dimensional Control Model Parameter space **X** be established and normalized.

Let the relevant forward modelling operator *M*[**x**] be defined on it.

Let also the simply connected sub-space $\mathbf{X}^A \subset \mathbf{X}$ (search sub-space) of all possible **x** vector positions be *a priori* defined in agreement together with the variation ranges for each of the **x** components.

Let the production data vector **p** be composed and transformed into the Data Comparison space **P** in form of simply-connected subset {ε**p**}⊂**P** defined around its most likely position **p**∈**P**.

The specific inverse problem for the data vector **p** with regard to the model sub-space $\mathbf{X}^A \subset \mathbf{X}$ is then reduced to constructing of the operator M^{-1}[**p**] defined on the Data Comparison space **P**, which is formally inverse to the forward operator *M*[**x**] and which is targeted for reverse mapping of any given data vector **p**∈**P** back to the origin Model Parameter sub-space $\mathbf{X}^A \subset \mathbf{X}$ (*Menke*, 1984). Since the operator *M*[**x**] is essentially non-linear we will define inverse operator M^{-1}[**p**] in class of non-unique data fitting operators. This is the formal reason why generally several possibilities exist in Control Model Parameter sub-space $\mathbf{X}^A$ to find an acceptable model parameter vector $\underline{\mathbf{x}} \in \mathbf{X}^A$ which delivers a fit between the data vector **p**∈**P** and its synthetic prototype $\mathbf{p}^\wedge = M[\underline{\mathbf{x}}] \in \mathbf{P}$ with the predefined fitting accuracy *ε*.

Generally, an inverse operator for the given data fitting problem can be defined within the class of conditional optimization procedures (*Menke*, 1984) on the Data Comparison space **P** as soon as a data fitting criterion is defined in terms of conditioned misfit level.

The standard misfit measure or the goal function in deterministic approach to an inverse problem solution is given by (*Tikhonov, Arsenin*, 1979):

$$R[(\underline{\mathbf{x}}, \mathbf{p})] = (\mathbf{p} - M[\underline{\mathbf{x}}])^T C_p^{-1}(\mathbf{p} - M[\underline{\mathbf{x}}]) = (\mathbf{p} - \mathbf{p}^\wedge)^T C_p^{-1}(\mathbf{p} - \mathbf{p}^\wedge), \tag{A1.3}$$

where **p**, $\mathbf{p}^\wedge \in \mathbf{P}$ in context of considered problem are production data and its synthetic prototype vectors respectively; C_p – the covariation matrix.

The misfit definition could be extended by adding some model constraints in soft or hard forms (*Tarantola*, 1987). Therefore the advanced misfit definition could be expressed in the form:

$$R[(\underline{\mathbf{x}}, \mathbf{p})] = (\mathbf{p} - M[\underline{\mathbf{x}}])^T C_p^{-1}(\mathbf{p} - M[\underline{\mathbf{x}}]) + (\underline{\mathbf{x}} - \underline{\mathbf{x}}^{\mathrm{ML}})^T C_x^{-1}(\underline{\mathbf{x}} - \underline{\mathbf{x}}^{\mathrm{ML}}), \tag{A1.3*}$$

for soft constraints, or

$$R[(\underline{\mathbf{x}}, \mathbf{p})] = (\mathbf{p} - M[\underline{\mathbf{x}}])^T C_p^{-1}(\mathbf{p} - M[\underline{\mathbf{x}}]) \text{ at } \underline{\mathbf{x}}^{\mathrm{MIN}} \le \underline{\mathbf{x}} \le \underline{\mathbf{x}}^{\mathrm{MAX}} \tag{A1.3**}$$

for hard constraints.

Thus, the inverse operator M^{-1}[**p**] at the deterministic approach to a model fitting problem can be finally defined based on the above introduced definitions as follows:

$$M^{-1}[\underline{\mathbf{p}}] = \underline{\mathbf{x}}_j; \; argmin \; [R(\mathbf{p}^\wedge, \mathbf{p})] \; \forall \underline{\mathbf{x}} \in \mathbf{X}^A \subset \mathbf{X}; \; \forall \; \mathbf{p}, \mathbf{p}^\wedge \in \mathbf{P}, \tag{A1.4}$$

Still, the inverse problem (A1.4) remains ill-posed in Hadamard sense (*Hadamard*, 1932) unless the minimization operator *argmin[ ]* is replaced in (A1.4) on conditional operator as follows:

$$M^{-1}[\underline{\mathbf{p}}] = \underline{\mathbf{x}}: R(\mathbf{p}^\wedge, \mathbf{p}^*) \le \varepsilon, \; \forall \underline{\mathbf{x}} \in \{\underline{\mathbf{x}}\} \subset \mathbf{X}^A \subset \mathbf{X}; \; \forall \; \mathbf{p}, \mathbf{p}^\wedge \in \mathbf{P}, \tag{A1.4*}$$

where *ε* is the acceptable misfit measure defined in advance (see fig. A1.1, A1.2).

The solution of the specific IP conditioned as it does (A1.4*) can be represented in form of *ε*-equivalent sub-set {**x**}⊂$\mathbf{X}^A \subset \mathbf{X}$ become stable and unique (*Tikhonov, Arsenin*, 1979).

The stochastic approaches to the specific IP solution consist in its Bayesian formulation. Here the solution is formulated in terms of a Probability Distribution Function (PDF), which is defined in connection with an inherently uncertain position of the vector **x** within the search sub-space $\mathbf{X}^A \subset \mathbf{X}$. Therefore *a priori* – P(*x*) and *a posteriori* – P(**p**|*x*) PDFs are introduced into consideration instead of the unknown **x** vector position in the Model Parameter space. As a matter of fact P(*x*) and P(**p**|*x*) represent a level of our knowledge about the reservoir fluid system in a stochastic model sense before and after, respectively, accounting for the production history data.

The standard goal function at the Bayesian approach to the inverse problem solution is defined in agreement with definition of a likelihood function (*Sambridge*, 1999; *Christie et al.*, 2002):

$$P(\underline{x} \mid \mathbf{p}) = [P(\mathbf{p}|\underline{x}) P(\underline{x})] / [\sum_{i=1,N} P(\mathbf{p}|x_i) P(x_i)]. \tag{A1.5}$$

This function is usable for updating *a priori* probability as a result of getting an observation P(**p**|*x*). The denominator in the right-hand side of (A1.5) is referred as normalization since the sum of posterior probability P(**p**|*x*) on all possible realizations must be equal to one.

The relevant optimization problem is associated here with maximization of goal function (A1.5) on Model Parameter and Data Comparison spaces. Thus unconditional data inverting operator in Bayesian formulation of history matching problem can be expressed in form:

$$W[\underline{p}] = \underline{x}_j: \text{argmax} [P(\underline{x}|\underline{p})] \ \forall \underline{x} \in \mathbf{X}; \ \forall \ \underline{p} \in \mathbf{P}, \quad \text{(A1.6)}$$

where $P(\underline{x}|p)$ – the likelihood function that defines probability of improving of our *posteriori* knowledge about reservoir model after getting observation $P(\underline{p}|\underline{x})$ (*Okano et al.*, 2005).

By analogy with conditionally correct formulation of deterministic inverse problem one should replace maximization operator *argmax*[ ] in (A1.6) with conditional operator. Then one will get:

$$W[\underline{p}] = \underline{x}: [P(\underline{x}|\underline{p})] \geq \delta, \ \forall \underline{x} \in \mathbf{X}; \ \forall \ \underline{p} \in \mathbf{P}, \quad \text{(A1.6*)}$$

where $\delta \leq 1$ is the predefined likelihood upper limit.

The inverse operators implemented at the deterministic and the stochastic formulations of a data inversion problem differ by strategy of optimization methods. Normally, the traditional local optimization approaches are implemented in the deterministic operators (A1.4-4*) (*Lépine*, 1999) (see also Appendix 2). The

Fig. A1.3 The scheme of consequent shrinking of the model parameter space $\mathbf{X}_i$ during parametrical upscaling and relevant creeping of a single synthetic data vector image around a conditionally exact solution $\underline{p}^\wedge = M[\underline{x}_0]$

maximization operator $W[\underline{p}]$ at Bayesian formulation of an inverse problem deals with consequent sampling (re-sampling) of a reservoir model and with repeated simulation runs. Here the Monte Carlo approach or the Markov Chain approach or its combination (MCMC) are applicable (*Gilks et al.*, 1996) as well as Neighbourhood Approximation (*Sambridge*, 1999).

Nowadays the global optimization techniques are often implemented in both strategies in inverse problem solutions. Namely the multi-start Evolution Algorithms (*Cavicchio*, 1970; *Goldberg*, 1989; *Bäck*, 1996) and Simulated Annealing algorithms (*Ingber*, 1993; *Kirkpatrick et al.*, 1983). Nevertheless it does not exclude combining of step-wise descent techniques belonged to the local optimization methods with the multi-start global optimization techniques at the deterministic conditionally correct formulation of an inverse problem (A1.4*).

We refer readers to Appendix 2 for more details about different approaches to the solution of optimization problems. Below the parametrical upscaling and the RM calibration sub-problems are considered in more detail based on the above introduced terminology and formalism.

A1.1. The parametrical upscaling of effective reservoir models

Homogenisation or upscaling of a reservoir model generally can be treated as a procedure by means of which the field scale flow and transport properties can be inferred from the effectively compressed petrophysical information available in a fine-scale RM and reproduced with accuracy constrained by some reference solution (*Moulton et al.*, 1998). In the case of parametrical upscaling this constraint is associated with finite accuracy of reference production data expressed in the vector norm measure ε defined on data comparison space $\mathbf{P}$ (see fig. A1.1-2).

Following to the above introduced notifications and assumptions let us suppose that the framework of some effective reservoir model is fixed on the fine scale. This means that the total amount of the litho-facies units and number of the core model parameters per unit correspond to a maximal available knowledge about the given reservoir geo-fluid system at the given stage of a production history. Despite unavoidably missing some of the details we will treat such a fine scale model as an excessively complete RM description. We also will recognize the relevant Model Parameter space $\mathbf{X}_0$ as a Universal one, which has maximal possible dimension at the given RM specification rules (see part 1 of the given article for more details). Neglecting some numerical errors we will associate $\mathbf{X}_0$ with the conditionally exact mapping $\underline{p}^\wedge = M[\underline{x}_0]$. In other words the relevant synthetic prototype $\underline{p}^\wedge$ of some real data vector produced is assumed to be free from modelling errors (see fig. A1.3). It is intuitively clear that such a speculative case can only be considered hypothetically. Practically however, the grid coarsening as well as skipping some of the secondary important physical mechanisms during simulation can be considered as decreasing the dimension and variability ranges for relevant workable Model Parameter space. Obviously it now can be treated as a sub-set of a universal one, i.e. $\mathbf{X}_j \subset \mathbf{X}_0$. From this point of view, for example, the class of reservoir models associated with the two-phase fluid dynamics is accommodated as a sub-set inside the more universal class of coupled geomechanics – fluid dynamics reservoir models (*Chin et al.*, 2002; *Vidal et al.*, 2002). Amount and quality of data available for calibration should dictate real needs in usage of either particular class of reservoir models.
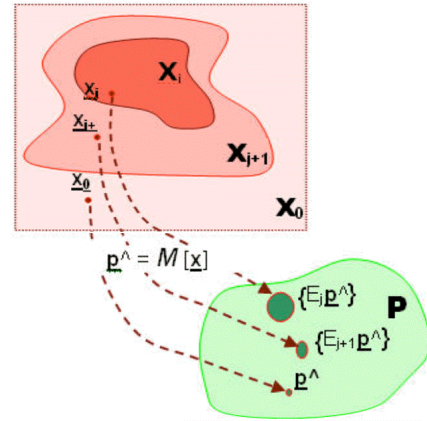
Formally speaking, any model simplification inevitably leads to introduction of some inherent modelling uncertainty (apparent modelling error) due to omission of some secondarily important (non-sensitive to the result) components in Model Parameter space $\mathbf{X}_0$. This apparent modelling error can formally be associated with a fuzzy forward problem solution $\{E\underline{\mathbf{p}}_0{}^\wedge\} \subset \mathbf{P}$ (fig. A1.3) interpretable as a group of all possible "exact" solutions, which were achieved on Data Comparison space $\mathbf{P}$ by using exact mapping $M[\underline{\mathbf{x}}_0]$ if all of omitted $\underline{\mathbf{x}}$ components will be recovered individually. Taken into account the continuity of the forward operator $M[\underline{\mathbf{x}}]$ and the low sensitivity of mapping results to the omitted $\underline{\mathbf{x}}_0$ components we could formally interpret the cumulative effect of simplification of the initially universal Model Parameter space $\mathbf{X}_0$ as a substituting of exact mapping result $\underline{\mathbf{p}}^\wedge$ on fuzzy mapping result $\{E\underline{\mathbf{p}}_0{}^\wedge\}$. Here $\{E\underline{\mathbf{p}}_0{}^\wedge\}$ denotes a closed simply-connected subset (*E*-locality of the exact solution $\underline{\mathbf{p}}_0$) caused by coarsening the Model Parameter space $\mathbf{X}_0$ and comparable with the above-introduced ε-locality of the real data image $\{\varepsilon\underline{\mathbf{p}}\}$ on the Data Comparison space $\mathbf{P}$. Clearly, the more components of the exact vector $\underline{\mathbf{x}}_0$ are omitted the higher extent of the apparent modelling error and the bigger size of *E*-locality around the exact solution $\underline{\mathbf{p}}_0$. Formally this effect of a model simplification can be expressed as follows:

$$\mathbf{X}_0 \subset \mathbf{X}_1 \subset \mathbf{X}_2 \subset \mathbf{X}_3 \Rightarrow \underline{\mathbf{p}}_0{}^\wedge \in \{E_1\underline{\mathbf{p}}_0{}^\wedge\} \subset \{E_2\underline{\mathbf{p}}_0{}^\wedge\} \subset \{E_3\underline{\mathbf{p}}_0{}^\wedge\}, \forall \underline{\mathbf{x}}_0 \in \mathbf{X}_0: M[\underline{\mathbf{x}}_0] = \underline{\mathbf{p}}_0{}^\wedge. \qquad (A1.7)$$

Now simplifying the reference universal Model Parameter space $\mathbf{X}_0$ in parametrical upscaling can be associated with the process of consequent omission of some of its coordinates, namely those which correspond to regularly low sensitive $\underline{\mathbf{x}}$ components. According to (*Madatov*, 1991; *Madatov, Sereda* 2005) an optimally simple Model Parameter space $\mathbf{X}_j \subset \mathbf{X}_0$ can be achieved during such consequent process as soon as the diameter of subset $\{E\underline{\mathbf{p}}_0{}^\wedge\}$ will be equal to the diameter of subset $\{\varepsilon\underline{\mathbf{p}}\}$, or just if $E_j = \varepsilon_p$. Thus, the algorithm of parametrical upscaling can be formally expressed in the following operator form:

$$J = \text{MIN}: M[\underline{\mathbf{x}}_j] = \{E_j\underline{\mathbf{p}}^\wedge\}, \exists \underline{\mathbf{x}}_j \in \mathbf{X}_j \text{ and } \{E_j\underline{\mathbf{p}}^\wedge\} \subseteq \{\varepsilon\underline{\mathbf{p}}\} \Rightarrow (E_j \le \varepsilon_\mathbf{p}), \underline{\mathbf{p}}^* \ \forall \in \mathbf{P}. \qquad (A1.8)$$

Essentially the expression (A1.8) implies comparison of an IP solution accuracy with an apparent modelling error. Practically there are no needs for exploration all of local IP solutions, but rather in checking whether or not at least one local solution is achievable at the given dimension *j* of model sub-space $\mathbf{X}_j$ (*Madatov, Sereda*, 2003).

A1.2. Calibration of an effective reservoir model

RM calibration is treated here as a field scale model-to-data fitting process, in the result of which such a RM specification should be obtained that yields sufficiently accurate simulations of the available production history data. Therefore calibrated reservoir model can be used for forecasting of future production profiles and for assessment of drilling risks at different continuation scenarios. Regardless of the method of calibration the model-to-data fitting process should be recognised as data inversion. Therefore all aspects of this ill-posed math problem – non-stability, high CPU time consumption and a non-uniqueness (*Hadamard*, 1932) shall be taken into account. In the context of reservoir performance prediction we will pay especial attention to the last aspect – a non-uniqueness of the IP solution.

The above-introduced conditionally correct formulation of the specific inverse problem ensures its stable and regular solution (*Tikhonov, Arsenin*, 1979) as soon as the data quality stop criterion is finite and the forward modelling operator is continuous. The CPU time consumption problems could be partially resolved by means of the parametrical upscaling. Still, even a model constrained IP solution conditioned according to (A1.3-4) remains local and not unique. This is because of highly non-linear behaviour of the relevant forward modelling operator $M[\underline{\mathbf{x}}_j]$ within the model constrained sub-space $\mathbf{X}^A \subset \mathbf{X}$.

Practically it means that the L_2 distance inside the vector space $\mathbf{P}$ between two forward problem solutions $\underline{\mathbf{p}}_i = M[\underline{\mathbf{x}}_j]$ and $\underline{\mathbf{p}}_j = M[\underline{\mathbf{x}}_i]$ can belong to a common ε-equal locality $\{\varepsilon\underline{\mathbf{p}}\} \subset P$, whereas $\underline{\mathbf{x}}_j$ and $\underline{\mathbf{x}}_i$ do not belong to any simply connected δ-locality defined on $\mathbf{X}^A$ (fig. A1.4). This situation can be recognised as a *Global non-uniqueness problem*.

In accordance with general definitions (A1.3-4) let us define this problem in more formal way as following:

$$\{\underline{\mathbf{x}}\}^n \not\subset \{\delta\underline{\mathbf{x}}\} \subset X^A : \|M[\underline{\mathbf{x}}_j] - \underline{\mathbf{p}}^\varepsilon\| \le \varepsilon, \forall \ \underline{\mathbf{x}}_j \in \{\underline{\mathbf{x}}^N\}. \quad (A1.9)$$

Here $\{\underline{\mathbf{x}}\}^n$ presents the set of *n* local solutions achievable for the inverse problem (A1.3-4).
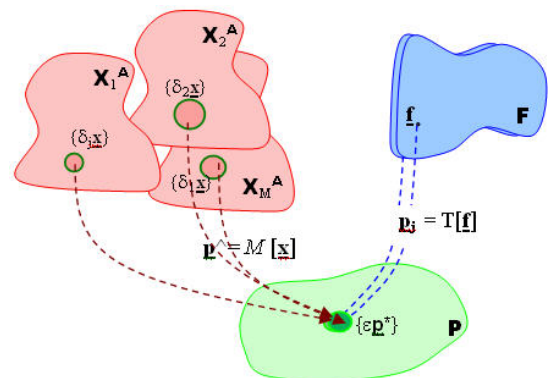

Fig. A1.4 The operator scheme of the data fitting problem solution during calibration of an effective reservoir model

Let us distinguish between *Global non-uniqueness* and *Local non-uniqueness* of the desired IP solution. The last can be defined on the sub-space of *a priori* possible solutions $\mathbf{X}^A$ as following:

$$\{\underline{\mathbf{x}}\}^m \subset \{\delta\underline{\mathbf{x}}\} \subset \mathbf{X}^A : \|M[\underline{\mathbf{x}}_j] - \underline{\mathbf{p}}^\varepsilon\| \leq \varepsilon, \ \forall \ \underline{\mathbf{x}}_j \in \{\underline{\mathbf{x}}\}^M. \tag{A1.10}$$

Now $\{\underline{\mathbf{x}}\}^m$ is a compact simply-connected subset, which accommodates *m* local solutions achievable for the inverse problem (A1.3-4).

Clearly the local *non-uniqueness* problem for the Model Parameter space $\mathbf{X}$ can be resolvable in the same way as for the Data Comparison space $\mathbf{P}$, i.e. by introducing of the corresponding δ-locality (compact $\{\delta\underline{\mathbf{x}}\}$) of some most likely inverse problem solution $\underline{\mathbf{x}}_\delta \in \mathbf{X}^A$. Indeed, any pair of solutions (A1.8) will get non-detectable difference between each other and consequently will belong to the δ-equivalent locality within the model space $\mathbf{X}$.

The continuity of a locally unique solution (A1.8) and the simple-connectivity of the subset $\{\underline{\mathbf{x}}\}^m \subset \{\delta\underline{\mathbf{x}}\}$ follows from locally monotonous behaviour of the forward modelling operator $M[\underline{\mathbf{x}}]$ in the reservoir engineering. Formally it can be expressed in form:

$$0<\delta_1<\delta_2<\delta_3: \ 0< \varepsilon_1< \varepsilon_2< \varepsilon_3 \ldots \Rightarrow .. \varnothing \neq \{\delta_1\underline{\mathbf{x}}\} \subset \{\delta_2\underline{\mathbf{x}}\} \subset \{\delta_3\underline{\mathbf{x}}\}. \tag{A1.11}$$

Conversely, the sub-set $\{\underline{\mathbf{x}}\}^n$ of globally non-unique solutions (A1.7) is not simply-connected. It can be presented as following:

$$\{\underline{\mathbf{x}}\}^n = [\{\delta_1\underline{\mathbf{x}}\} \cup \{\delta_2\underline{\mathbf{x}}\} \cup \{\delta_3\underline{\mathbf{x}}\} \ldots \cup \{\delta_n\underline{\mathbf{x}}\}], \tag{A1.12}$$

where $\{\delta_k\underline{\mathbf{x}}\} \neq \varnothing, \forall \ k =1,2,\ldots,n;$ and $[\{\delta_i\underline{\mathbf{x}}\} \cap \{\delta_j\underline{\mathbf{x}}\}] = \varnothing, \ \forall \ i \neq j: i,j \in (1,2,\ldots n).$

Our many years experience in connection to an Effective Basin Model calibration (*Madatov, Sereda*, 2000; 2005; *Madatov, Williams*, 2010) reveals that condition (A1.12) is the most common case in real practice. By analogy we could expect that the same situation pertains at the reservoir modelling applications.

Thus, calibration of an effective reservoir model implies exploration of the search sub-space $\mathbf{X}^A \subset \mathbf{X}$, detecting on it the set of the ε-equivalent local IP solutions and recognition among them of a sub-set of globally non-unique solutions, which satisfy conditions (A1.12).

The *posteriori* separation of locally and globally non-unique solutions can be based on sensitivity analysis of solution vectors (*Madatov, Sereda*, 2000). Indeed, if $\|\underline{\mathbf{S}}(\underline{\mathbf{x}}_j) - \underline{\mathbf{S}}(\underline{\mathbf{x}}_i)\| \leq \varepsilon$ at $\|\underline{\mathbf{x}}_j - \underline{\mathbf{x}}_i\| > \delta$, then the Freshet derivatives $\partial M[\underline{\mathbf{x}}]/\partial x_k$ will approach to zero for any of components of the vector $\underline{\mathbf{x}} \in \{\delta\underline{\mathbf{x}}_j\}$ or $\underline{\mathbf{x}} \in \{\delta\underline{\mathbf{x}}_i\}$ and condition (A1.11) will be satisfied. So, the checked vectors $\underline{\mathbf{x}}_j$ and $\underline{\mathbf{x}}_i$ will be mapped to the same ε-locality on the Data Comparison space $\mathbf{P}$ despite that they do not belong to the δ-locality on the Model Parameter space $\mathbf{X}$. In other words they can be treated as globally unique and *vice versa*. Here ε and δ are infinitesimally small values defined on spaces $\mathbf{P}$ and $\mathbf{X}$ respectively and $\underline{\mathbf{S}}(\underline{\mathbf{x}})$ is the sensitivity vector defined on space $\mathbf{X}$ as following (see part 1 of this paper for more details):

$$\underline{S}(\underline{\mathbf{x}}) = \frac{M[\underline{\mathbf{x}} + \partial x_k] - M[\underline{\mathbf{x}}]}{M[\underline{\mathbf{x}}]} \tag{A1.13}$$

It is also possible to ensure globally non-unique solutions. It can be based on splitting the search sub-space $\mathbf{X}^A$ on set of the compact sub-spaces $\mathbf{X}_1^A$, $\mathbf{X}_2^A$,... $\mathbf{X}_M^A$ (see fig. A1.4) predefined on conditions $\|\underline{\mathbf{S}}(\underline{\mathbf{x}}_j) - \underline{\mathbf{S}}(\underline{\mathbf{x}}_i)\| > \varepsilon; \ \forall \ \underline{\mathbf{x}}_j \in \mathbf{X}_m^A; \ \forall \ \underline{\mathbf{x}}_i \in \mathbf{X}_n^A$ and $n \neq m$ (*Madatov, Sereda*, 2005). Each of the compact sub-spaces $\mathbf{X}_1^A$, $\mathbf{X}_2^A$,... $\mathbf{X}_M^A$ could be interpreted in this context as an *m*-th cluster of the search multidimensional sub-space $\mathbf{X}^A$ with its own specific sensitivity signature.

Therefore, calibration can be treated as a multistage exploration of Model Parameter space $\mathbf{X}$ within the predefined compact sub-spaces targeted at getting globally non-unique IP solutions, each of which has a specific sensitivity signature and a model interpretation (see s/section 6.2 to get some model examples).

APPENDIX 2: Review of involved minimization strategies

In connection with resolving of the production data inverse problem defined in Appendix 1 we will distinguish between step-wise and multi-start strategies of minimization. There are numerous literature descriptions of both strategies in applications to the local and global optimization problems (see for example reviews: *Himmelblau*, 1972; *Luenberger*, 1984; *Bunday*, 1984; *Press et al.*, 1992; *Bäck*, 1996; *Orlyansaya*, 2002). Here only general descriptions are presented for those methods which reveal good results on the 2D testbed (see s/section 5.2) and then were selected for extended tests of parametrical upscaling and calibration (see s/sections 3.2 and 4 respectively). The Evolution Algorithms are considered in more detail since they were selected as a basis for the combine algorithm targeted for multiple non-local solution of the specific inverse problem (see s/section 5.3).

A2.1. Step-wise strategies

All step-wise strategies of a local optimization regardless to the modification could be expressed in form of the following recursive routine:

$$\underline{\mathbf{x}}_{k+1} = \underline{\mathbf{x}}_k + q_k \underline{\mathbf{d}}_k, \ \forall \ \underline{\mathbf{x}}_k \in \mathbf{X}^S \subset \mathbf{X}; \ k=0,1,2,\dots.K. \tag{A2.1}$$

Here $\underline{\mathbf{x}}_{k+1}, \underline{\mathbf{x}}_k$ are consequent positions of unknown vector associated in our context with adjustable model parameters; index k indicates the number of steps in a consequent descent process. Note that the start ($\underline{\mathbf{x}}_0$), finish ($\underline{\mathbf{x}}_K$) and all intermediate positions of the vector-solution should belong to the same predefined multi-dimensional search window $\mathbf{X}^S$ within the model parameter space $\mathbf{X}$. Vector $\underline{\mathbf{d}}_k$ defines the direction of the descent on k-th step of the process. Scalar q_k defines the length of the given k-th step.

The descent process can be stopped or interrupted before any next k-th step as soon as one of stop criteria is satisfied. The solution is considered to be achieved and termination of the process is considered to be successful only if a misfit function $R(\underline{\mathbf{x}})$ yields a target condition (see Appendix 1 for more details). There are several reasons for an unsuccessful termination of the consequent process (A2.1): exceeding of an iteration limit, too slow convergence, too many penalty steps etc.

Checking the stop criteria and elaboration of the next step require multiple re-computation of the goal function $R(\underline{\mathbf{x}})$. Since the forward solution operator $\mathbf{M}\ [\underline{\mathbf{x}}]$ is enabled at this process, the total amount of $R(\underline{\mathbf{x}})$ computing required for elaboration of the each next step in (A2.1) is a critically important characteristic of any particular step-wise strategy.

A2.1.1. One-dimensional minimization

This family of strategies incorporates methods in which the direction vector $\underline{\mathbf{d}}_k$ do not need defining in the context of the general definition (A2.1) since there is only one possible direction – along x axis (see fig. A2.1).

The main idea behind all one-dimensional minimizations consists in consequent bracketing of an initial search interval on condition that each new bracket contains a minimum and it is narrower than the previous one. Of course, the initial search interval (X^S) must contain a point where a goal function is less than it is at the left and right edge positions.

Fig. A2.1. Getting minimum in 6 steps for 1D problem by using DSK Powell method (after *Himmelblau*, 1972)

So, when minimum of the goal function $R(x)$ is known to be bracketed within the $[x_a, x_c]_k$ interval, where $x_a < x_c$, the problem is reduced to finding the intermediate point x_b, such that $R(x_a) > R(x_b) < R(x_c)$ and $x_a < x_b < x_c$. The process of bracketing is continuing until one of the stop criteria will be achieved: deep enough minimum $R(x_b)$ is found (successful termination) or the distance between the two outer points $[x_a, x_c]_k$ is tolerably small (unsuccessful termination). Then triplet $[x_a, x_b, x_c]$ (a target bracket) is used for parabolic approximation and getting refinement of a draft solution. If a doubling of current step length does not yield a decrease of a goal function up to crossing of the search limit, then the step length is initialized and process is restarted in the reverse direction from the edge point reached. If a certain number of consequent reverses is exceeded, then the solution is associated with the minimum of two edge points.

In context of the general strategy (A2.1) defining the scalar q_k as a function of k is the main problem of a one-dimensional optimization. Indeed, the total amount of subdividing K of start interval $[x_a, x_c]_0$, which is necessary and sufficient for capturing of an acceptable $R(x)$ is critically important from a computing point of view. There are many modifications of this common strategy: Golden bracketing, gradient intersections, parabolic interpolations etc. (see, for example, *Himmelblau*, 1972), which all implement the general bracketing approach. The difference between these methods roughly consists of the particular algorithm recursively getting each next step length q_{k+1}, subdividing the relevant interval $[x_a, x_c]_{k+1}$ and capturing the draft solution x_b within the target bracket $[x_a, x_b, x_c]$.

The one-dimensional minimization strategy is illustrated in fig. A2.1 by the example of application of DSK Powell method with linear acceleration of the step length ($q_{k+1} = 2q_k$) at minimization of $R(x)$ function by using six consequent steps.

A2.1.2. Gradient methods of minimization

Gradient methods (see for ex. *Luenberger*, 1984) include a wide range of iterative procedures where the direction of the next step in the search strategy (A2.1) is reversed to the current gradient of a goal function $R(\underline{\mathbf{x}})$. It allows approaching to a nearest local minimum along steepest ascent by using locally linear approximation of $R(\underline{\mathbf{x}})$.

In terms of common definition of the search strategy, the gradient methods imply determining a full gradient as a direction vector $\underline{\mathbf{d}}_k = \underline{\nabla} R(\underline{\mathbf{x}}_k)$ and computing the simple step length q_k. The scalar q_k is defined as a solution of the one-dimensional minimization problem at each next k-th point of the descent process (see above). Thus, the steepest descent strategy in agreement with the general definition (A2.1) can be expressed as:
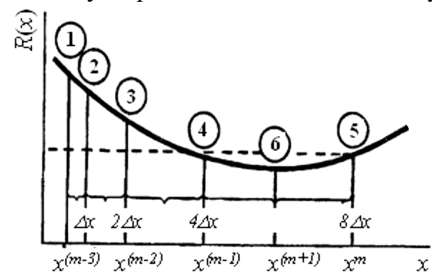
$$\begin{cases} \underline{\mathbf{x}}_{k+1} = \underline{\mathbf{x}}_k - q_k \underline{\nabla} R(\underline{\mathbf{x}}_k), \ \ k = 0,1,2,.....K. \\ q_k: R\left(\underline{\mathbf{x}}_{k+1}\right) = \min_{q>0}\{ \ R\left(\underline{\mathbf{x}}_{k+1}\right) \ \} \end{cases} \tag{A2.2}$$

The following final difference approximation of the Freshet derivatives is used for the numerical calculation of the gradient components:

$$\nabla_j R(\underline{\mathbf{x}}) = \partial R(\underline{\mathbf{x}})/\partial \mathrm{x}_i \cong R(\underline{\mathbf{x}} + \delta x_i \underline{\mathbf{e}}_i) - R(\underline{\mathbf{x}})/\partial \mathrm{x}_i, \ i = 1,2,\ldots,N, \tag{A2.3}$$

where $\partial \mathrm{x}_i$ – the small increment of the *i*-th component of the model parameter vector $\underline{\mathbf{x}}$; $\underline{\mathbf{e}}_i$ – the single vector having only *j*-th component non equal to zero but unity; N – the dimension of the relevant minimization problem.

From a computing point of view it is important to stress that every new iteration in (A2.3) requires N forward problems to be resolved (the operator $M[\underline{\mathbf{x}}]$ to be applied) for defining the direction vector $\underline{\mathbf{d}}_k$. Additionally some more forward problem solutions are obviously needed for defining the simple step length q_k according to a one-dimensional minimization routine.

Normally, few first gradient steps from randomly assigned start point lead to significant reduction in a misfit value $R(\underline{\mathbf{x}}_k)$. However, a descending process could easily be trapped by a nearest local minimum. When the multi-dimensional landscape of a goal function in local minimum vicinity is represented by an elongated gully the recursive iteration process (A2.2) becomes slowly converging. Unfortunately this is quite a typical situation for a multidimensional minimization problem, and a common practice for escaping from such a local trap is changing the direction from current gradient to conjugated one.

The main idea behind all conjugated gradient methods is based on the statement that the motion along some new direction $\underline{\mathbf{C}}$ does not spoil the minimization along the predefined direction $\underline{\nabla} R(\underline{\mathbf{x}})$ in the only case when the gradient stays perpendicular to it. Thus, new direction must be conditioned as follows:

$$\{ \ \underline{\mathbf{C}} ,\underline{\nabla} R(\underline{\mathbf{x}}) \ \} = 0, \tag{A2.4}$$

where operator $\{\underline{\mathbf{a}},\underline{\mathbf{b}}\}$ denotes scalar product of two vectors $\underline{\mathbf{a}}$ and $\underline{\mathbf{b}}$.

In the general case the relevant descent strategy comes up with a set of N linearly independent, mutually conjugated directions.

Combining of a steepest descent with a descent in conjugated direction allows quick and effective analysis of a local multidimensional landscape of the goal function $R(\underline{\mathbf{x}}_k)$. This is especially true when the conjugated direction coincides with locally most sensitive component of the desired vector $\underline{\mathbf{x}}$.

A2.2. Multi-start strategies

The common disadvantage of any step-wise strategy is the trapping of the iteration process (A2.1) by the nearest local minimum. Thus, strictly speaking, only one specific class of goal misfit functions – unimodal continuous – can be accepted as being favourable for step-wise minimization methods. Most probable shape of a goal function given in the form of model constrained for misfit representation (see A1.3-A1.3**) or likelihood function (see A1.5) is a multi-modal shape. At this, the amount of local minima increases exponentially with size of the dimension of the relevant minimization problem. In addition, all step-wise strategies fail when a goal function $R(\underline{\mathbf{x}})$ at the current $\underline{\mathbf{x}}$ position reveals some singularity and thus its derivatives cannot be computed.

The approach to the minimization of broader class of goal functions $R(\underline{\mathbf{x}})$, which include discontinuities and multi-modal behaviour implies vast exploring of a complex multi-dimensional $R(\underline{\mathbf{x}})$ topography, where many local minima will be most probably explored before stop criterion conditions will be satisfied[5].

Multi-start strategies of optimization incorporate a wide class of the optimization methods oriented on similar vast exploration of a potentially multi-modal goal function $R(\underline{\mathbf{x}})$, defined on multi-dimensional vector space $\mathbf{X}$ (see for example *Bäck*, 1996).

Formally any of multi-start strategies could be expressed by analogy with (A2.1) in the following recurrent form:

$$\{\underline{\mathbf{x}}\}_{k+1} = \mathrm{E}_k^1[\ \mathrm{E}_k^2 [\ \ldots [\mathrm{E}_k^L \ \{\underline{\mathbf{x}}\}_k \ ]]], \ \forall \ \{\underline{\mathbf{x}}\}_k \in \mathbf{X}^S \subset \mathbf{X}; \ k=0,1,2....K, \tag{A2.5}$$

where $\{\underline{\mathbf{x}}\}_k$ denotes a local sub-set of vector-solutions of a minimization problem, which belongs to the search sub-space $\mathbf{X}^S$ within the vector solution space $\mathbf{X}$ at any stage k of the process; $E_k^{1,2,...L}$ denotes the suite of operators which generate modifications in the positioning of $\{\underline{\mathbf{x}}\}_k$ at every new successive stage k of the recurrent process.

---

[5] Note that the search for the global minimum is senseless in the context of conditionally unique inverse problem solution considered in Appendix 1.

Multi-start term implies that in contrast with the step-wise strategy this group of methods deals with the set of vectors $\{x\}_k$ at every stage instead of tracing a single trajectory of the descent. Therefore, relevant approaches do not require the derivative information.

Apart from pure random search (Monte-Carlo method), all other modifications of a multi-start strategy at the stage $k>1$ use the information achieved at all previous k-1 stages for updating of the operators $E_k^{1,2,...L}$. Thus, initially randomly exploring topography of a goal function gradually obtains more and more regular features.

In connection to the considered specific inverse problem we will review two groups of multi-start methods: Simulated Annealing approach (SA) (*Kirkpatrick et al.*, 1982; *Ingber*, 1993) and Evolutionary Algorithms (EA) (*Bäck*, 1996; *Pohlheim et al.*, 1999; *Voigt, Anheyer*, 1994). Note, that the last group will be described in more detail since it represents the core of a newly developed combined algorithm (see s/section 5.3).

A2.2.1. Simulated annealing

Simulated Annealing (SA) is a global optimization strategy, which exploits an analogy between the way in which a metal cools and freezes into a minimum energy crystalline structure (the annealing process). The essence of the slow cooling process is allowing ample time for redistribution of the atoms as they lose mobility. If the current system energy at certain stage of cooling can be associated with the goal function $R(\mathbf{x}_k)$ at k-th stage of minimization, then the system temperature can be by analogy treated as a main external control factor. This forms the basis of the considered optimization technique and its main distinction from all the above-considered techniques.

Indeed, all of the above-considered step-wise minimization algorithms are searching for immediate downhill descent without any external control on "system energy". Thus, to continue the metal cooling analogy, they are all "quenching" the initially hot system.

From this point of view the external control principle as a key feature of the SA approach can be applied for any step-wise strategy on condition that some of the bad (uphill) steps could become acceptable depending on external controlling factor. Still, historically the SA approach appeared as a random-search technique (*Metropolis et al.*, 1953). So, in context of this review we include it in the broad group of multi-start strategies.

Regardless of the form of the recurrent process (A2.5), it essentially performs two main repeated tasks between its initialization and its termination: updating current the solution set $\{\mathbf{x}_{k+1}\}$ and checking the stop criteria. The Simulated Annealing approach replaces the start-stop descent strategy by gradual descent strategy where uphill steps could be accepted and achievement of rather low value for goal function does not mean termination of the process. Instead the recurrent process (A2.5) is entirely controlled by annealing schedule that can be expressed in form of Boltzmann probability distribution:

$$Prob\,(E) \sim \exp(-dE/kT)\,, \tag{A2.6}$$

where k is the Boltzmann's constant, that relates the temperature T with the energy E in the thermal equilibrium state.

This formula expresses the idea that a system in thermal equilibrium has its energy probabilistically distributed among all different energy states E. Even at low temperature, there is a chance, albeit very small, of a system being in a high energy state. Therefore, there is a corresponding chance for the system to get out of a local energy minimum in favour of finding a better, more global, one. In context of the considered minimization process (A2.5) this schedules the descent process in such a way that it sometimes goes uphill as well as downhill; but the lower the temperature, the less likely is any significant uphill excursion.

Accepting each particular perturbation step $d\underline{\mathbf{x}}_k = \|\underline{\mathbf{x}}_{k+1} - \underline{\mathbf{x}}_k\| / \|\underline{\mathbf{x}}_k\|$ in the Simulated Annealing approach (see for example algorithm *Dueck, Scheuer*, 1990) is according to the probability $Prob_k$ given by:

$$Prob_k\,[R_k(\underline{\mathbf{x}})] = \exp\,[-dR_k(\underline{\mathbf{x}})\,/Td\underline{\mathbf{x}}_k], \tag{A2.6*}$$

where $dR_k(\underline{\mathbf{x}})$ is the current increment of the goal function; $d\underline{\mathbf{x}}_k$ is the average step size; T is the cooling temperature, scheduled in agreement with a particular annealing method. So that $-dR_k(\underline{\mathbf{x}})/d\underline{\mathbf{x}}_k$ is a measure of the effectiveness of the change made by the given step. As the size of the step taken is considered in calculating $Prob_k$, it does not need to be adjusted when T is changed.

The annealing schedule determines the degree of uphill movement permitted during the search and is thus critical to the algorithm's performance. The principle underlying the choice of a suitable annealing schedule is easily stated: the initial temperature should be high enough to "melt" the system completely and should be reduced towards its "freezing point" as the descent progresses. Choosing an annealing schedule for practical purposes is something of an art. The standard implementation of the SA algorithm is one in which homogeneous Markov chains of finite length are generated at decreasing temperatures.

The verity of SA approaches is essentially defined by means of specification of the initial temperature T_0; the final temperature T_f or a stop criterion; a length for the Markov chains and the rule for decrementing the temperature.

The major advantage of SA over other methods is an ability to avoid trapping of a minimization process by a local minimum. It has been proved by carefully controlled rate of the cooling, SA can find the global

optimum (*Goffe et al.*, 1994). However, this could require infinite time. Fast Annealing and Very Fast Simulated Re-annealing (VFSR) or Adaptive Simulated Annealing (ASA) are each exponentially faster and overcome this problem (see review *Ingber*, 1993 for more details).

A2.2.3. Evolutionary methods

Evolutionary Algorithms (EAs) are mimic global minimization methods that take their inspiration from natural selection and survival of the fittest in the biological world. EAs differ from step-wise optimization techniques in that they involve a search from a "population" of solutions, not from a single point. In contrast with the above-reviewed Simulated Annealing methods, which formally could follow start-step descent strategy, an EA strategy is the originally multi-start descent strategy. Each EA iteration involves a competitive selection that weeds out poor solutions. The solutions with high "fitness" are "recombined" with other solutions by substituting parts of a solution with another. Solutions are also "mutated" by adding a small change to every single element of the solution (see Table A2.1). Recombination (crossover) and mutation are used to generate new solutions that are biased towards regions of the search sub-space for which good solutions have already been seen.

Like the Simulated Annealing, the Evolutionary Algorithms got their development impetus at the problem of global minimization of a multi-dimensional goal function with singularity features. Here their multi-start origin reveals ability to explore vast areas within the search sub-space and allocate there a great number of local minima suitable as conditionally correct solutions of the relevant minimization problem.

Additionally the Evolutionary Algorithms are intrinsically parallel in contrast with above-reviewed step-wise strategies, which are serial[6]. Since EAs have multiple offspring, they can explore a solution space **X** in multiple directions at once. So, implementation of the parallel computation releases relevant computer programs from the main drawback of multidimensional optimization, i.e. high computing cost (*Pohlheim et al.*, 1999).

Here we briefly describe the main stages of an EA strategy while focusing on crucial topics important in the context of the considered specific inverse problem. We recommend the following reviews of evolutionary approaches in minimization (*Goldberg*, 1989; *Voigt, Anheyer*, 1994; *Bäck*, 1996).

The upper level workflow diagram (fig. A2.2) illustrates the main elements of any evolutionary strategy in their interaction. In agreement with this scheme and in connection to the general formula (A2.5) one can highlight 3 consequently applied operators:

Fig. A2.2. Upper level work-flow diagram of generalized Evolutionary Algorithm

- selection – getting vector $\underline{x}^*$, that is most suitable for continuation, from the current solution subset $\{\underline{x}\}_k$ based on some fitness criteria;
- crossover and mutation – generation of next solution subset $\{\underline{x}\}_{k+1}$.

Recursive repetition of this main operator triplet should ensure successive improvement of a fitness function given as $Aver[R\{\underline{x}\}]$, where *Aver* denotes an operator of averaging of the $R(\underline{x})$ values over the population $\{\underline{x}\}$. Each round of this process should clean up population of vector solution sub-set $\{\underline{x}\}_k$ from unsuccessful members. If we use analogy $\{\underline{x}\}_k \rightarrow \{\underline{x}\}_{k+1}$ with set of pairs "parents → child", then such solution refinement on average is accomplished by involving in continuation just a few of the parents, namely those which reveal the best fitness quality. A child solution could be randomly perturbed (mutated) before being selected to the next stage of the evolution in order to avoid trapping of a current solution sub-set in a local minimum vicinity.

We suggest the following translation table in order to establish link between the specifics of considered forward and inverse problems and generally accepted EA terminology.

Table A2.1. The EA terminology vs. specific inverse and forward problems terms

| EA term	Context of considered inverse and forward problems
Phenomenon	Absolute and relative permeability distributions within the 3D grid representing the given framework and texture of reservoir model
Encoding of phenomenon	Specification of an effective RM by using control model parameters set (see part 1 of this paper)
Population at the given evolution stage	Subset of control model parameter vector $\{\underline{x}\}_k$ involved in simulation of the real production data and an effective RM fitting at k-th stage of an evolution process
Individual	$\underline{x}_{ik} \in \{\underline{x}_1, \underline{x}_2, …, \underline{x}_L\}_k = \{\mathbf{x}\}_k$
Population size	L – the total number of vectors $\underline{x}$ belonged to the subset $\{\underline{x}\}_k$ (can be varied during minimization progress)
Parents	Input sub-set $\{x^\wedge\}_k \subset \{x\}_k$ selected from subset $\{x\}_k$ for producing offspring at k-th stage of an evolution process
Offspring	Output subset of control model parameter vector $\{\underline{x}\}_{k+1}$ at k-th stage of an evolution process
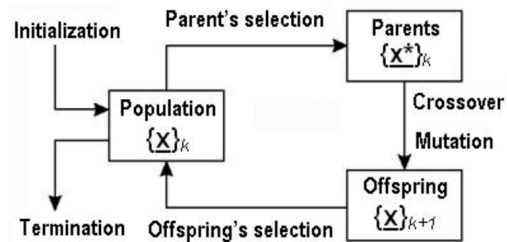
---

[6] The serial algorithm can only explore the solution space in one direction at a time, and if the solution turns out to be suboptimal, there is nothing to do but abandon all previous work and start again.

Fitness function   R($\underline{\mathbf{x}}$) – the goal function used at the data inverting (misfit)

Population diversity   Diameter of the subset $\{\underline{\mathbf{x}}\}_k$ given by metric introduced on the Model Parameter space $\underline{\mathbf{x}}$[7]

Each of the EA operators indicated in the general scheme (fig. A2.2) is considered below in more detail.

Selection

The aim of selection in the minimization context is to successively improve data fitting quality in average over the sub-set $\{\underline{\mathbf{x}}\}$ during evolution stages. This operator implemented on the whole population is essentially stochastic in terms of measure, which it uses for producing output. Namely, based on the fitness estimation over the given subset $\{\underline{\mathbf{x}}\}_k$ it involves the best individuals into the new generation $\{\underline{\mathbf{x}}\}_{k+1}$ with subsequent higher probability. In other words, the better the parents are in terms of fitness function, the higher their chance of participating in producing the offspring.

Let every individual of *k*-th population have its unique rank based on its fitness estimation. So, the sorting $\{\underline{\mathbf{x}}_1,\underline{\mathbf{x}}_2,….,\underline{\mathbf{x}}_L\}_k$ is according to sequence $[R_1<R_2,…. R_{L-1}<R_L]_k$. Let the full range of the goal function variation from acceptable *Rmin* up to non-considerable *Rmax* misfit levels be subdivided into *n* equal intervals. The empirical distribution function $S_k(R)$ indicates how many individuals of the given population belong to each of *n* intervals at the given evolution stage *k*. Clearly, if the population size $L\rightarrow\infty$ and the number of intervals $n\rightarrow\infty$ then estimation $S_k(R)$ will approach to a probability density function – PDF. Thus, the PDF estimation $S_k(R)$ is called *Fitness Distribution* of *k*-th population. The number of the $S_k(R)$ extremes and their configuration gives a modality of relevant PDF estimation.

At the beginning of process (*k*=0), when start positions of individuals within the search sub-space are random, the function $S_0(R)$ obviously has fuzzy and multi-modal form where maximum is close to *Rmax* (the black curve in fig. A2.3). This form could not be changed on average during the EA progress if a random selection is used. In contrast, when selection follows certain evolution scenarios, then obviously the relevant fitness distribution will transform into the unimodal distribution form with the maximum tending to *Rmin*. The faster such transformation comes the more successful the corresponding evolution is. A typical example of a successful EA implementation at the minimization of 2D Rozenbrock function is represented in fig. A2.3.

The following quality indicators are important for comparison of different approaches:

Reproduction Rate

This indicator is represented by ratio of $S_k(R)$ estimations for adjacent stages in evolving:

$$Rt(k) = \begin{cases} \dfrac{S_{k+1}(R)}{S_k(R)}, & \text{if } S_k(R) > 0 \\ 0, & \text{if } S_k(R) = 0 \end{cases}$$

Fig. A2.3. The PDF estimation of $S_k(R)$ for standard EA testing at minimization of Rozenbrock function *L*=100; *N* = 10 (after *Bäck*, 1996)

Population diversity

Evidently during the selection some of the low ranked vectors (parents) will be removed from the subset $\{\underline{\mathbf{x}}\}_k$ (current population). This could lead to shrinking of the population around its best representative(s) and finally to a population collapse. The measure of this process is called diversity *Ds*(*k*). It is expressed as following:

$$Ds(k) = \sum_{l=1}^{L} [(\| \underline{\mathbf{x}}_l^k - \overline{\overline{\underline{\mathbf{x}}}}^k \|) / \| \overline{\overline{\underline{\mathbf{x}}}}^k \|]/L,$$

where $\underline{\mathbf{x}}_l^k$ denotes *l*-th individual and $\overline{\overline{\underline{\mathbf{x}}}}^k$ median vector of *k*-th population (see also table A2.1).

Selection Intensity (*Whitley*, 1999) is a measure of average improvement of the fitness rank for the given population at *k*-th stage of an evolutionary process, given by:

$$Is(k) = \frac{\overline{R^{k+1}} - \overline{R^k}}{\sigma} \quad ,$$

---

[7] The bigger the population diversity at later stage of the minimization process the higher the probability to get multiple solution of relevant minimization problem.
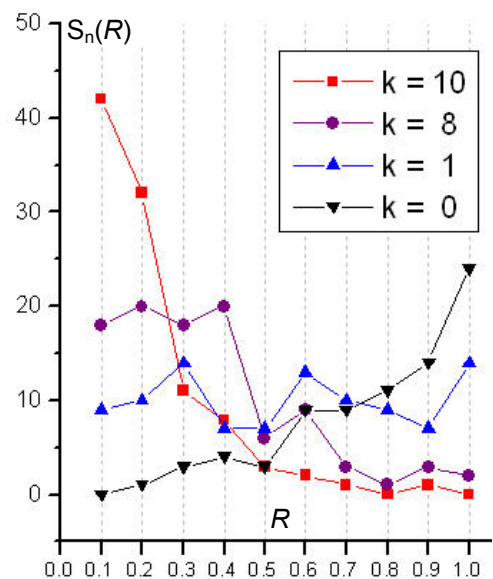
where $\overline{R}^k = \dfrac{1}{L}\int\limits_{R\min}^{R\max} S_k(R)R\,dR$; the upper accentuation denotes averaging over population size; σ is the standard increment level.

The standard selection algorithms, which have been tested during development of the combined algorithm include the following (*Holland*, 1975; *Jong*, 1975; *Syswerda*, 1989; *Grefenstette*, 1986; *Grefenstette, Baker*, 1989):

1. *Elitist selection* implies copying part of the parents which has upper ranks into the offspring population with no changes. When this part $M < L$ is constant during all stages, the corresponding selection is called *Steady State selection*.

2. In *Proportional selection* the probability of the given parent's individual ($\underline{x}_l \in \{\underline{x}\}_k$) to be selected for producing next generation ($\{\underline{x}\}_{k+1}$) is proportional to its fitness values.

3. The similarity measure defined on space **X** is added to the misfit value to form fitness function in the *Selection with Thickening*. In terms of goal function formulation this kind of selection implies the $R(\underline{x})$ form extended by soft model constraining (see formula A1.3*).

4. The algorithm of *Tournament selection* includes two successive steps: a random sample set with the size M ($2 < M < L-2$) is selecting from current sub-set $\{\underline{x}\}_k$; the best individual ($\underline{x}_l \in \{\underline{x}\}_k$) is selected from this sample set based on rank analysis. This process is repeating L times while the new generation $\{\underline{x}\}_{k+1}$ will be completed.

5. In the *Linear Ranking selection)* the individuals are sorted within the population inversely to the relevant rank. I.e. the worst $\underline{x} \in \{\underline{x}\}_k$ gets $Rank = \eta_1$ and the best gets $Rank = \eta_L$. The probabilities of selecting these two vectors into the parents' set are $P_1 = \eta_1/L$ and $P_L = \eta_L/L$ respectively. So, the probability of selecting any $\underline{x} \in \{\underline{x}\}_k$ with intermediate rank is then assigned according to linear function:

$$P_l = \frac{\left[\eta_l + (\eta_L - \eta_1)\dfrac{l-1}{L-1}\right]}{L}.$$

We recommend also the review (*Bäck*, 1996) for getting more details about this selection of algorithms.

Recombination or Crossover

This operator targets to produce an offspring population based on parent's sub-set selected in the previous stage. It is successively implemented to the least pair of the selected parent individuals. In the general case the amount M of parents used to produce single offspring representatives (children) could be more then two (which is the natural choice)

Formally, this operator can be expressed as follows:

$$Ec[\{\underline{x}^*_m\}]: (\underline{x}_l)^{k+1} = (\underline{x}^*_1 \cup \underline{x}^*_2 \cup \ldots \cup \underline{x}^*_m)_k, \forall \underline{x}^*_{1,2-m} \in \{\underline{x}^*_m\} \subset \{\underline{x}\}_k \subset \mathbf{X}^S; \ \forall \underline{x}_l \in \{\underline{x}\}_{k+1} \subset \mathbf{X}^S \tag{A2.7}$$
$$l = 1,2,\ldots L; \quad m = 1,2,..M; \quad M \le L.$$

As it follows from this formula, a crossover aims to keep every offspring individual $\underline{x}_l \in \{\underline{x}\}_{k+1}$ combination of components inherent to the selected parents' sub-set $\{\underline{x}^*_m\} \subset \{\underline{x}\}$ with the probability proportional to the rank of each parent ($\underline{x}^*_{1,2-m} \in \{\underline{x}^*_m\}$).

So, this operator is also essentially stochastic.

Mutation

This operator targets for random perturbation in components of vector $\underline{x} \in \{\underline{x}\}$. Normally, it is successively implemented on every single individual in the offspring sub-set $\{\underline{x}\}_{k+1}$.

Formally, this operator can be expressed as follows:

$$Em[\underline{x}_l]: \underline{x}_l = \underline{x}^*_1 + \underline{r}(\sigma), \ \forall \underline{x}_l \in \{\underline{x}\} \subset \mathbf{X}^S; \forall [\underline{x}_l + \underline{r}(\sigma)] \in \{\underline{x}\} \subset \mathbf{X}^S; l = 1,2,\ldots L, \tag{A2.8}$$

where $\underline{r}_k(\sigma)$ is a random perturbation vector defined on vector space **X**. The amplitude of a perturbation is controlled by the current dispersion level σ and by the type of function $\underline{r}(\sigma)$ (*Neighbourhood Function*). In particular, it could be the Gauss function symmetrically distributed around "0" with the dispersion σ.

The important element of any mutation vector is adaptation of the dispersion level σ to the rank of input vector $\underline{x}_l$ and the type of selection, which has been implemented for its generation. For example, in the Reachenberg's algorithm (Strategy "1+1") the perturbation step increases when more than 20 % of new generation individuals reveal better then parents' fitness and vice versa. In the self-learning Bäck's algorithm (*Bäck*, 1996) the perturbation step can be kept, or increased, or decreased with equal probability 0.3. Therefore,

the population {**x**} is subdivided into three equal size sub-sets – regressive, stable and progressive – based on rank updating. The relevant perturbation step $\mathbf{r}_k(\sigma)$ is implemented in each sub-set.

Both operators Crossover and Mutation are normally combined in the way to provide maximal mutation steps and crossover success along the main axis of hyper-ellipse, which defines probability of maximal evolution progress (*Voigt, Anheyer*, 1994).

APPENDIX 3: The reference reservoir model and the simulation results used as a data prototype

Here the plausible earth model is introduced together with the production & injection well system. The structure and texture features of this model are prototypes of the real oilfield system attached to one of Gulf of Mexico mini-basin (*Bridges et al.*, 2001), whereas the multi-well system is doubly speculative. This input was served for initialization of the relevant reference reservoir model and then for simulation of the production history data set, which was assumed to be applicable for data inversion at both parametrical upscaling and calibration.

As seen on the framework model of fig. A3.1 the prototype oilfield is represented by the anticline trap subdivided by five fault segments into four blocks. The reservoir has six-layered texture where the fine flow units R1-3 are gradually replaced with the poor flow units C1-3 in a north-east direction (see cross section in fig. A3.1b). In agreement with the basic concept the fault displacement and the reservoir geometry, that potentially control a geo-fluid system during production, are assumed to be fixed. The oil in place and the initial OWC position are supposed to be defined based on the results of the first stage history matching (see s/section 2.3.1). The PVT properties of water and oil are normally predefined based on lab investigations (*Zeltov*, 1998; *Denesh*, 1998). The initialization of a porosity cube defined on the given 3D grid was based on Athy's compaction models (see s/section 5.1 in part 1 of this paper).

The porosity cube is used for conversion to an absolute permeability cube based on an anisotropic model for continuous parts of the reservoir and on a transmissibility model for the discontinuous parts of the reservoir (see s/sections 5.1 and 5.2 respectively in part 1 of this paper). Computing and updating of the relative permeability values were performed in agreement with the modified Muskat's formula (see s/section 5.3 in part 1 of this paper).
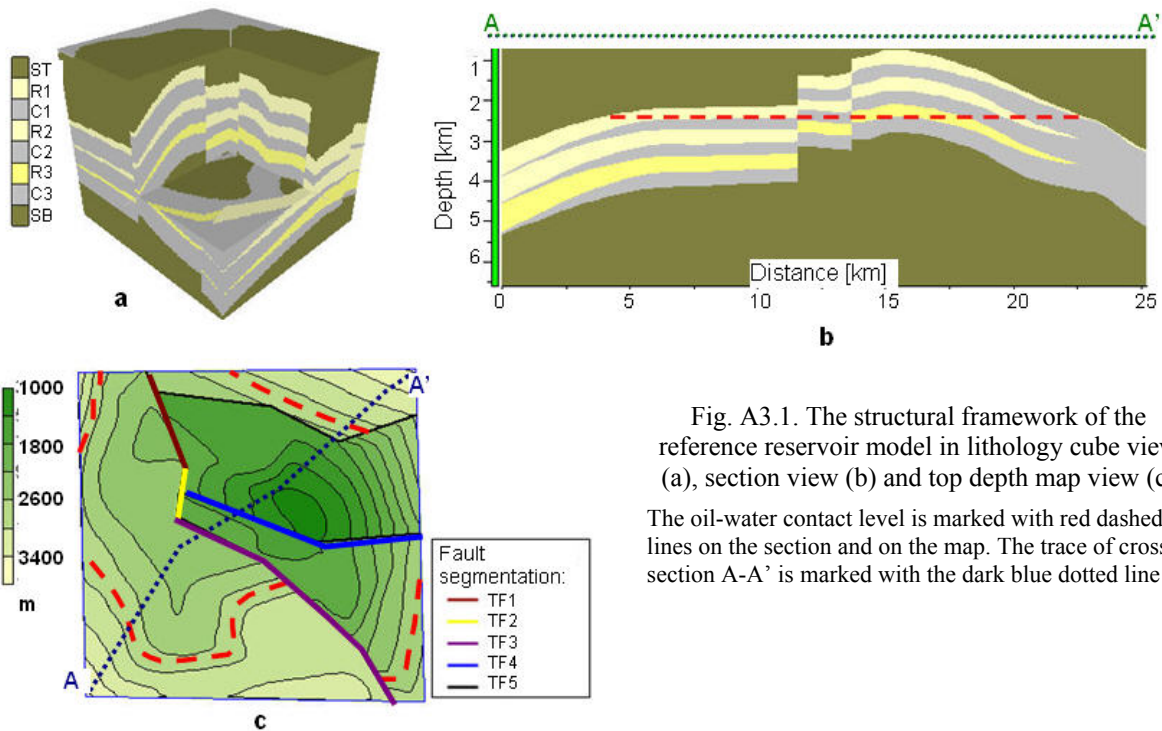

Fig. A3.1. The structural framework of the reference reservoir model in lithology cube view (a), section view (b) and top depth map view (c)

The oil-water contact level is marked with red dashed lines on the section and on the map. The trace of cross-section A-A' is marked with the dark blue dotted line

Table A3.1. The specification of the CMP components for the reference reservoir model

Specification type	index	CMP components				
		A0	A1	A2	Log(Ss)	Log(Sm)
Flow units	R1	0.25	0.11	-0.06	-2.51	-1.40
	S1	0.54	0.15	-0.06	-1.63	-0.72
	R2	0.05	0.00	0.00	-1.93	-1.62
	S2	0.55	0.21	-0.90	-0.98	-0.82
	R3	0.14	0.30	-0.05	-2.75	-2.11
	S3	0.67	0.5	-0.16	-2.04	-0.97

Flow barriers	index	Uncertain Clay Content Fraction
	TF1	0.22
	TF2	-0.58
	TF3	0.24
	TF4	-0.75
	TF5	0.05

The shaliness required for both absolute and relative permeability calculations was uniquely defined by a linear trend function assigned for each of the six flow units via relevant coefficients ($A0$, $A1$, $A2$).

The control model parameters necessary for specification of the reference reservoir model are collected in Table A3.1. As it follows from Table A3.1 the full dimension of the CMP space introduced for the given reference reservoir model is given by $N = 6\times5 + 5\times1 = 35$.

The **well schedule** description during a speculative production history was supposed to be available in following form suitable for simulation.

Phase 1
(13 time steps):

Production well P1 – shut;
Production well P2 – open BHP = 1020 psia;
Production well P3 – open BHP = 1020 psia;
Injection wells I1 – open with rate = 50000stb/day;
Injection wells I2 – open with rate = 50000stb/day;
Injection wells I3 – shut;
Injection wells I4 – open with rate = 50000stb/day;

Phase 2
(6 time steps):

Production well P1 – open BHP = 1020 psia;
Production well P2 – open BHP = 1020 psia;
Production well P3 – shut;
Injection wells I1 – shut;
Injection wells I2 – open with rate = 70000stb/day;
Injection wells I3 – open with rate = 70000stb/day;
Injection wells I4 – open with rate = 70000stb/day;

Phase 3
(8 time steps):

Production well P1 – open BHP = 1020 psia;
Production well P2 – open BHP = 1020 psia;
Production well P3 – open BHP = 1020 psia;
Injection wells I1 – shut;
Injection wells I2 – open with rate = 90000stb/day;
Injection wells I3 – open with rate = 90000stb/day;
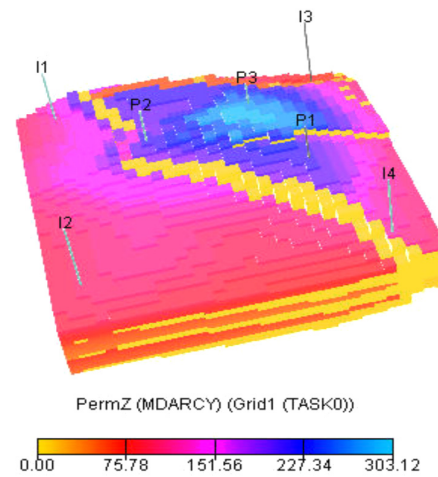Injection wells I4 – open with rate = 90000stb/day.

Fig. A3.2 The absolute permeability model (Z-component) with the well positions marked
Permeability scale is given in mD

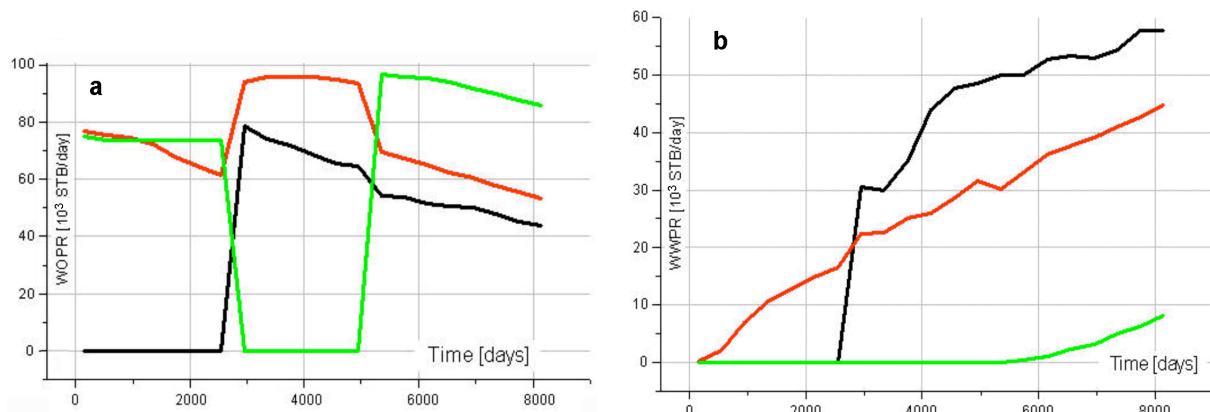The results of simulation are represented in fig. A3.2 and A3.3.

Fig. A3.3. The well production rate curves for oil (a) and water (b)
used as a real production data prototype for model tests

540