History Matching In Parallel Computational Environments

Final Report

Authors: Dr. Steven Bryant, Dr. Sanjay Srinivasan, Alvaro Barrera, Yonghwee Kim and Sharad Yadav

Date Issued: December 2006

DOE Award Number: DE-FC26-03NT15410

Submitting organization: Center for Petroleum and Geosystems Engineering
The University of Texas at Austin
Austin, Texas 78712-0228

Summary

In the probabilistic approach for history matching, the information from the dynamic data is merged with the prior geologic information in order to generate permeability models consistent with the observed dynamic data as well as the prior geology. The relationship between dynamic response data and reservoir attributes may vary in different regions of the reservoir due to spatial variations in reservoir attributes, fluid properties, well configuration, flow constrains on wells etc. This implies probabilistic approach should then update different regions of the reservoir in different ways. This necessitates delineation of multiple reservoir domains in order to increase the accuracy of the approach.

A general procedure for gradual updating of geological models within an assisted history matching framework was developed as part of this project. We have also developed a robust scheme for identifying reservoir regions using principal components analysis of sensitivities that will result in a more robust parameterization of the history matching process. In order to render the domain delineation procedure feasible regardless of the flow simulator used for the purpose of history matching, a unique new scheme was developed that utilizes the variance of grid block pressure values calculated over a suite of realizations. Several examples of application of this new approach to domain delineation and its integration within the history matching framework is demonstrated in this report.

A key accomplishment of the research team was the development of a integrated software for history matching Pro-HMS. The software interface of Pro-HMS integrates several codes that have been developed to render the history matching procedure efficient. These include:

1. An algorithm for calculating sensitivities using an upscaled ensemble of reservoir realizations
2. Principal component analysis for the determination of most sensitive, least interacting sub-domains
3. Probability perturbation within subdomains

The software modules were themselves developed in C++. The software interface is written in Qt programming language to render it feasible to execute Pro-HMS in Mac, PC and Unix/Linux environments.

A brief description of the new scheme for domain delineation using an ensemble of realizations is included in this report. That is followed by a discussion of several case examples performed using the Pro-HMS software. Finally, a user manual for Pro-HMS is included in the Appendix to this report.

An Improved Domain Delineation Procedure

The traditional approach for determining sensitive regions i.e. locations where the permeability value has an important influence on the observed flow response is based on the Hessian matrix that is calculated internally by the flow simulator. The Hessian of a function is the matrix whose components are the second partial derivatives of the function. In the case of this presentation, the function is the objective function and the partial derivatives are evaluated with respect to permeability.

The Hessian matrix is defined as follows:

$$S = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right) & \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial z}\right) & \mathrm{K} \\ \dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial x}\right) & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial z}\right) & \mathrm{K} \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} \\ \Lambda & \Lambda & \Lambda & \Lambda \end{bmatrix} \quad (1)$$

Where: f is the objective function
x,y, and z are parameters

There are two issues with the use of the Hessian for domain delineation:

1. The computation of the Hessian is cpu expensive.
2. Not all commercial simulators have the facility to output the Hessian matrix

In order to circumvent these issues and render the domain delineation procedure feasible regardless of the flow simulator used, an alternate scheme for determining sensitivities was formulated.

Description of Sensitivity using an ensemble of reservoir models

Sensitivity is defined as the degree to which a physical quantity (or the objective function in the case of an iterative solver) may be affected by another physical quantity (permeability). In other words, it describes how the pressure varies as the permeability in a specific grid block is changed. An additional drawback in using the Hessian matrix as a measure of sensitivity is that the pressure response is assumed to be sensitive to the permeability values at a maximum of two locations. The connectivity of permeability is thus ignored while calculating the sensitivities.

The grid block pressures reported by the flow simulator are obtained by solving the finite difference form of the flow equations. The pressures therefore take into account the connectivity of permeability. At locations close to the wells, the variations in underlying permeability will cause large variations in the corresponding grid block pressures because of the high fluxes in those regions. At locations far away from the wells, variations in permeability will cause only minor variations in grid block pressures due to the low flux in those regions. The variance of pressure computed over a suite of realizations, is therefore a good measure of the sensitivity of the well response to variation in the underlying reservoir properties.

The off-diagonal terms in the Hessian (Eq. 1) measure the redundancy between the sensitivity of the response variable to one variable (or permeability value at one location) and that due to another variable (or the value at another location). An alternate way to represent this redundancy is through the covariance between the pressure values at two different grid block locations. Thus the covariance matrix of pressure:

Sensitivity Coefficients (Variances)

$$\begin{pmatrix} \text{cov}(P_1,P_1) & \text{cov}(P_1,P_2) & \Lambda & \Lambda & \text{cov}(P_1,P_{n-1}) & \text{cov}(P_1,P_n) \\ \text{cov}(P_1,P_2) & \text{cov}(P_2,P_2) & \Lambda & \Lambda & \text{cov}(P_2,P_{n-1}) & \text{cov}(P_2,P_n) \\ \text{M} & \text{M} & \Lambda & \Lambda & \text{M} & \text{M} \\ \text{cov}(P_1,P_{n-1}) & \text{cov}(P_2,P_{n-1}) & \Lambda & \Lambda & \text{cov}(P_{n-1},P_{n-1}) & \text{cov}(P_{n-1},P_n) \\ \text{cov}(P_n,P_1) & \text{cov}(P_n,P_2) & \Lambda & \Lambda & \text{cov}(P_n,P_{n-1}) & \text{cov}(P_n,P_n) \end{pmatrix}$$

is a surrogate to the Hessian shown in Eq. 1. Note that the covariance matrix is also of order n x n just like the Hessian.

The procedure for calculating the sub-domains using the covariance matrix of pressure therefore consists of the following steps:

1. **Generate an ensemble of initial reservoir permeability models.**
   Sequential indicator simulation is used to get initial models conditioned to static data.

2. **Perform flow simulations on an ensemble of initial realizations that have been upscaled**
   - After running flow simulations, the pressure response at each grid block of each realization is obtained.
   - The pressure data are calculated based on the permeability distribution of initial realizations that have been input to the flow simulator.

3. **Construct covariance matrix**
   The covariances are calculated using the ensemble of realizations. The expected values in the definition of the covariances are calculated over the suite of realizations.

4. **Perform principal component analysis of eigenvalues obtained from covariance matrix**
   - Principal component analysis (PCA) is a multivariate data reduction method. The factors or groupings of pressure regions are constructed in a way that reduces the overall complexity of the problem and takes advantage of inherent interdependencies in the data. Thus regions of the reservoir that have a similar behavior in grid block pressures will be grouped together.
   - The covariance matrix is subjected to PCA and the five highest eigenvalues and their corresponding eigenvectors are retained. The highest five eigenvalues are enough to describe the sensitivity of covariance matrix.

5. Scaling and sorting of eigenvalues and eigenvectors

1) Scale and rank the eigenvalues based on their magnitude.
2) Compare one eigenvector component which is from the first eigenvalue with the other ones which are from the rest of eigenvalues (Note: This comparison should be done at the same location with different eigenvalues.)
3) Choose the biggest one and sort them according to their magnitude.
4) Mark what eigenvalues have the largest eigen-component.
5) Store eigenvector components and corresponding eigenvalues (these eigenvalues and eigenvector components are ones obtained in part 3) & 4)) in an array.
6) Apply threshold (volume cutoffs) to an array that we have in part 5) to define size of sub-domains.
7) The size should be optimal so as to enhance effective perturbation process. (Note: If the size is too big, increased interaction between sub-domains is expected. On the other hands, if it is too small, the history matching will be ineffective.)

6. Obtain sub-domains

– Since sub-domains are most sensitive and least correlated (reasons are explained in below), we can perturb conditional probability distribution describing permeability values of the sub-domains simultaneously.

  The parallel computing environment enables perturbations of different sub-domains at the same time. This method reduces computational cost.

– Since the sub-domains are obtained by manipulation of sensitivity matrix, the sub-domains identified are always most sensitive. This is further reinforced by picking the highest eigenvalues after the PCA procedure.

– The eigenvalues obtained by PCA are orthogonal (at least in a covariance sense). The physical meaning of orthogonality is that the regions are independent or uncorrelated. Any residual dependence between the identified sub-domains can be removed by performing eigen-rotation procedures.

7. Simple example

The conditions and procedures are explained in below. Figure 15 is the schematic of the process.

- Suppose the size of a reservoir is 4x4 (i.e. a 2D reservoir with 4 grid blocks in the x and y directions)..
- Permeability data for each grid block are given. They are generated by sequential indicator simulation.
- Run a flow simulator and get pressure data for four grid blocks.
- Calculate the covariance of permeability between one grid block with the rest three grid blocks.
- The size of the matrix will be 4x4.
- The diagonal terms will be variances of the pressure because covariance is calculated with respect to the same block.

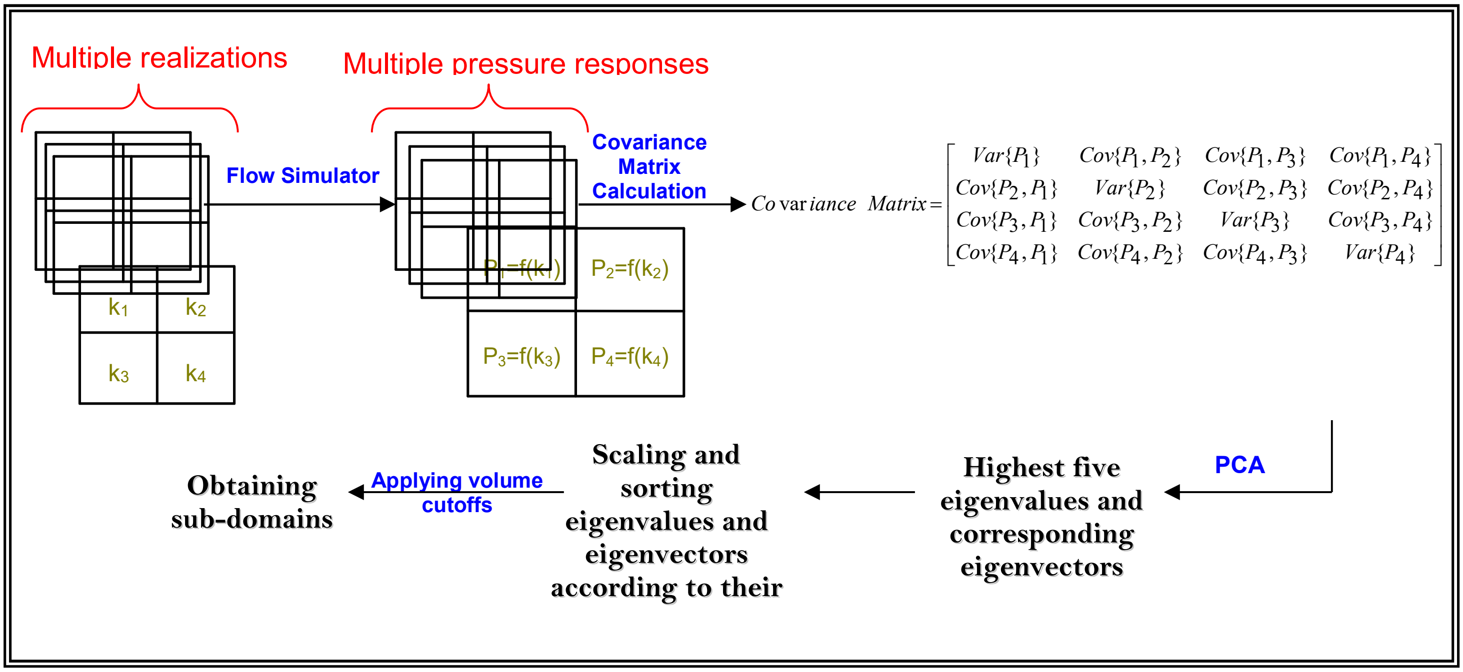


Figure 15. The schematic of sub-domain obtaining process

Case Examples Using Pro-HMS

Objective

A set of synthetic cases was devised in order to test the robustness of the Pro-HMS software. The sequential perturbation of the permeability probability distribution within the sub-domains is implemented to see the resultant reservoir permeability map for each step. For comparison, the results obtained without using sub-domains is also presented to demonstrate the effectiveness of the Pro-HMS Algorithm.

Conditions for test cases

Reference and initial models are generated based on the following conditions.

1. The reference reservoir models are generated by sequential indicator simulation. The dimension of the reservoir is 100 X 100 X 5, and each grid is 50ft X 50ft X 10ft. Three wells are located at grid locations (30, 30), (50, 50), and (70, 70). One injector is at (90, 90). Figure 1 illustrates slices of the reference permeability model.

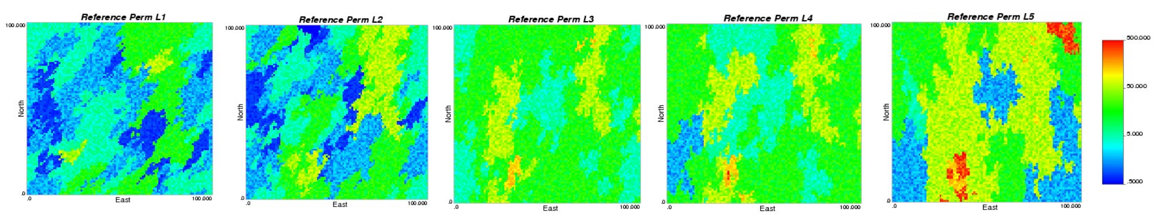


Figure 1. Reference permeability models from top layer to the bottom (from left to right)

2. Pro-HMS generates an ensemble of initial models in order to get sub-domains. Figure 2. shows slices from one of the initial models.

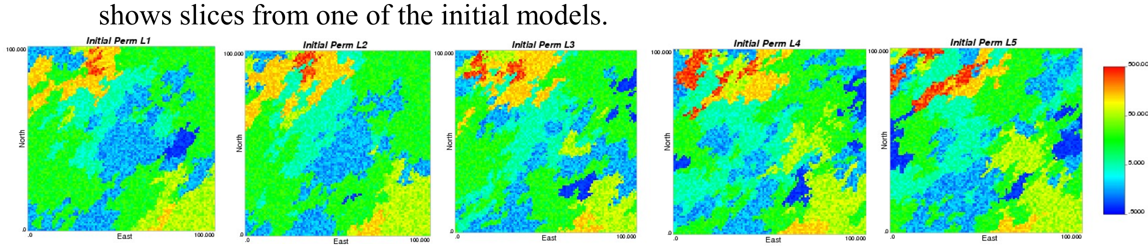


Figure 2. An initial permeability model from top layer to the bottom (from left to right)

1st Case: History matching with limited duration data & sub-domains

The first test is intended to demonstrate the history matching capability of Pro-HMS. The production data over a three year period was used for the history matching.

The sub-domains, which are the most influential yet least correlated regions, are obtained by covariance matrix calculation and principle component analysis (PCA). (See the explanation on ‘An Improved Domain Delineation Procedure’ Section) The sub-domains obtained for this case after 40% volume cutoffs are show in Figure 3.

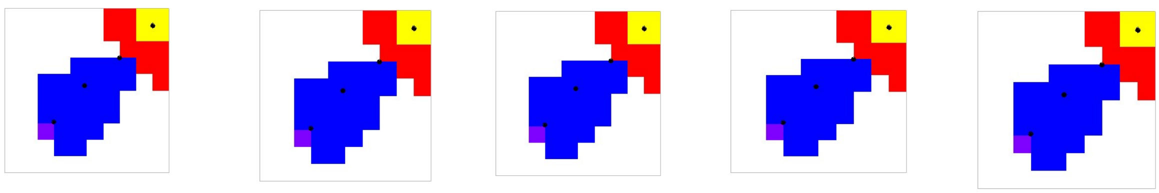


Figure 3. Sub-domains for tests after applying 40% volume cutoffs.

The final history matched permeability reservoir models are shown in Figure 4. Note that the reference model exhibits varying angles of anisotropy in the different layers. The limited duration of the production data matched, is insufficient for imparting the correct variations in permeability anisotropy to the different layers.

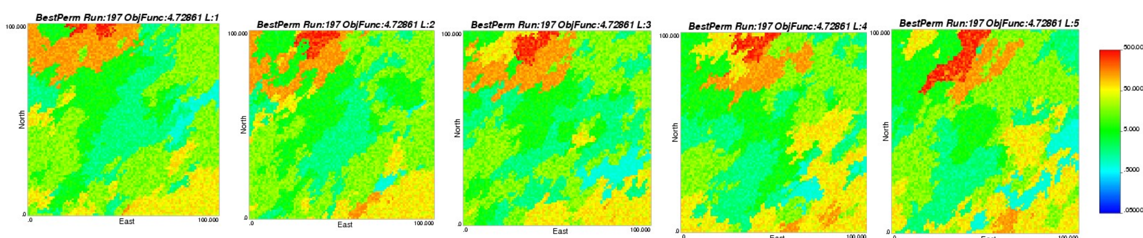


Figure 4. The permeability model after three years history matching

The results indicate that the representation of spatial heterogeneity has improved in the vicinity of the production wells. However, the permeability distributions in regions away from three wells do not resemble that observed in the reference models. Predictions of future performance for wells in those regions are therefore likely to be inaccurate. Figure 5 shows the matches to the production data observed at the wells. As is evident, the initial model exhibits production response that is distinctly different from the reference response. The final model matches the reference response closely.

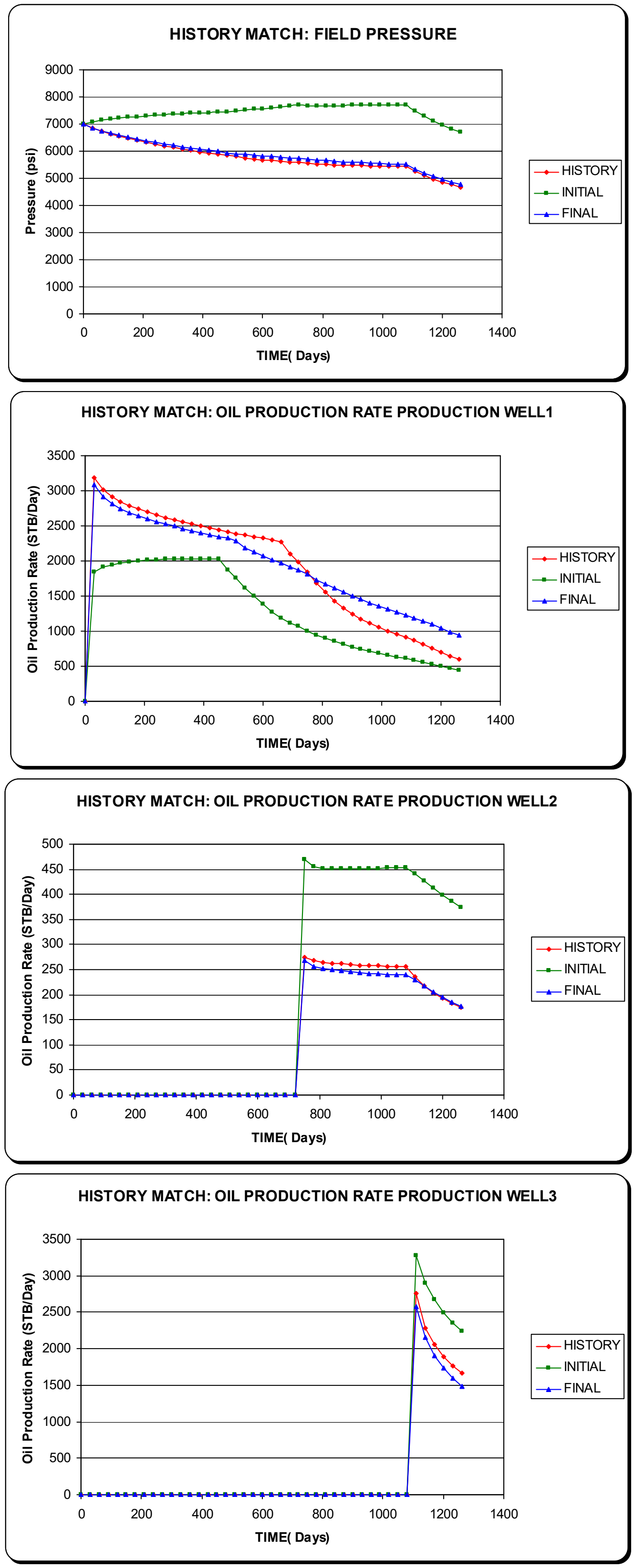


Figure 5. History matching result for test 1

The generated parameter files from Pro-HMS for this case are shown in Tables 1, 2, and 3.

Table 1. The 'hmissim.par' generated by Pro-HMS

'hsissim.par'

```
VariableType        1
NumThresholds/Categories      7
ThresholdV/CategoriesValues   0.55 0.9 2.5 7.55 26.85 84.2 157.95
PriorCDF/PDF        0.05 0.1 0.25 0.5 0.75 0.9 0.95
CondDataFile        Conditional.txt
CondFileDescrip     1 2 3 4
SoftDataFile        direct.ik
SoftFileDescrip     1 2 3 4 5 6 7 8 9 10
Markov-Bayes        0
Mark-BCalibBs       0.61 0.61 0.61 0.61 0.61 0.61 0.61
TrimmingLimits      -1.0e21 1.0e21
Min/MaxDataVal      0.005 500
LowerTailOpt        1 0.005
MiddleOption        1 1
UpperTailOpt        1 500
TabulatedValFile    cluster.dat
TabFileDescrip      4     0
DebuggingLevel      0
DebuggingFile       sisim.dbg
IncSimPermFile      PERM.DAT
ProdHistFile        ProdHistory.txt
GridDescripX        100 0.5 1
GridDescripY        100 0.5 1
GridDescripZ        5 0.5 1
RandomSeedVal       93984
MaxCondDataKrig     12
MaxSimDataKrig      12
MaxSoftDataKrig     1
AssignCondData      1
MultiGridOpt        0     3
NumDataPerOct       0
MaxSearchRadii      1000 1000 200
SearchAngles        341 0 0
Full/MedianK        0 50
Simple/Ordinary     1
InitRealizOpt       0
InitRealizFile      InitRealiz.dat
SubdomainsOpt
SubdomainFile       RegionIndices.txt
InnerLoopIter       3
OuterLoopIter       4
1-st Variogram      1 0
1 1 45 0 0
16 8 2
2-nd Variogram      1 0
1 1 45 0 0
16 8 2
3-rd Variogram      1 0
1 1 45 0 0
16 8 2
4-th Variogram      1 0
1 1 45 0 0
```

```
16 8 2
5-th Variogram    1 0
1 1 45 0 0
16 8 2
6-th Variogram    1 0
1 1 45 0 0
16 8 2
7-th Variogram    1 0
1 1 45 0 0
16 8 2
```

Table 2. The ‘sensregion.par’ generated by Pro-HMS

‘sensregion.par’

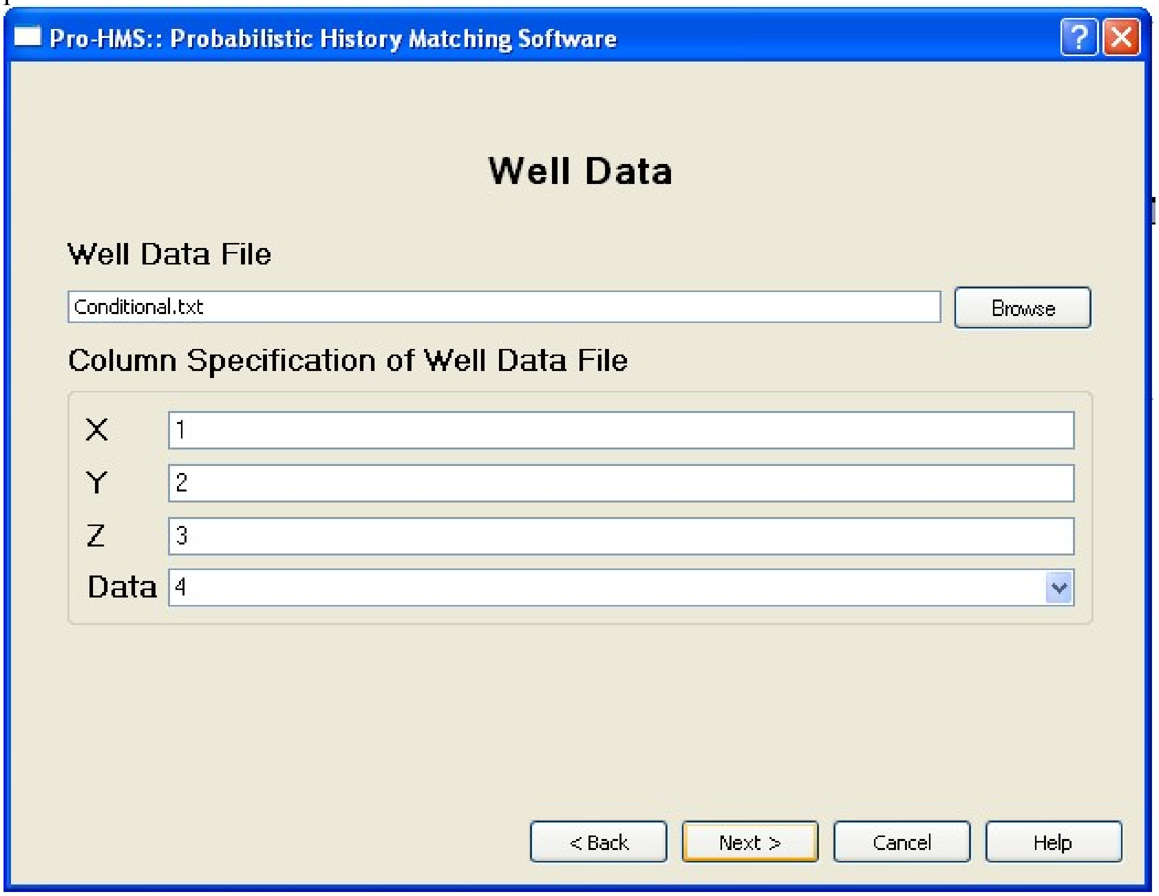
```
1                       - 1 = continuous(cdf), 0 = categorical(pdf)
7                             - number of thresholds/categories
0.55 0.9 2.5 7.55 26.85 84.2 157.95       - thresholds / categories
0.05 0.1 0.25 0.5 0.75 0.9 0.95           - global cdf / pdf
highcluster.dat                     - file with data
1 2 3 4                             - columns for X,Y,Z, and variable
direct.ik                           - file with soft indicator input
1 2 3 4 5 6 7 8 9 10                - columns for X,Y,Z, and indicators
0                             - Markov-Bayes simulation (0=no, 1=yes)
0.61 0.61 0.61 0.61 0.61 0.61 0.61          - calibration B(z) values
-1.0e21 1.0e21                              - trimming limits
0.005 500                           - minimum and maximum data values
1 0.005                             - lower tail option and parameter
1 1                                 - middle tail option and parameter
1 500                               - upper tail option and parameter
highcluster.dat                     - file with tabulated values
3 0                                 - columns for variable, weight
0                                   - debugging level: 0,1,2,3
sisim.dbg                     - debugging file for debugging output
sisim.out                           - file for simulation output
100 0.5 1                           - nx, xmn, xsiz
100 0.5 1                           - ny, ymn, ysiz
5 0.5 1                             - nz, zmn, zsiz
93984                               - random number esed
12                            - maximum original data for each krigging
12                            - maximum previous nodes for each krigging
1                       - maximum soft indicator nodes for krigging
1                             - assign data to nodes? (0=no, 1=yes)
0 3                     - multiple grid search? (0=no, 1=yes), num
0                                    -maximum per octant (0=not used)
1000 1000 200                       - maximum search radii
341 0 0                             - angles for search ellipsoid
51 51 11                            - size of covariance lookup table
0 50                          - 0=full IK, 1=median approx. (cutoff)
1                             - 0=SK, 1=OK
1 0                           -1 nst, nugget effect
1 1 45 0 0                    - it,cc,ang1,ang2,ang3
16 8 2                              - a_hmax, a_hmin, a_vert
1 0                                 -2 nst, nugget effect
1 1 45 0 0                    - it,cc,ang1,ang2,ang3
16 8 2                              - a_hmax, a_hmin, a_vert
1 0                                 -3 nst, nugget effect
```

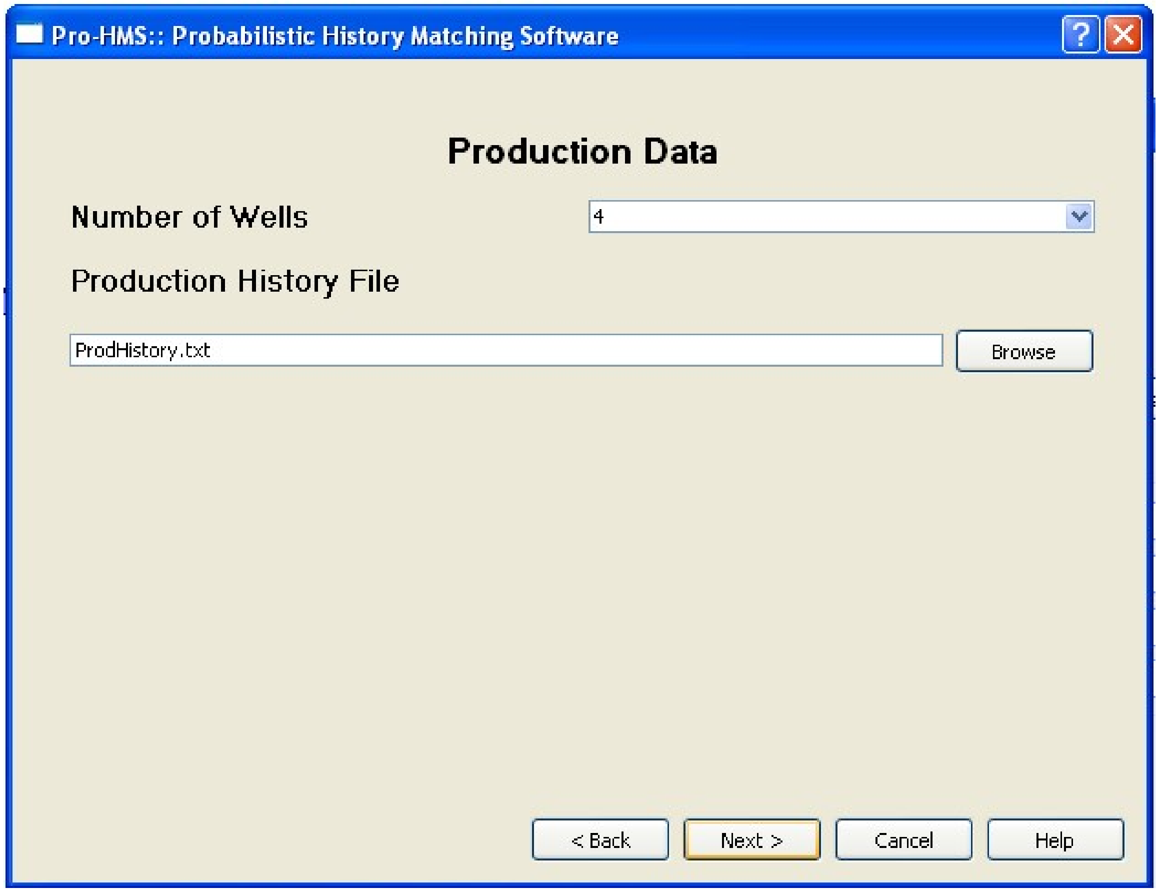
```
1 1 45 0 0                     - it,cc,ang1,ang2,ang3
16 8 2                               - a_hmax, a_hmin, a_vert
1 0                                  -4 nst, nugget effect
1 1 45 0 0                     - it,cc,ang1,ang2,ang3
16 8 2                               - a_hmax, a_hmin, a_vert
1 0                                  -5 nst, nugget effect
1 1 45 0 0                     - it,cc,ang1,ang2,ang3
16 8 2                               - a_hmax, a_hmin, a_vert
1 0                                  -6 nst, nugget effect
1 1 45 0 0                     - it,cc,ang1,ang2,ang3
16 8 2                               - a_hmax, a_hmin, a_vert
1 0                                  -7 nst, nugget effect
1 1 45 0 0                     - it,cc,ang1,ang2,ang3
16 8 2                               - a_hmax, a_hmin, a_vert
0.25 porosity2.out             - CONSTANT POROSITY VALUE FOR UPSCALING OR
PROSITYFILE
1                              - NUMBER OF SISIM OUT PUT MAPS THAT USER
WANTS TO HAVE USING PIXELPLT(FOR ALL LAYERS)
10 10 5                        - UPSCALED RESERVOIR GRID IN X,Y, AND Z
DIRECTION
data.dat                       - FLOW SIMULATOR FILE FOR GENERATING SUITE
OF FLOW RESPONSES CORRESPOND TO SUITE OF UPSCALED PERM
1260.00                        - SPECIFIC TIME AT WHICH THE PRESSURES ARE
CONSIDERED FOR COVARIANCE OR SENSITIVITY CALCULATION
50                             - NUMBER OF REALIZATIONS TO BE USED FOR
COVARIANCE CALCULATION
```

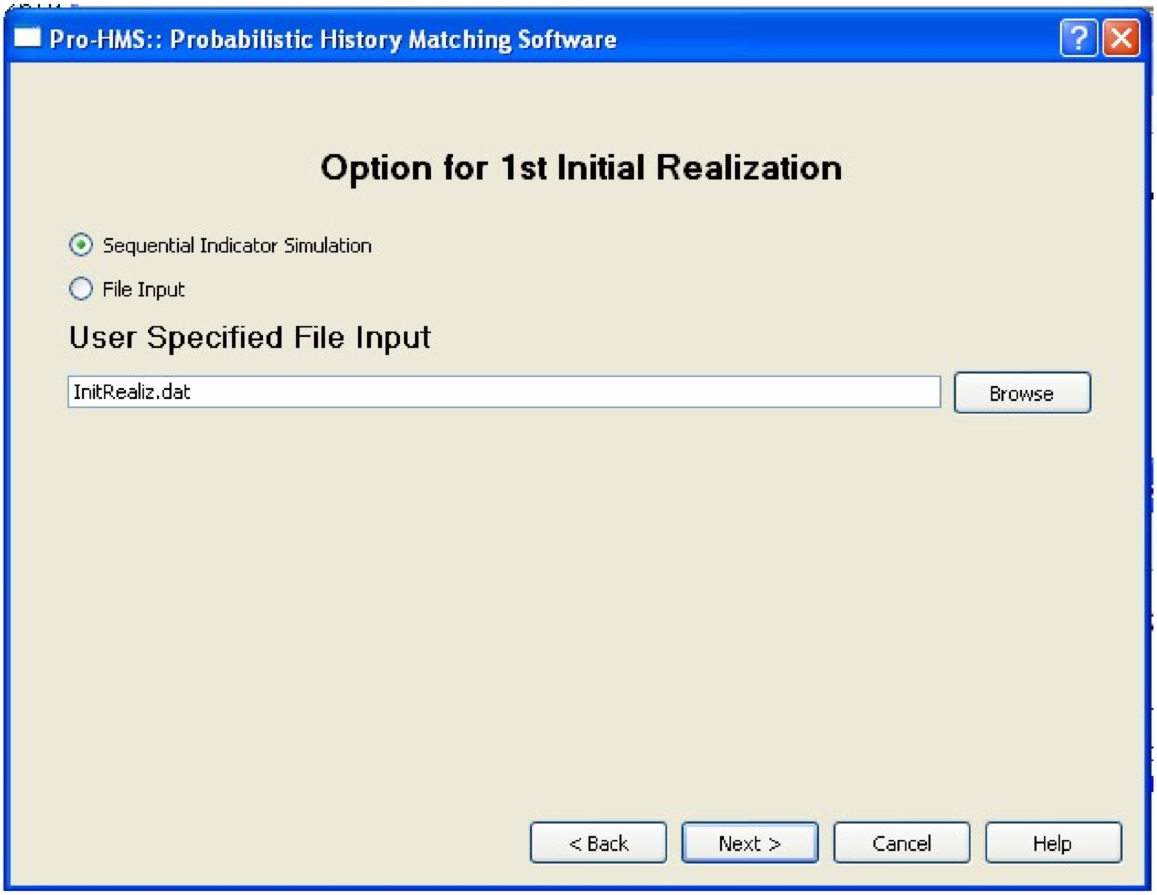
Table 3. The ‘dd.par’ generated by Pro-HMS

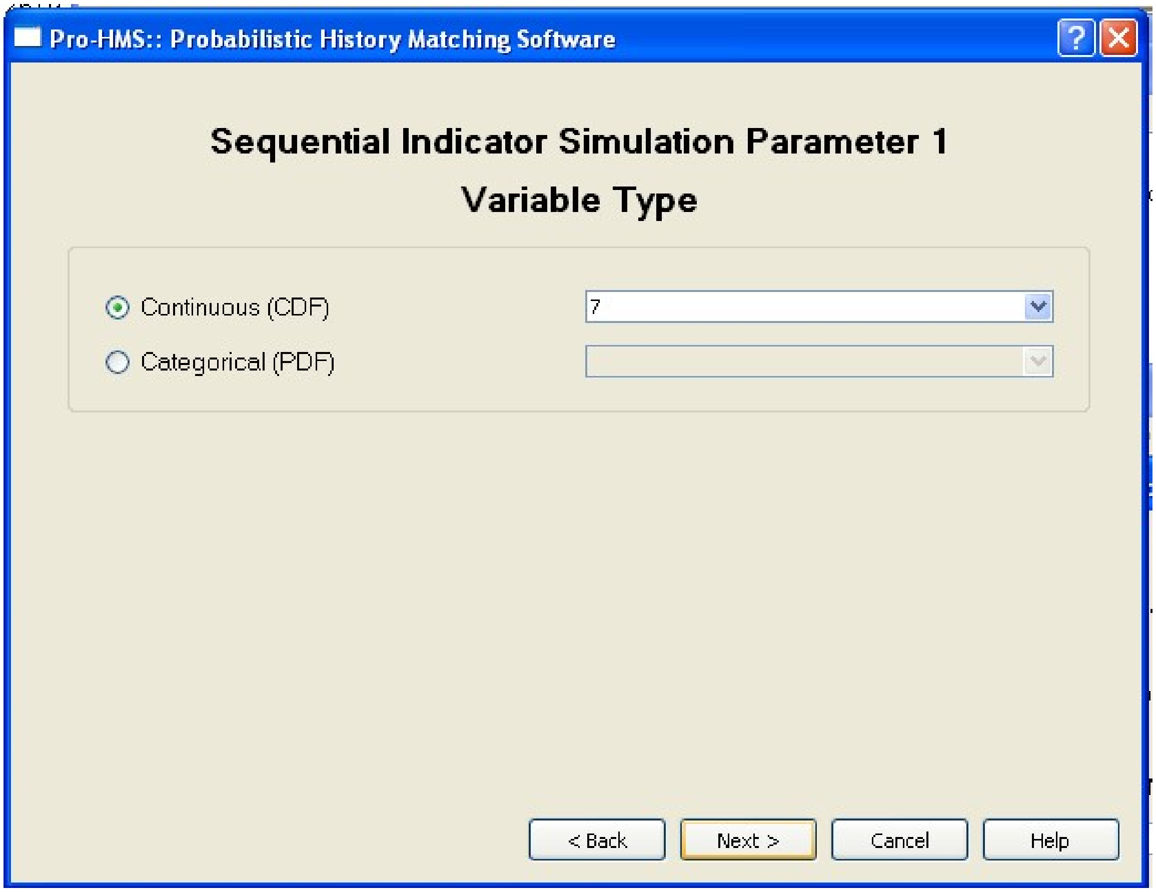
‘dd.par’
100
100
5
0.4

The following series of Pro-HMS screenshots are for this case. They show various input parameters.

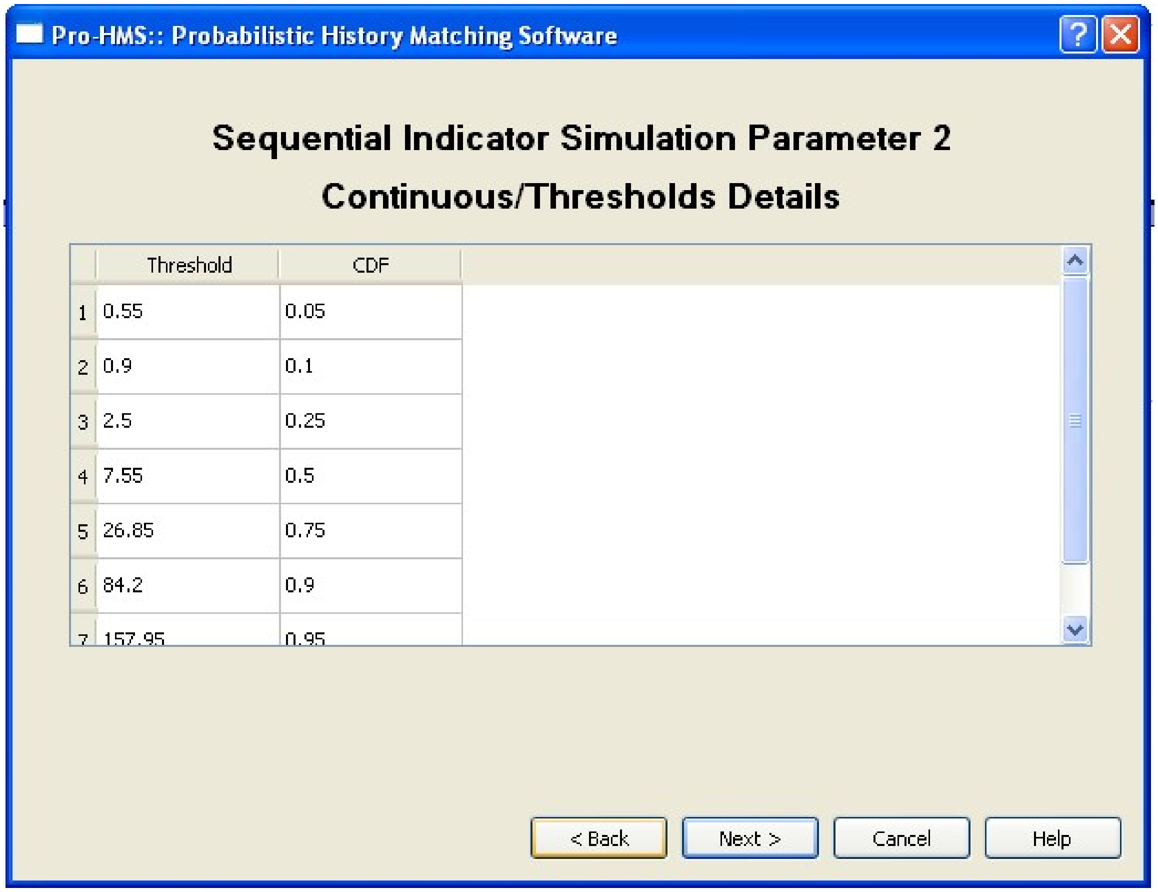


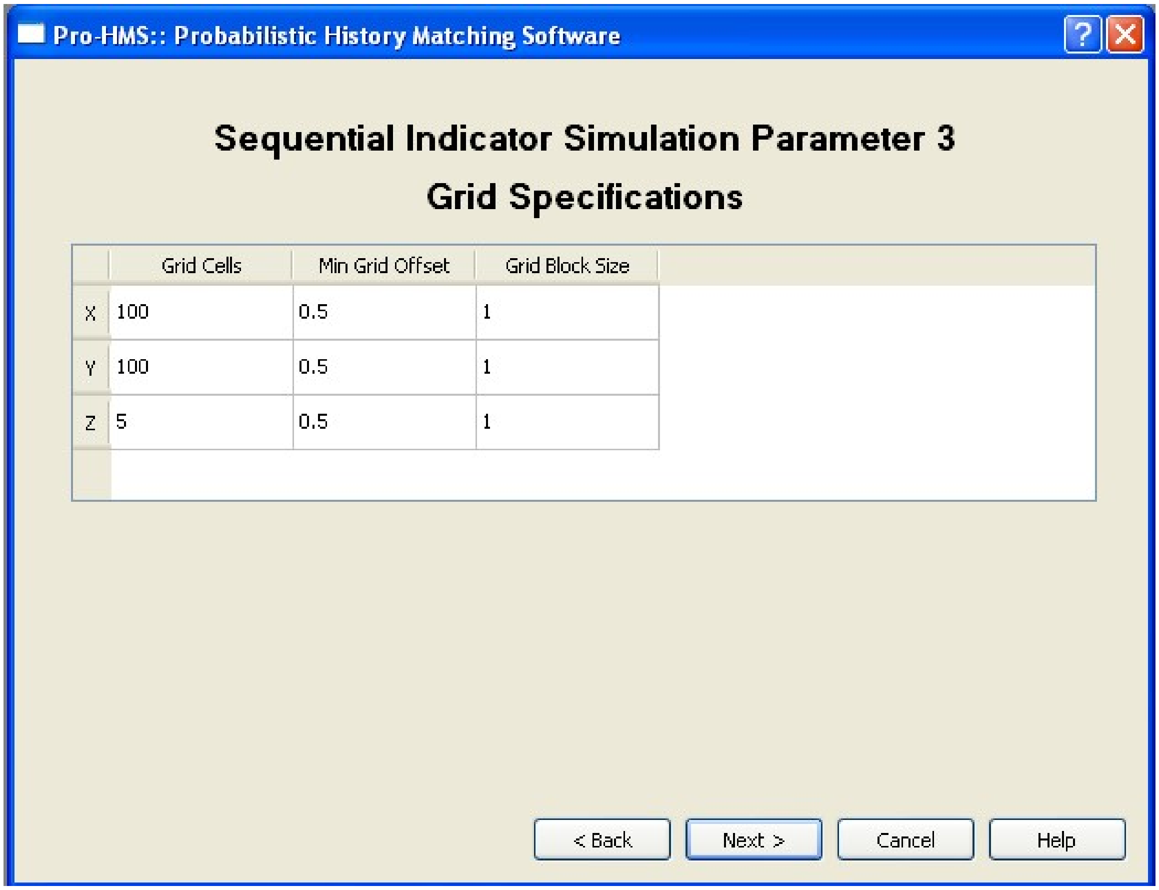

Pro-HMS:: Probabilistic History Matching Software
Production Data
Number of Wells
4
Production History File
ProdHistory.txt
Browse
< Back
Next >
Cancel
Help

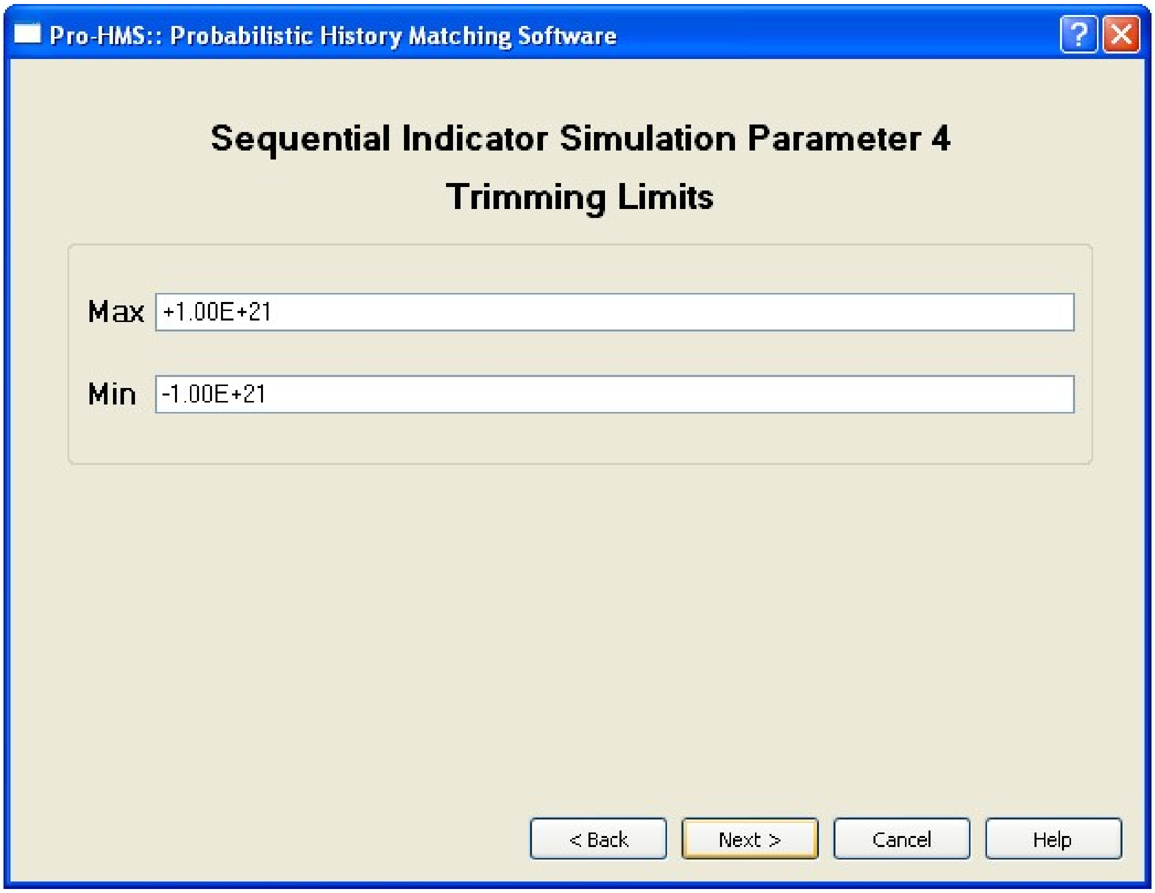

Pro-HMS:: Probabilistic History Matching Software
Option for 1st Initial Realization
Sequential Indicator Simulation
File Input
User Specified File Input
InitRealiz.dat
Browse
< Back
Next >
Cancel
Help

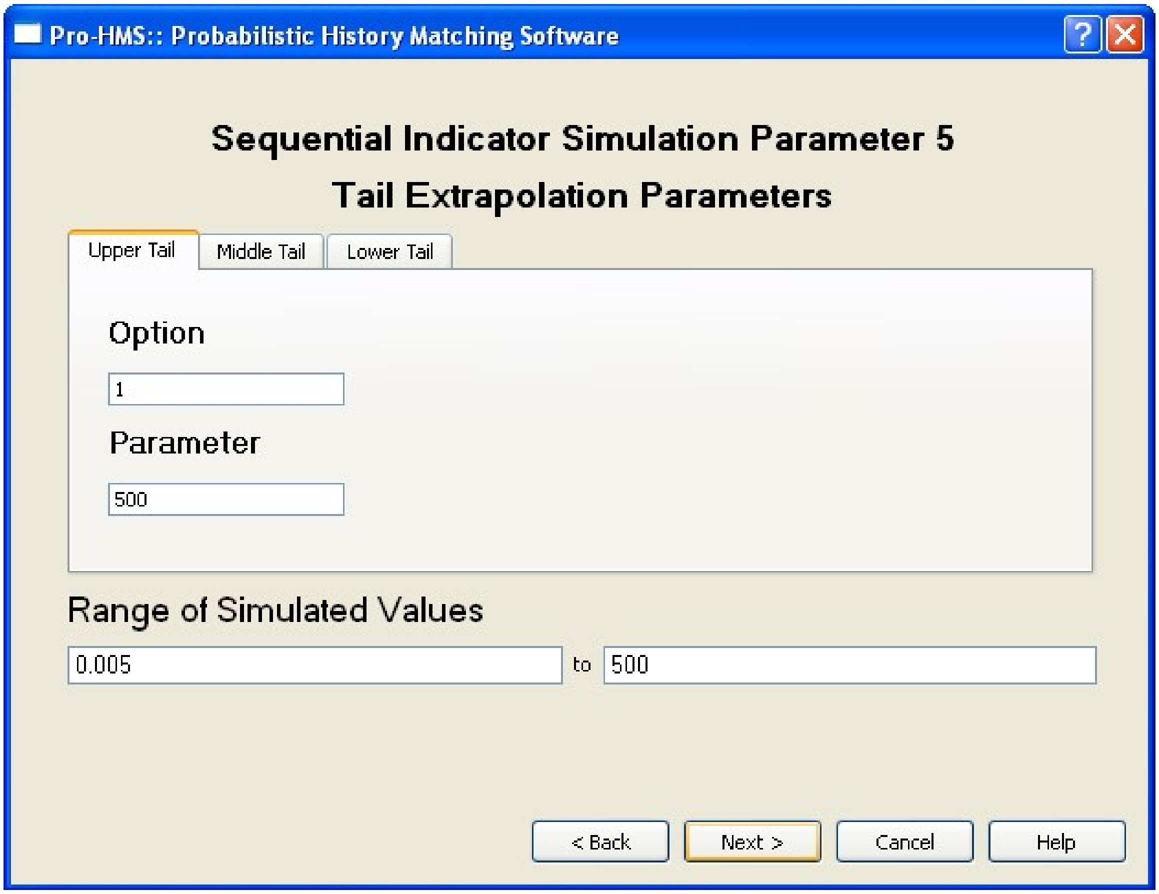


Pro-HMS:: Probabilistic History Matching Software
Sequential Indicator Simulation Parameter 1
Variable Type
Continuous (CDF)
7
Categorical (PDF)
< Back
Next >
Cancel
Help

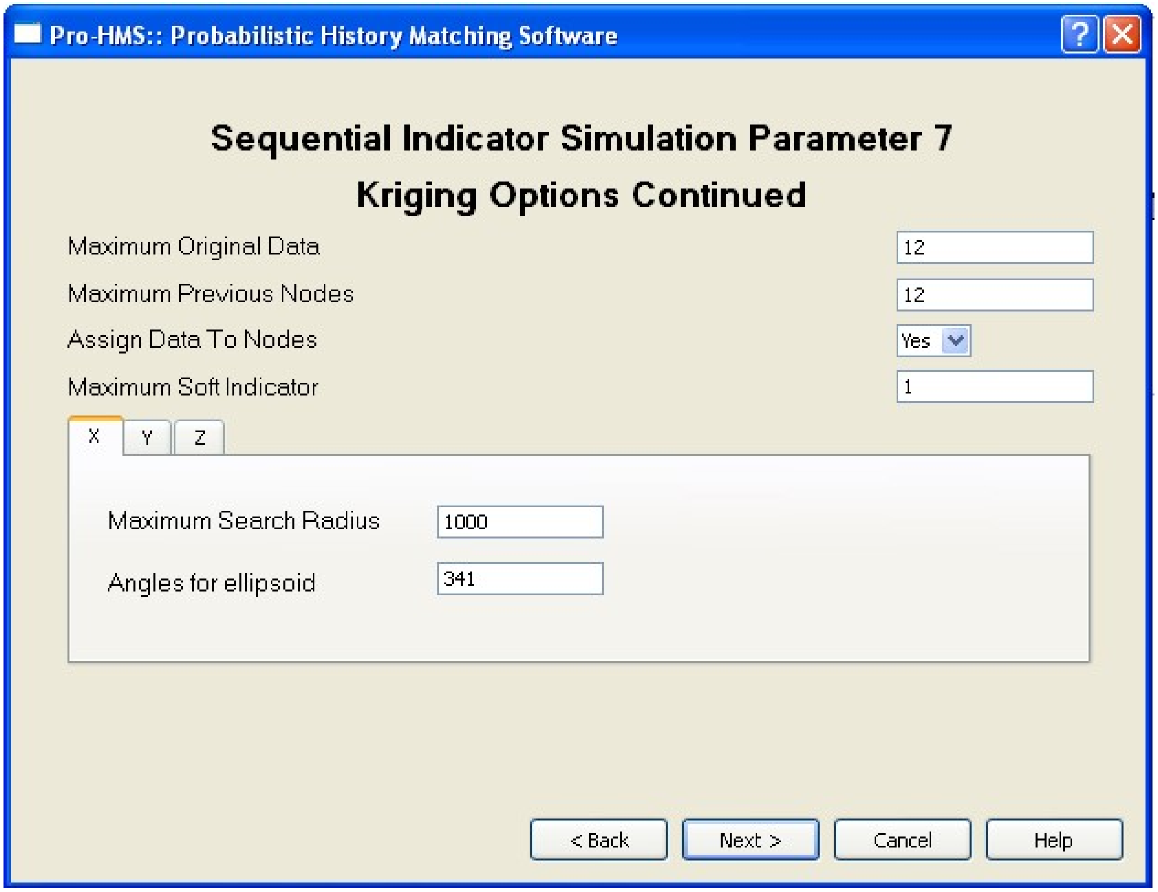

Pro-HMS:: Probabilistic History Matching Software
Sequential Indicator Simulation Parameter 2
Continuous/Thresholds Details
Threshold
CDF
1 0.55 0.05
2 0.9 0.1
3 2.5 0.25
4 7.55 0.5
5 26.85 0.75
6 84.2 0.9
7 157.95 0.95
< Back
Next >
Cancel
Help

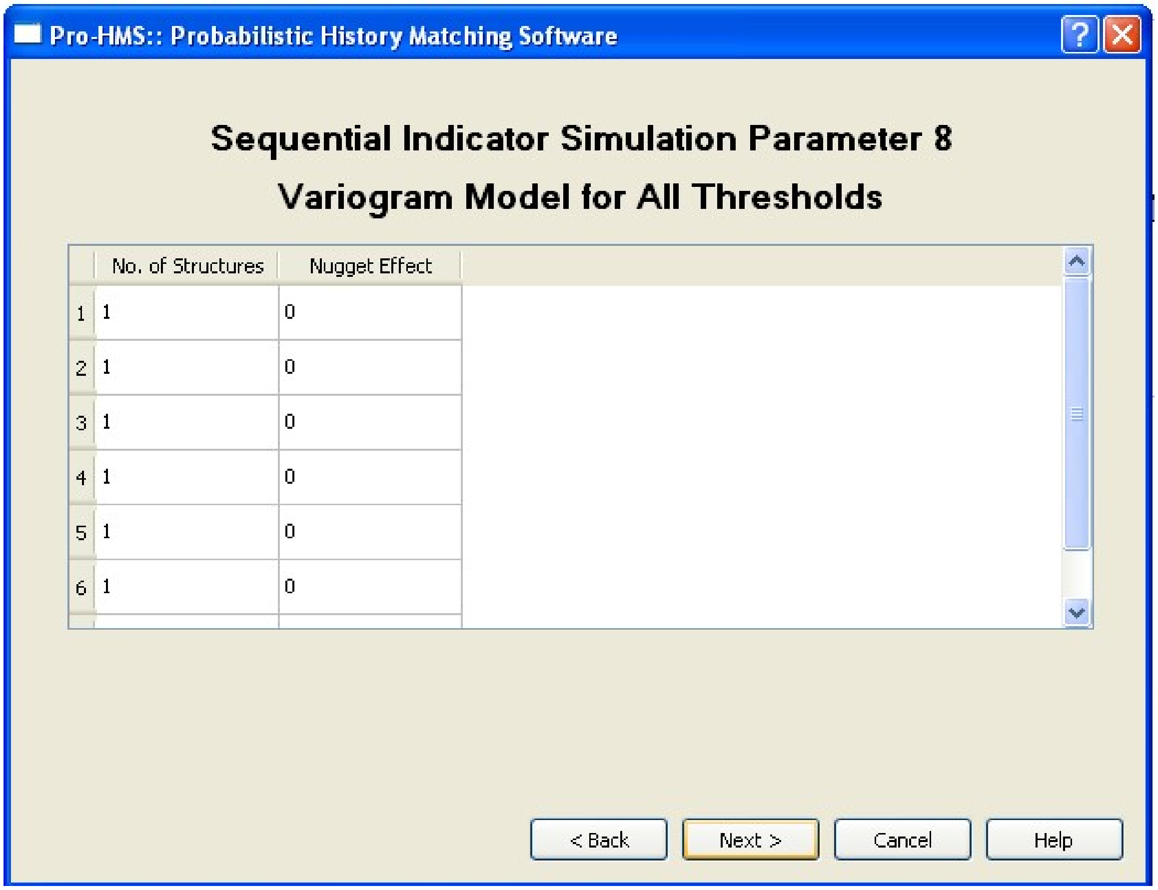

Pro-HMS:: Probabilistic History Matching Software
Sequential Indicator Simulation Parameter 3
Grid Specifications
Grid Cells
Min Grid Offset
Grid Block Size
X 100 0.5 1
Y 100 0.5 1
Z 5 0.5 1
< Back
Next >
Cancel
Help

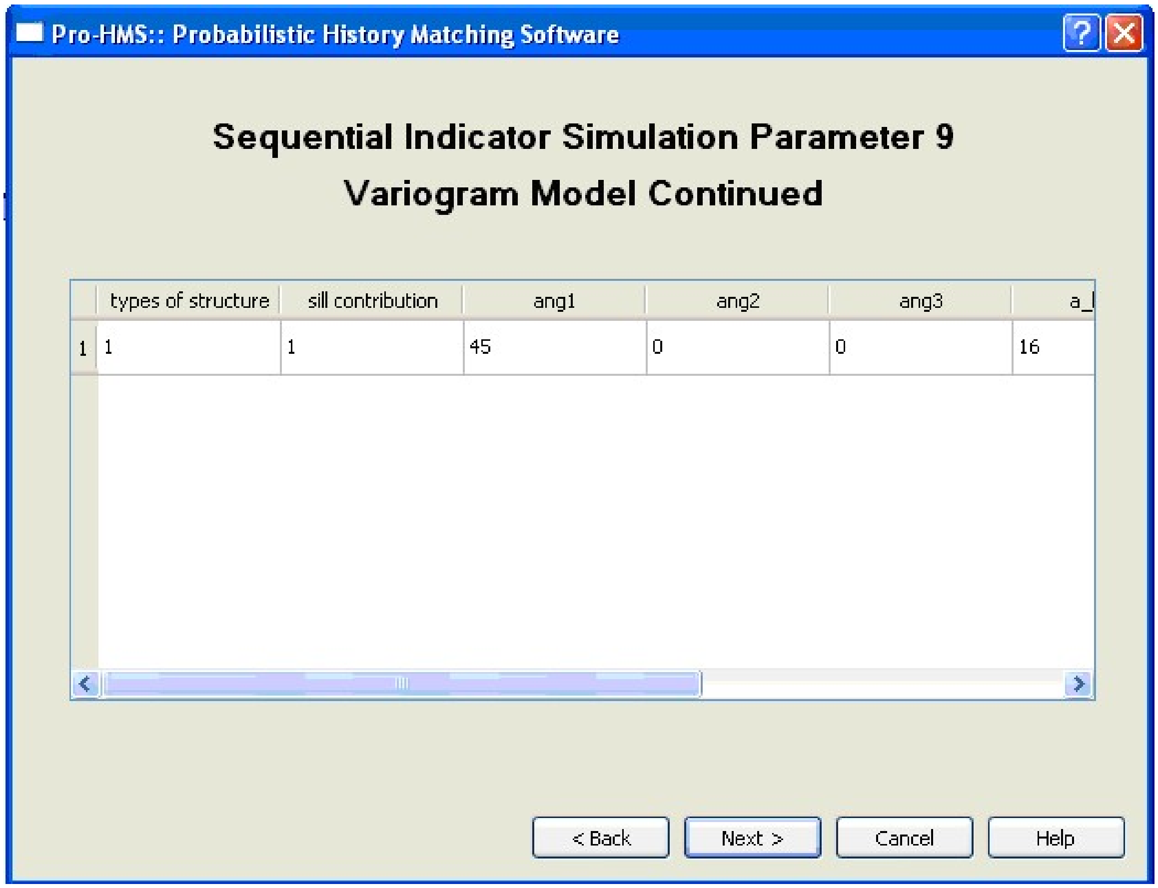

Pro-HMS:: Probabilistic History Matching Software
Sequential Indicator Simulation Parameter 4
Trimming Limits
Max
+1.00E+21
Min
-1.00E+21
< Back
Next >
Cancel
Help

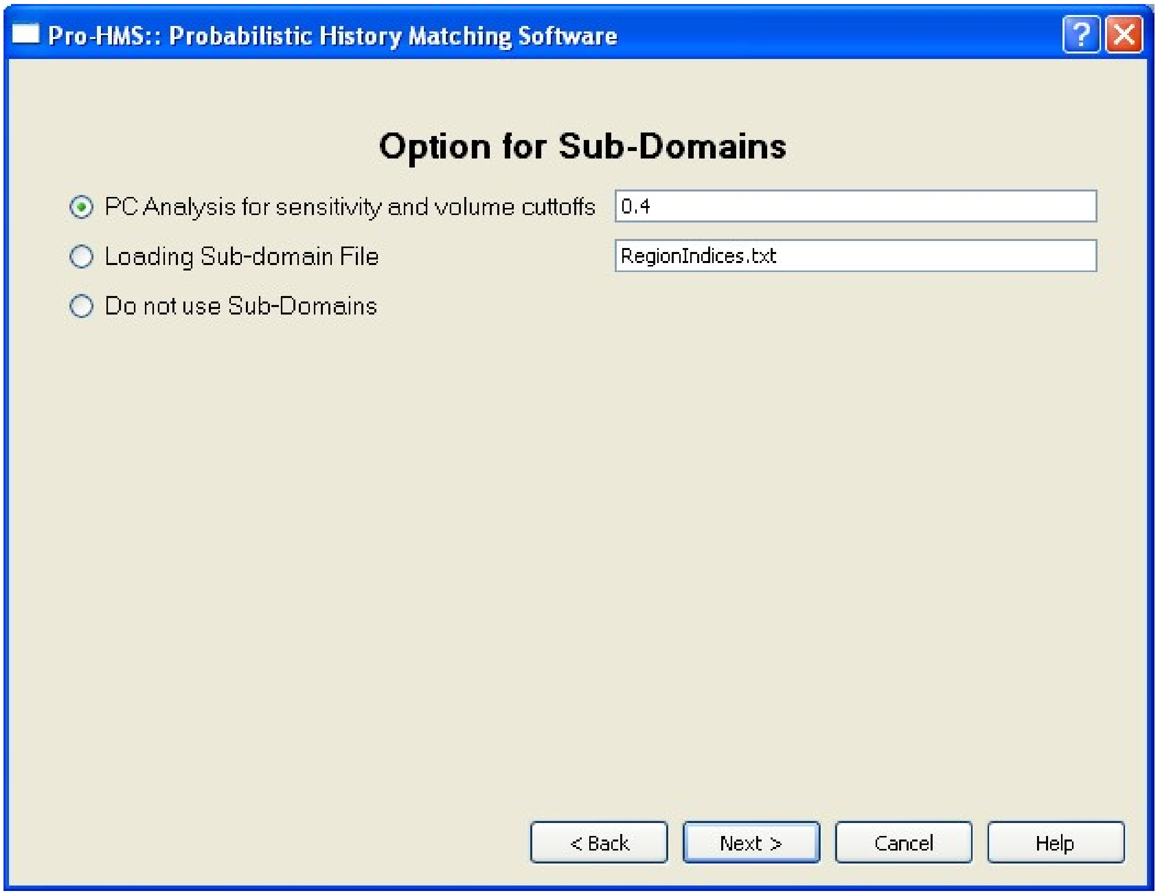

Pro-HMS:: Probabilistic History Matching Software
Sequential Indicator Simulation Parameter 5
Tail Extrapolation Parameters
Upper Tail
Middle Tail
Lower Tail
Option
1
Parameter
500
Range of Simulated Values
0.005
to
500
< Back
Next >
Cancel
Help

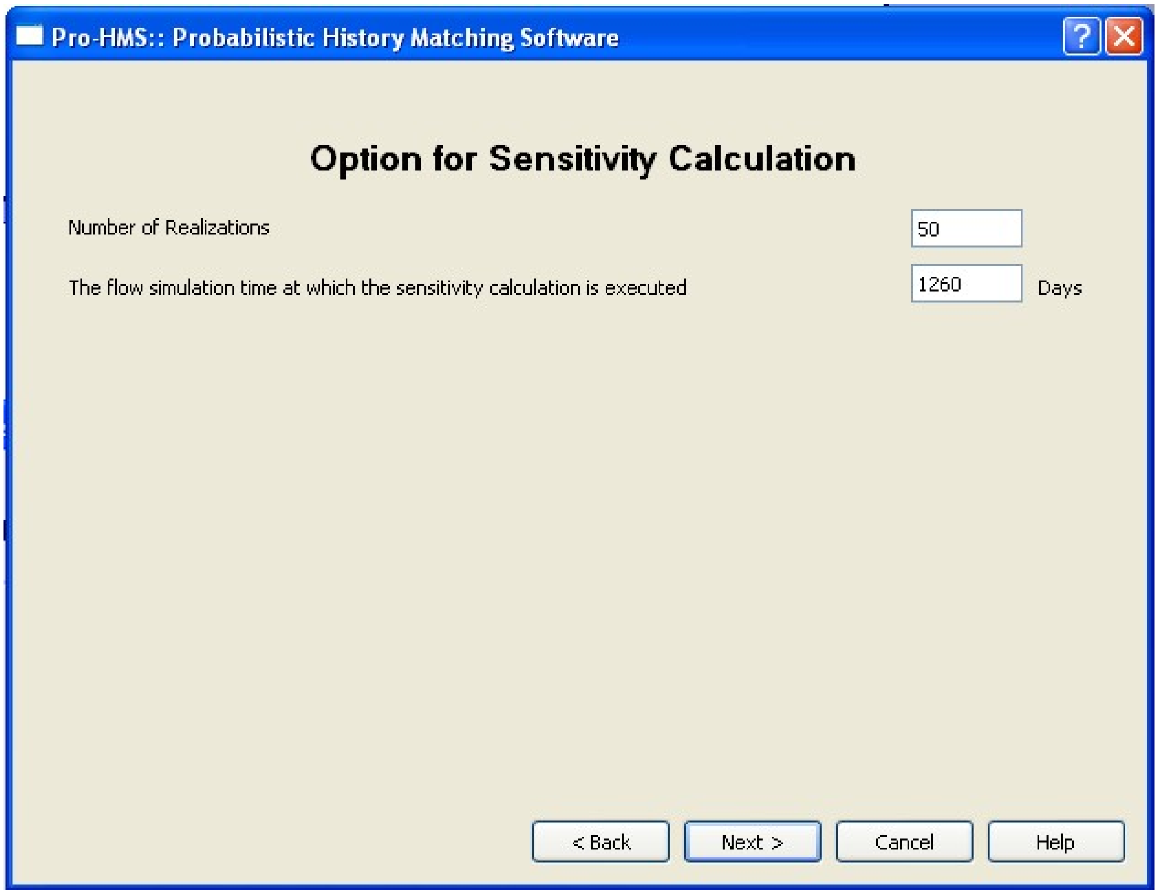

Pro-HMS:: Probabilistic History Matching Software
Sequential Indicator Simulation Parameter 6
Kriging Options
Indicator Kriging
Full
Median
Types of Kriging
Ordinary Kriging
Simple Kriging
Mean Value
2.5
< Back
Next >
Cancel
Help

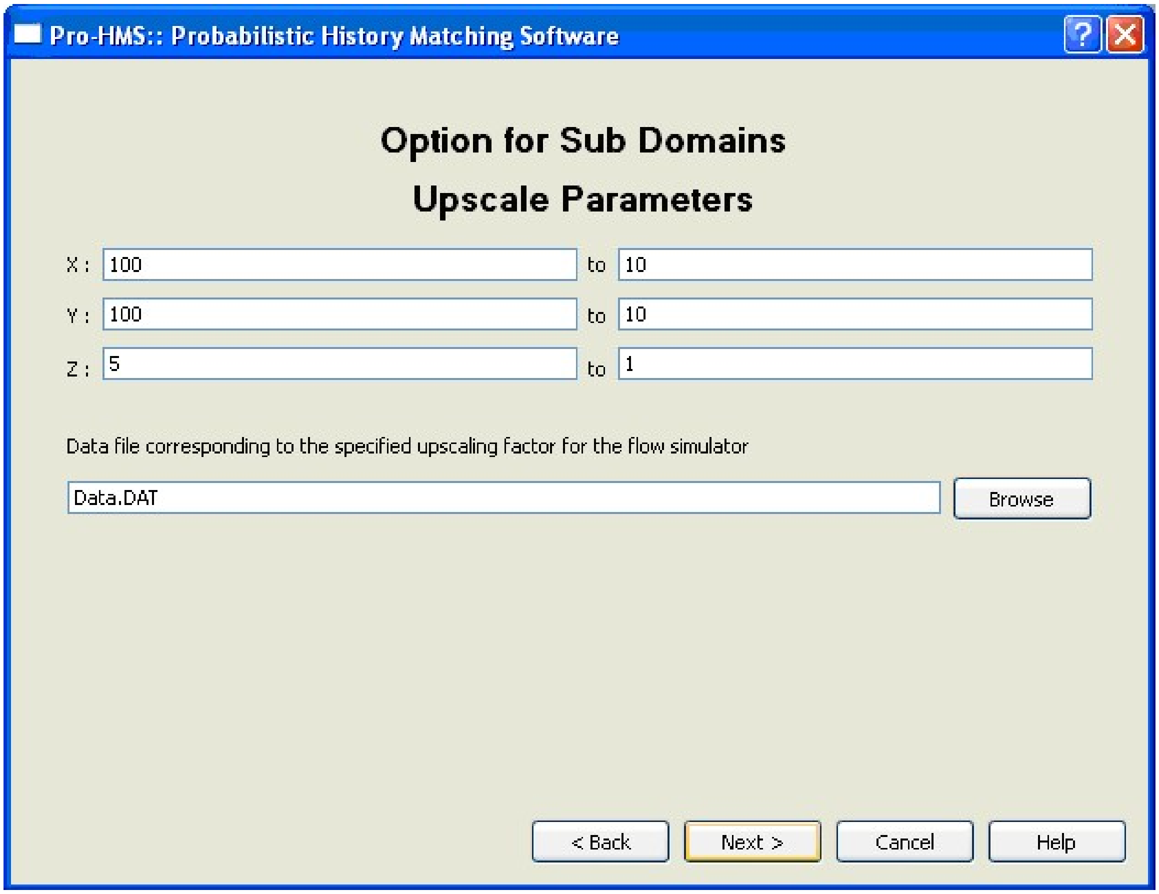

Pro-HMS:: Probabilistic History Matching Software
Sequential Indicator Simulation Parameter 7
Kriging Options Continued
Maximum Original Data
12
Maximum Previous Nodes
12
Assign Data To Nodes
Yes
Maximum Soft Indicator
1
X
Y
Z
Maximum Search Radius
1000
Angles for ellipsoid
341
< Back
Next >
Cancel
Help

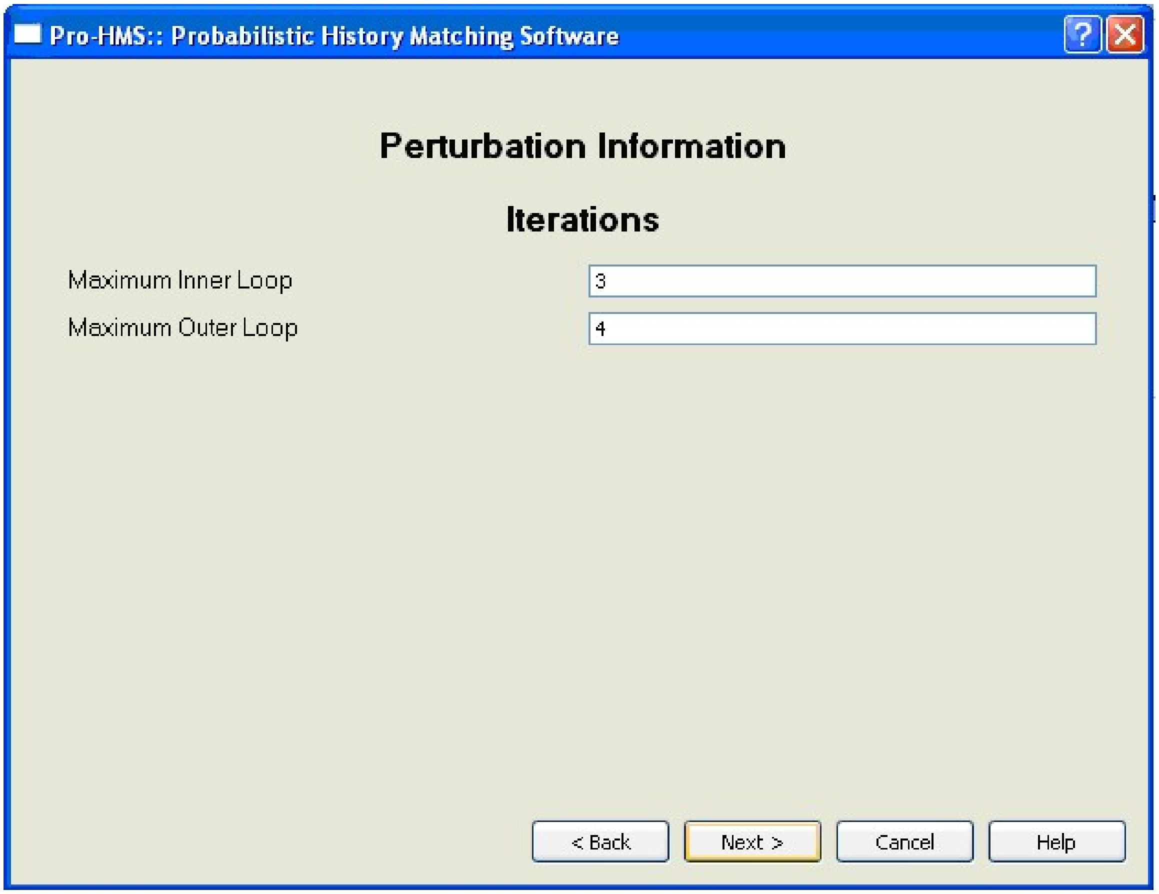

Pro-HMS:: Probabilistic History Matching Software
Sequential Indicator Simulation Parameter 8
Variogram Model for All Thresholds
No. of Structures
Nugget Effect
1 1 0
2 1 0
3 1 0
4 1 0
5 1 0
6 1 0
< Back
Next >
Cancel
Help


Pro-HMS:: Probabilistic History Matching Software
Sequential Indicator Simulation Parameter 9
Variogram Model Continued
types of structure
sill contribution
ang1
ang2
ang3
a_
1
1
1
45
0
0
16
< Back
Next >
Cancel
Help


Pro-HMS:: Probabilistic History Matching Software
Option for Sub-Domains
PC Analysis for sensitivity and volume cuttoffs
0.4
Loading Sub-domain File
RegionIndices.txt
Do not use Sub-Domains
< Back
Next >
Cancel
Help


Pro-HMS:: Probabilistic History Matching Software
Option for Sensitivity Calculation
Number of Realizations
50
The flow simulation time at which the sensitivity calculation is executed
1260
Days
< Back
Next >
Cancel
Help


Pro-HMS:: Probabilistic History Matching Software
Option for Sub Domains
Upscale Parameters
X : 100 to 10
Y : 100 to 10
Z : 5 to 1
Data file corresponding to the specified upscaling factor for the flow simulator
Data.DAT
Browse
< Back
Next >
Cancel
Help


Pro-HMS:: Probabilistic History Matching Software
Perturbation Information
Iterations
Maximum Inner Loop
3
Maximum Outer Loop
4
< Back
Next >
Cancel
Help

2nd Case: History matching using limited duration data without sub-domains

The importance of domain delineation for developing robust history-matched reservoir models is demonstrated in this second case. The perturbations of local conditional distributions are performed in all grid blocks. The result is shown in Figure 6. Compared to Figure 4, the permeability values far away from wells are also perturbed. This results in increased computational cost and also poor convergence to the target reservoir responses as evident in Figure 7.

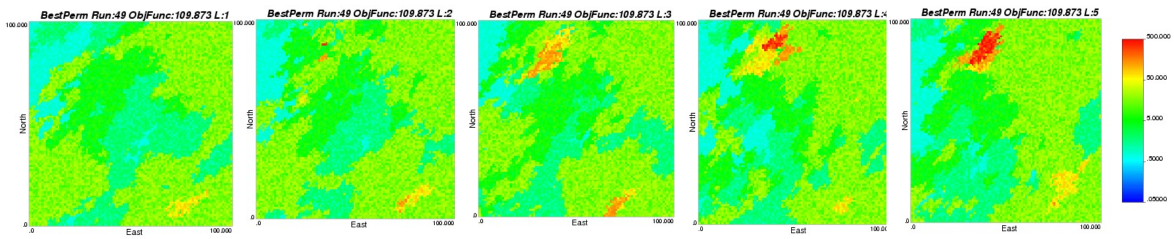
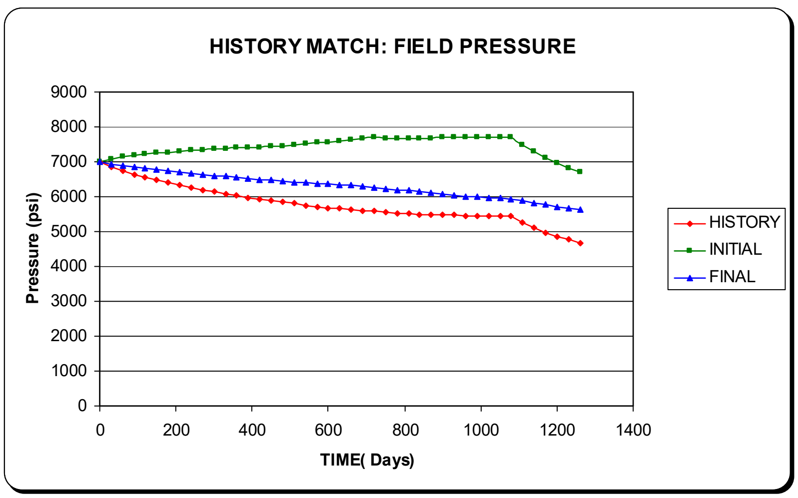
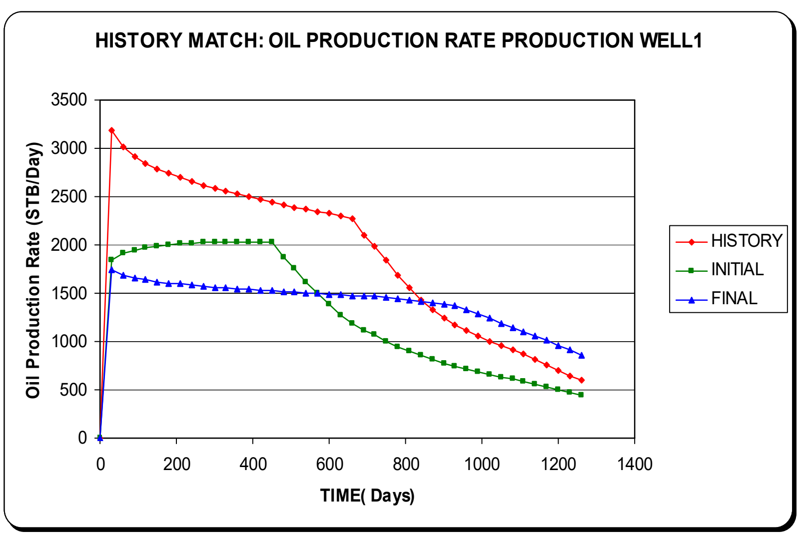


Figure 6. The permeability model after three years history matching without sub-domains





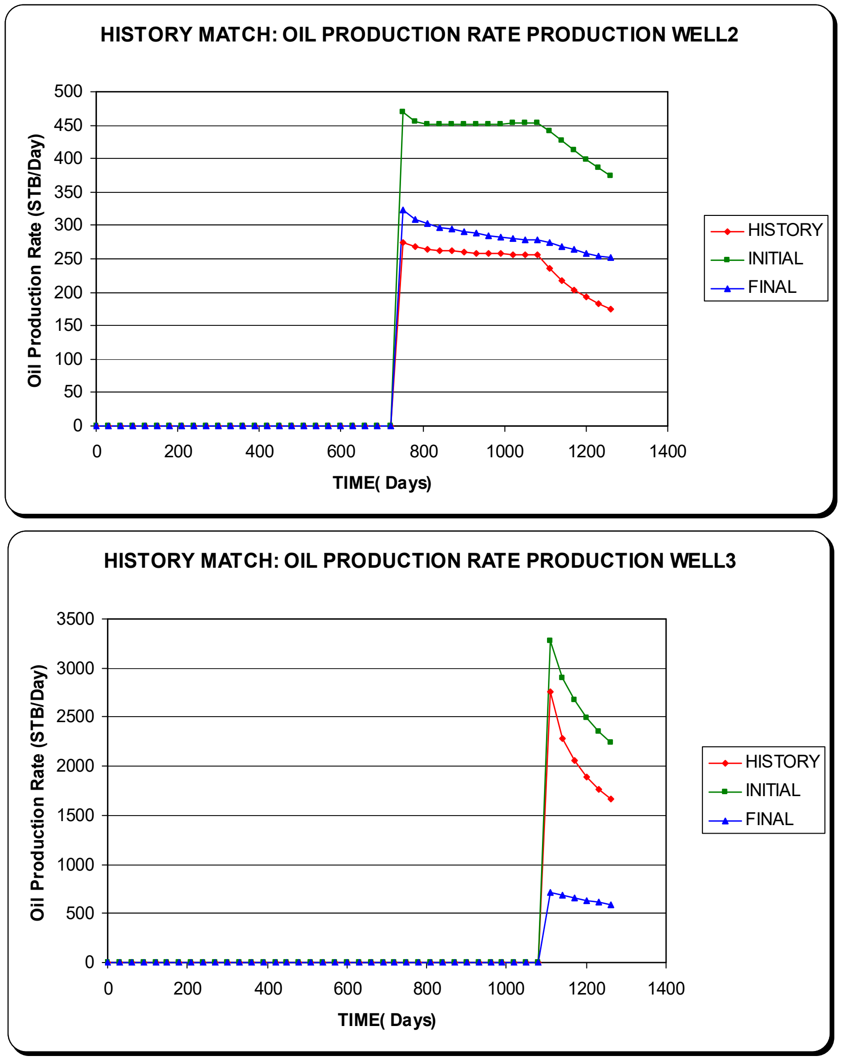


Figure 7. History match results when the permeability model is perturbed without using sub-domains

Conditions for Further Test Cases

It was observed in Figure 4 that though the production data is matched well by the updated models, the spatial characteristics of the permeability field in regions far away from wells are not consistent with the reference model. This can result in poor predictions of the reservoir performance in the future. To improve the characteristics of the reservoir models in regions away from existing wells and to test the efficacy of the software to handle addition of new wells on future dates, the following cases were attempted.

These further cases have not been done with Pro-HMS itself, but it has been tested with customized codes of Pro-HMS. The only change made in the codes for these cases is to alter the definition of conditioning data for simulating further nodes. These cases have been demonstrated in order to see the robustness of the Pro-HMS. Since all the other procedures are the same but the part of conditioning data, it is possible to test the effectiveness of Pro-HMS with these further cases.

Based on the result of 3 years history matching reservoir models, two more wells are drilled at locations (40, 90) and (90, 10). After the sensitivity test, the sub-domains are obtained and they are shown in Figure 8.

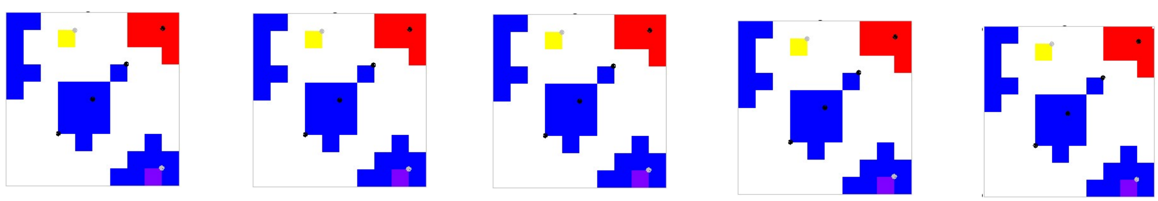


Figure 8. Sub-domains for the expanded case with two additional wells after applying 40% volume cutoffs.

3rd case: Expanded case with 2 additional wells and 3 years of production data

In this case, the permeability values within the two new sub-domains are perturbed conditioned to the previously perturbed regions in case 1. The local conditional distributions at locations inside the two new sensitivity regions are obtained conditioned to the updated permeability values in the sensitivity regions identified for case 1.

The result after the updating process is shown in Figure 9. This result is compared with the one obtained from 1st case, and then it can be concluded that the new sub-domains area are perturbed such that the high permeability values observed in Figure 4 are reduced to low permeability values. This is to be expected since the updated permeability values in the earlier sensitivity regions (case 1) are all medium valued (corresponding to the green regions in the map) and that data conditions the probability distributions in the two new sensitivity regions. These smooth updated reservoir models lead to poor reproduction of the historic production data as evident from Figure 10.

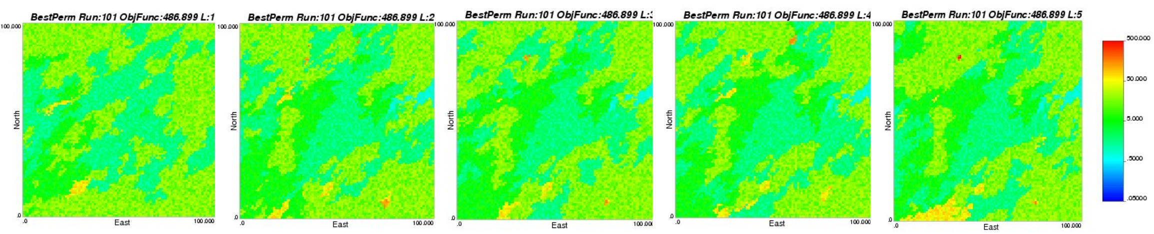
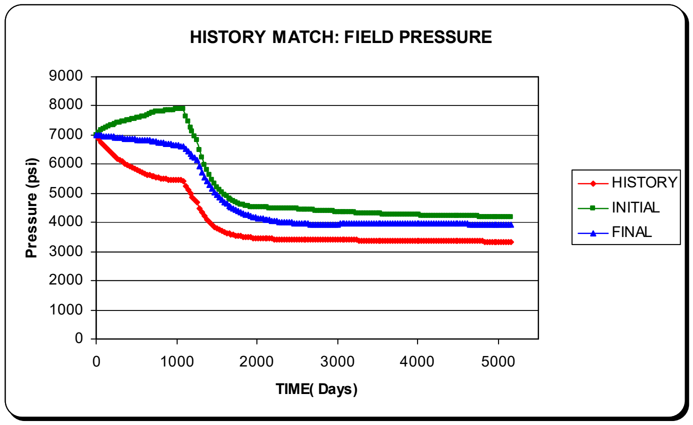
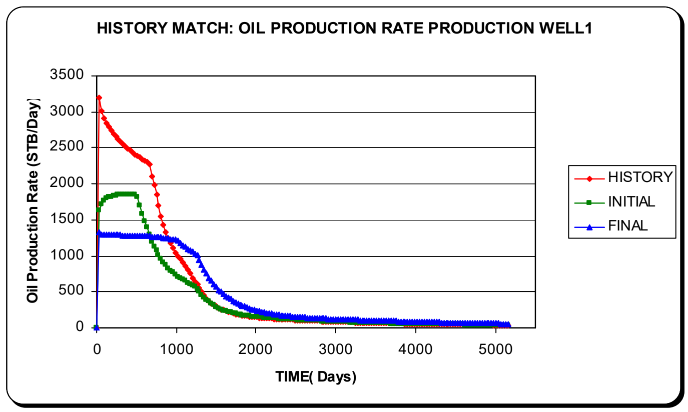
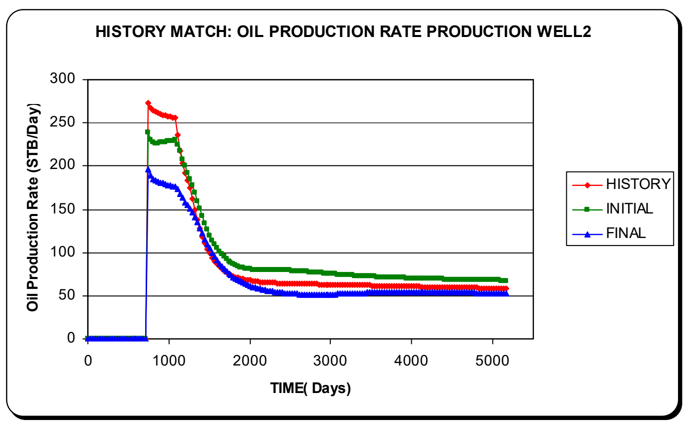
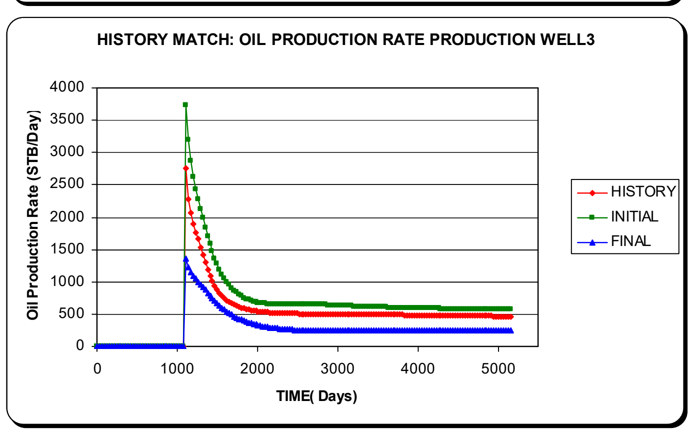


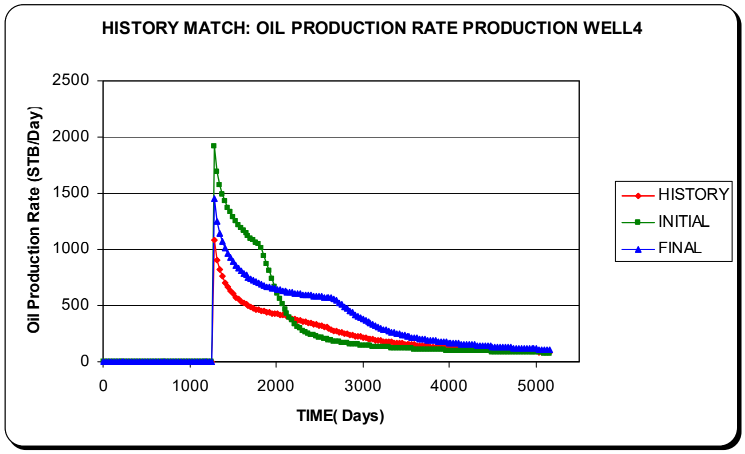
Figure 9. The permeability models after updating the permeability values in the sensitivity regions corresponding to the two new wells.


HISTORY MATCH: FIELD PRESSURE
Pressure (psi)
TIME(Days)
HISTORY
INITIAL
FINAL


HISTORY MATCH: OIL PRODUCTION RATE PRODUCTION WELL1
Oil Production Rate (STB/Day)
TIME(Days)
HISTORY
INITIAL
FINAL


HISTORY MATCH: OIL PRODUCTION RATE PRODUCTION WELL2
Oil Production Rate (STB/Day)
TIME(Days)
HISTORY
INITIAL
FINAL


HISTORY MATCH: OIL PRODUCTION RATE PRODUCTION WELL3
Oil Production Rate (STB/Day)
TIME(Days)
HISTORY
INITIAL
FINAL



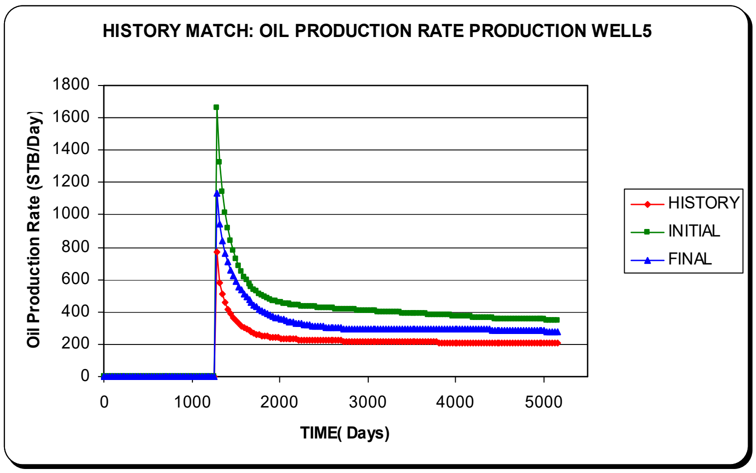


Figure 10. History matching results corresponding to the update of permeability values in the sensitivity regions corresponding to the two new wells. The updated values in the domains corresponding to the earlier configuration of wells (Figure 3) are used as conditioning data for updating the values in the two new sensitivity regions.

4th case: Updating of values in new sub-domains using only original "hard" data

In this case, the conditional probability distributions at locations within the two new sub-domains are only obtained using only the original "hard" data. The sequential indicator simulated permeability values in the less-sensitive regions are however obtained conditional to the values in the updated sub-domains.

The main feature of this case is that the computational time is tremendously reduced compared to the 3rd case. The updated reservoir models are shown in Figure 11. The permeability values near new wells are not similar to those for the 3rd case. The permeability values are higher. This can be attributed to the conditioning of values in the new sub-domains to the original data histogram that is not affected by the distribution of values within the previous sub-domains.

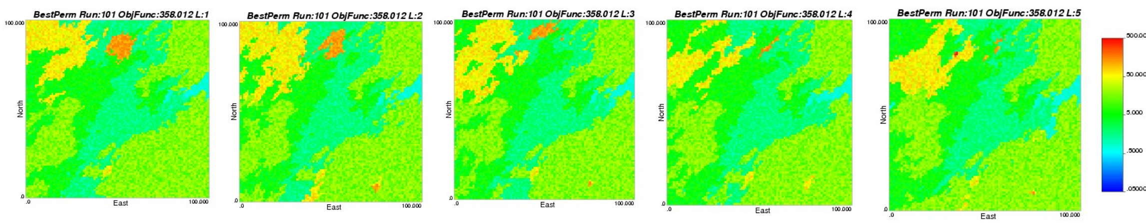
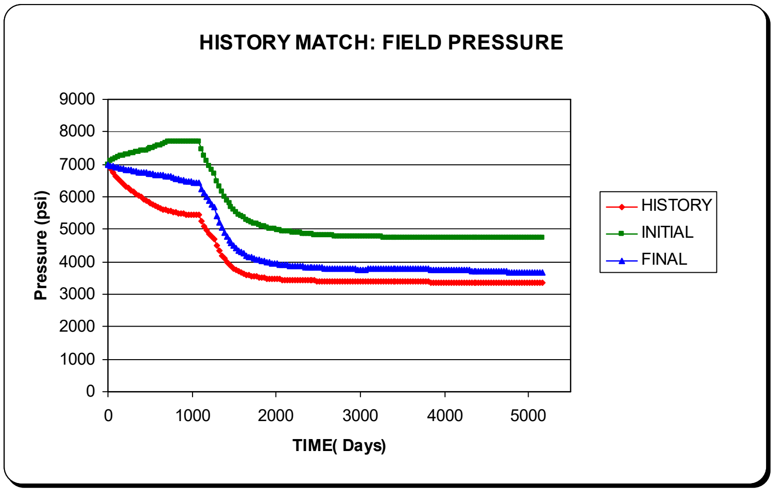
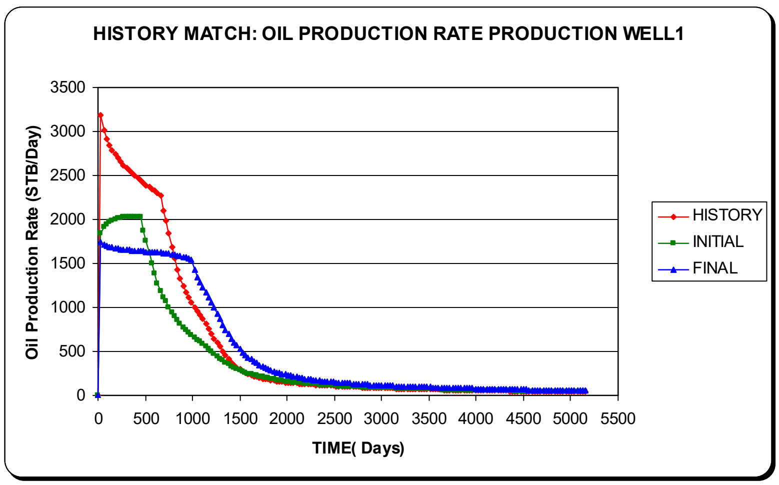
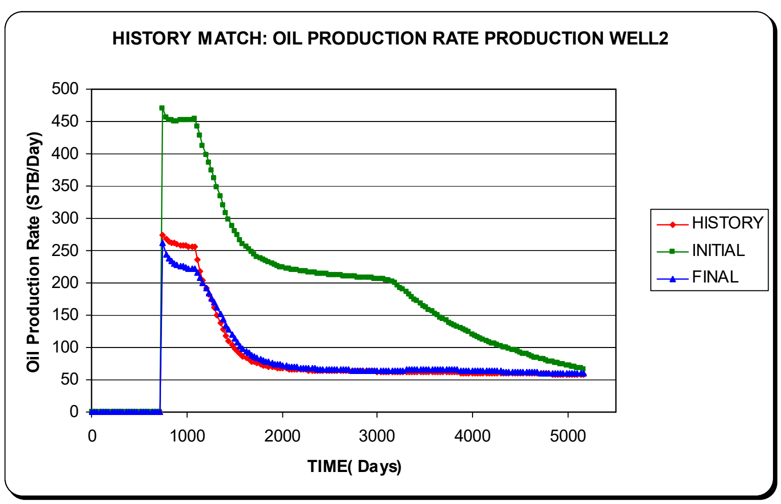


Figure 11. The permeability models after history matching using only the original "hard" data for updating the probability distributions within the new sub-domains and using the 19950 simulated values for conditioning the simulated values in the insensitive regions.

Figure 12 below indicates that much better reproduction of the production history information is possible using this approach.







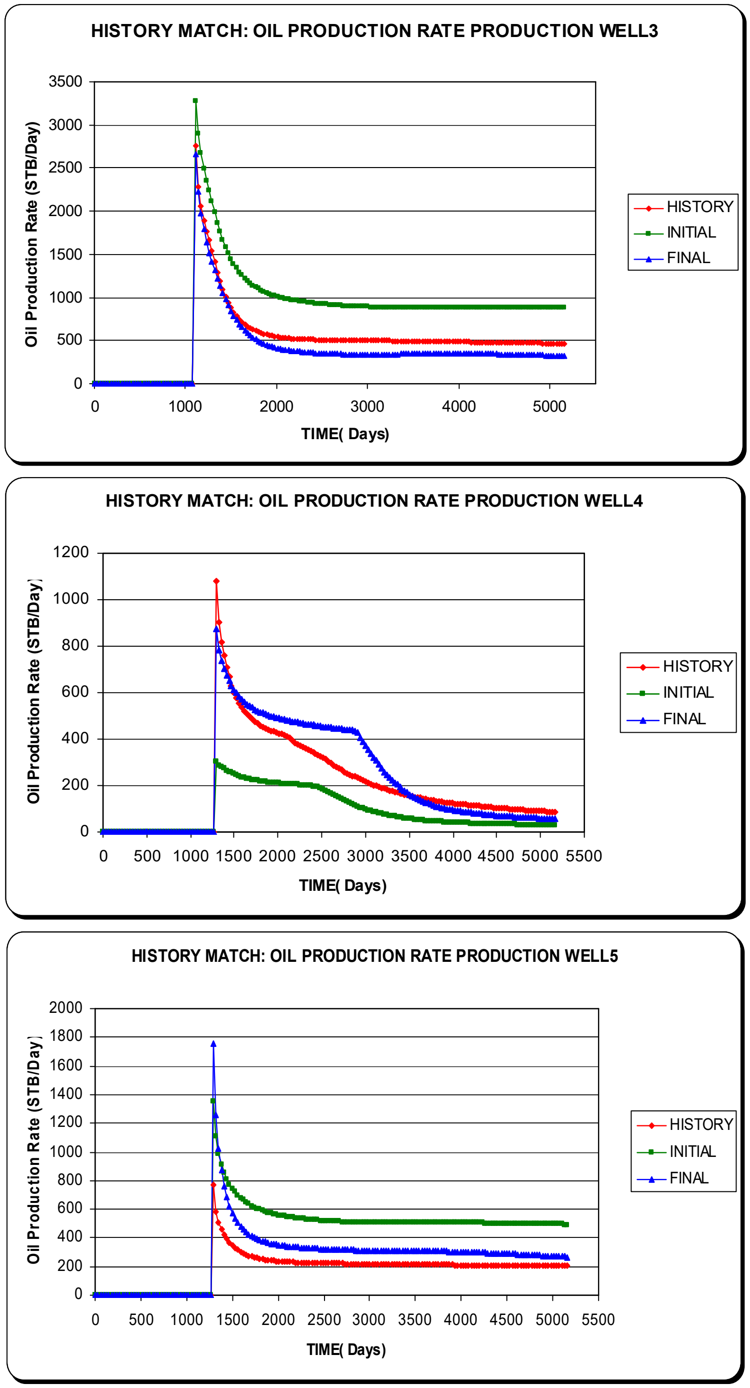


Figure 12. History matching result for the case where the permeability values in the new sub-regions are only conditioned to the original "hard" data.

5th case: Model updating using a reduced set (13000) of "hard" data simulating values in the less sensitive regions

This test uses the same procedures as for test 4, but the number of conditioning data from previously history matched sub-domains for simulating the values in the less sensitive regions is reduced from 19950 to 13000.

The results are illustrated in Figure 13 and Figure 14. Reducing the number of conditioning data has minimal effect on the reproduction of the production history information.

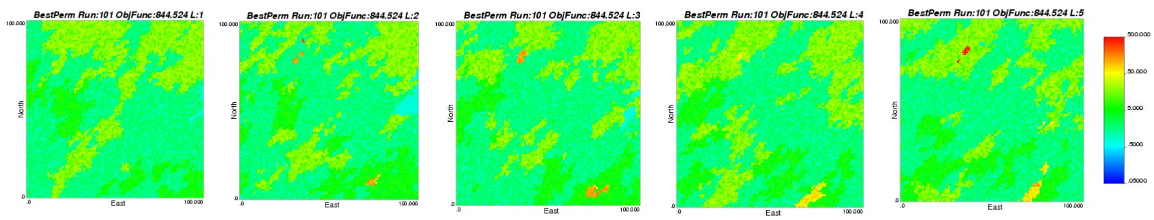
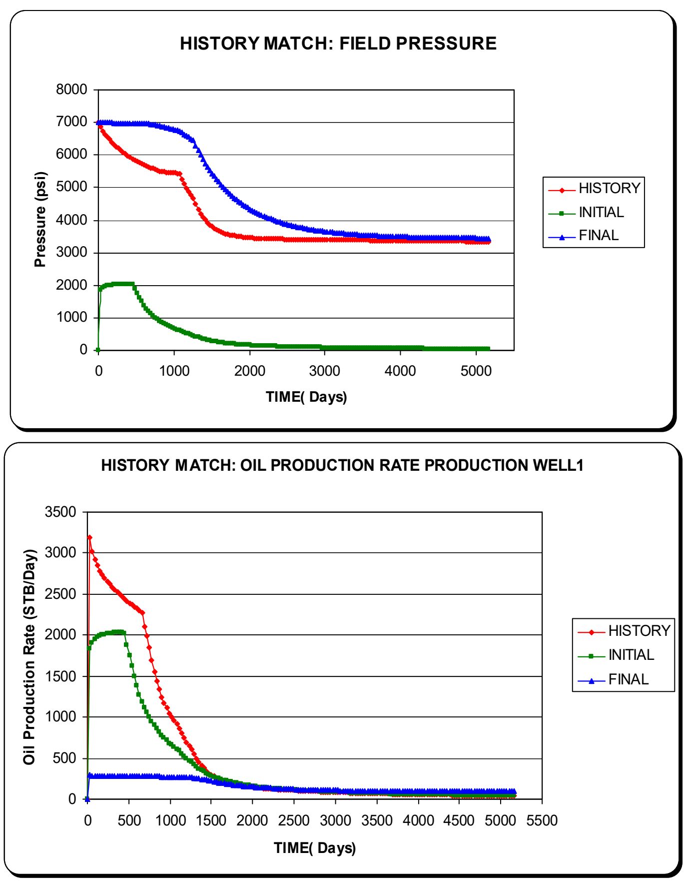


Figure 13. The updated permeability models obtained using a reduced number of conditioning data for simulating the permeability values in the less sensitive regions



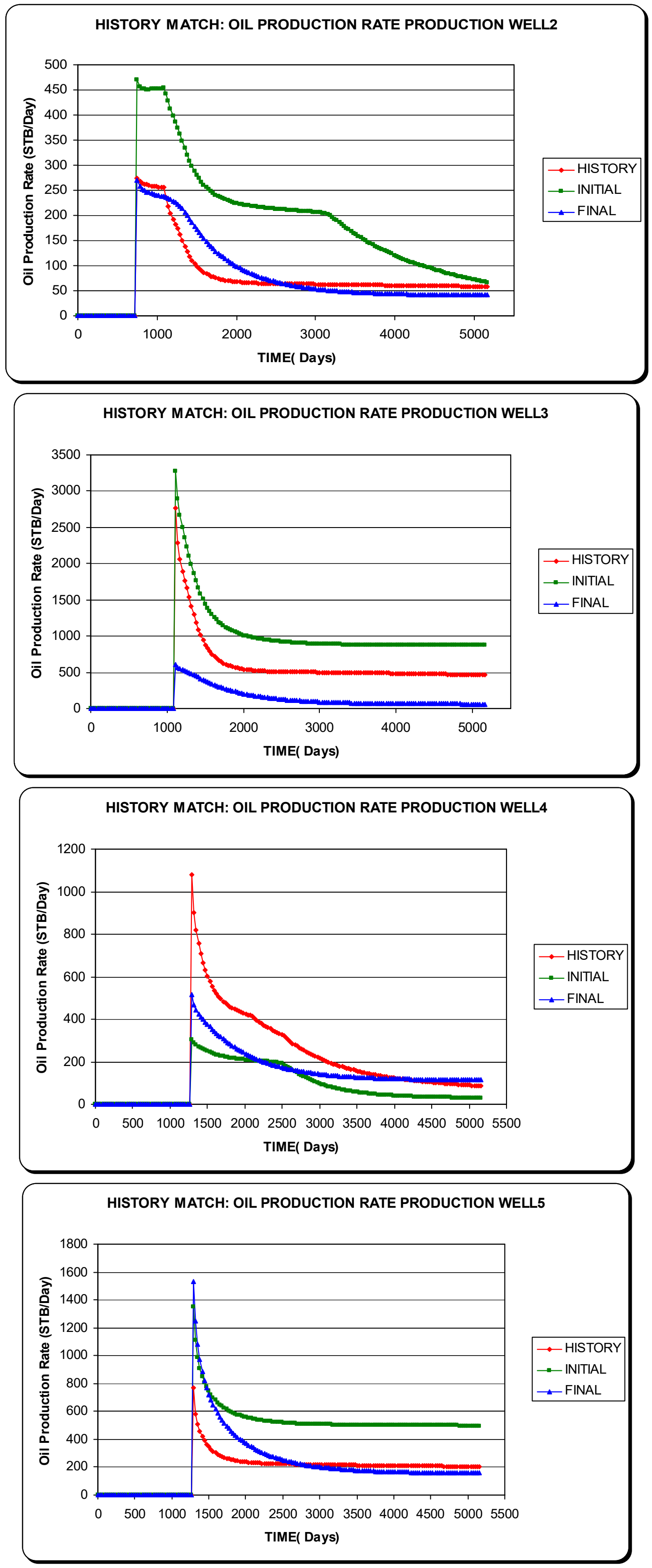


Figure 14. History matching result plots when the permeability values in the less sensitive regions are conditioned to a lesser number of updated permeability values in the sub-domains.

APPENDIX 1

Description of Pro-HMS Input/Output Files

Pro-HMS User Input Files		
File name	*Format*	*Instruction*
Conditional.txt	Conditional Data [Name of the file] 5 [Number of columns] X [Name of 1st column] Y [Name of 2nd column] Z [Name of 3rd column] Perm [Name of 4th column] Rank [Name of 5th column] 21 43 5 1.40405 0.05 [Data] 27 64 5 4.73353 0.05 54 41 1 1.05201 0.05 35 88 1 223.917 0.05	User specified conditioning data file. The column numbers of relevant data needs to be specified. If rank or weights for data are not specified, data are assumed equal weighted.
Prodhistory.txt	Primary Production History [Name of the file] 3 [Number of columns] T FPR WOPR-P1 [Name of each column] 43 [Number of data row] 0 7006.199 0 [Data] 30 6851.058 3190.944 60 6738.524 3019.498 90 6643.017 2917.788	It is mandatory for the user to provide this file. There is no maximum on the number of variables that can be included in the objective function.
RegionIndices.txt	-9999 [No index is required; file starts with data] -9999 -9999 -9999 1 -5 -9999 -2 -9999 3	It is generated by Pro-HMS only if the user wants to have sub-domains. If a user wants to provide this file, the name should always be 'RegionIndices.txt'
BASE.DAT	It should be consistent with the data file used for covariance matrix calculation.	Simulator data file for perturbation
DATA.DAT	It should be consistent with the data file used for perturbation. The only difference occurs in the reservoir size.	Simulator data file for covariance matrix calculation

Pro-HMS Output Files

Parameter files

<table>
<tr><th>File name</th><th>Description</th></tr>
<tr><td>sensregion.par</td><td>Parameter file for covariance matrix calculation</td></tr>
<tr><td>dd.par</td><td>Parameter file for domain delineation</td></tr>
<tr><td>hmissim.par</td><td>Parameter file for perturbation

[Note]
❖ The order of sub-domains perturbations is fixed in such a way that Pro-HMS perturbs sub-domains in the sequence of region indices.

For example, if the format of ‘RegionIndices.txt’ is
-9999
-2
3
1
.
.
.
, then Pro-HMS will perturb sub-domains whose indices are -2, 3, and 1 sequentially.

❖ If the perturbation order is to be changed, it is possible to change the order in ‘hmissim.par’ directly.

For example, if ‘hmissim.par’ has format(38th line) such as
SubdomainsOpt - 2 3 1 -9999
then Pro-HMS will perturb sub-domains whose indices are -2, 3, and 1 sequentially. However, if the sequence of perturbation has to be 1, then -2, and finally 3, then the 38th line has to be changed as
SubdomainsOpt 1 -2 3 -9999

❖ When loading a previously existing ‘hmissim.par’ file into Pro-HMS, the file should follow the format of the sample file. The default ‘hmissim.par’ file has only the key word ‘SubdomainsOpt’. This corresponds to performing the sub-domain perturbations in the order of indices in ‘RegionIndices.txt’. If the sequence is to be changed, the new sequence has to be entered into ‘hmissim.par’ by opening it in a text editor and adding the sequence following the keyword SubdomainsOpt as</td></tr>
</table>

	in above.

Covariance matrix related files

File name	*Description*
sisim.dbg	Debug file from SISIM
sisim.out	Ensemble of permeability values
sisim.par	Parameter file for generating multiply reservoir models
PR#1L#1.ps	Postscript permeability maps for one realization. One map per layer is generated.
upsc-por.out	Porosity data for upscaling
upscaledperm.out	Effective permeability data after upscaling
upscaler.par	Parameter file for upscaling the realizations
wholeoutput.out	Gridblock Pressure values extracted from flow simulation output file (*.PRT file in case of Eclipse)
perminclude.inc	Upscaled permeability values to be used in flow simulation.
upscaledpermX.out	Absolute permeability data after upscaling
upscalK #1Layer1.ps	Upscaled permeability map corresponding to the first realization. One map per upscaled layer will be generated.
covariance.out	File with computed covariance matrix
pressurearrays.out	Pressure data corresponding to the time specified by the user for covariance matrix calculation
covdiagonal.out	The diagonal terms of covariance matrix
sensitivitco L# 1.ps	The map of sensitivity coefficients. The region with high sensitivity coefficient will be the basis for defining sub-domains.
tridiagmat.out	Reduced tri-diagonal matrix to get eigenvalues
eigenvalue.out	Eigenvalues of the tri-diagonal matrix
eigenvectors.out	Eigenvectors of the tri-diagonal matrix
1wellflowev1.txt	Eigenvector corresponding to the first eigenvalue
1wellflowev2.txt	Eigenvector corresponding to the second eigenvalue
1wellflowev3.txt	Eigenvector corresponding to the third eigenvalue
1wellflowev4.txt	Eigenvector corresponding to the fourth eigenvalue
1wellflowev5.txt	Eigenvector corresponding to the fifth eigenvalue
1wellflowevalue.txt	Eigenvalues after Principal Component Analysis(PCA) listed in their magnitudes
sortedigenvalues.out	Sorted eigenvalues according to their magnitude

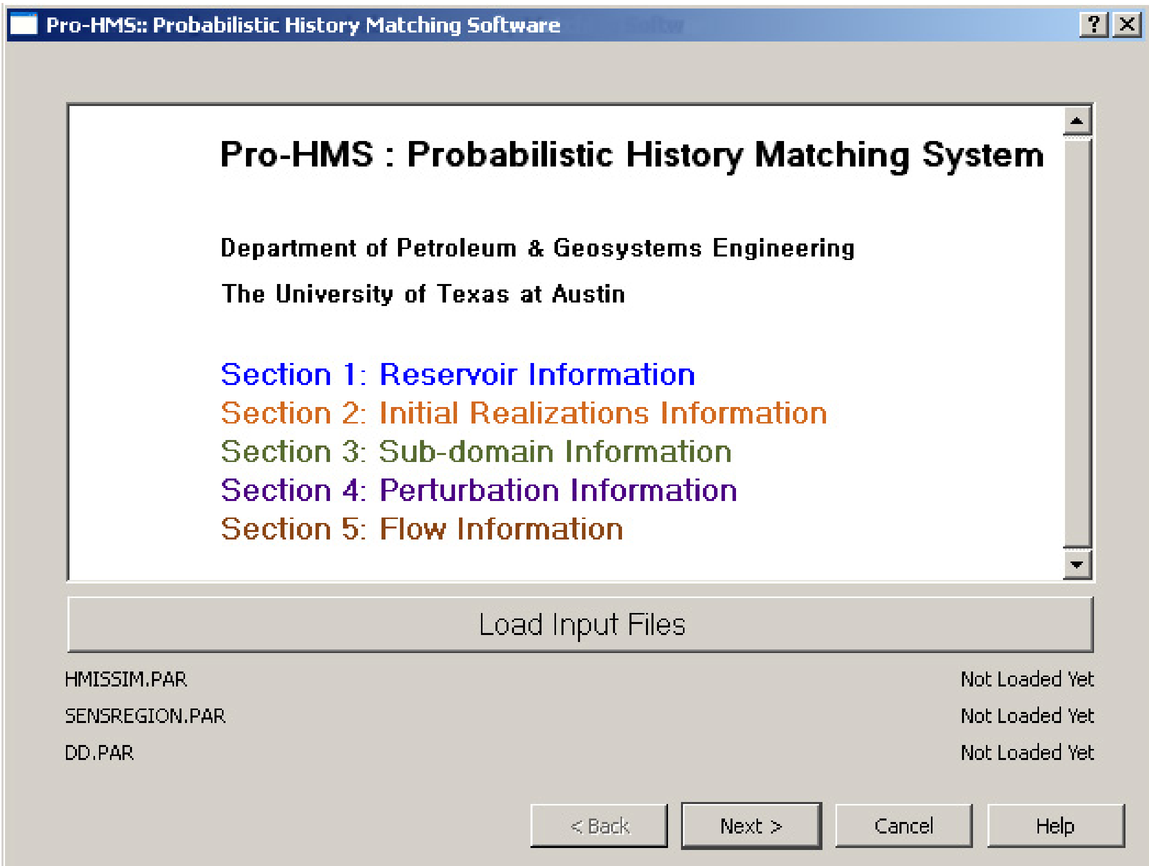
Domain Delineation related files

File name	*Description*
Output.txt	Highest eigenvector component and its position of occurrence
RegionIndices.txt	Region indices of the reservoir -9999 means not sensitive regions. Minus means negative eigenvector component. Plus sign indicates positive eigenvector component.

Perturbation related files	
File name	*Description*
ReadData.txt	The conditional data that is reading for perturbation
ReadParameter.txt	The parameters that is read from parameter file by Pro-HMS
PERM.DAT	Permeability data used as an include file for perturbation Overwriting on the same file.
PermMap_after_OuterLoop.out	Permeability data after each outer loop Overwriting on the same file
PermOptResult.txt	Permeability data after perturbations of each outer loop for each sub-domains Overwriting on the same file
ProdOptResult.txt	Production matching result after each outer loop for each sub-domains
RDout.txt	Objective function information after outer loop for each sub-domains

APPENDIX 2

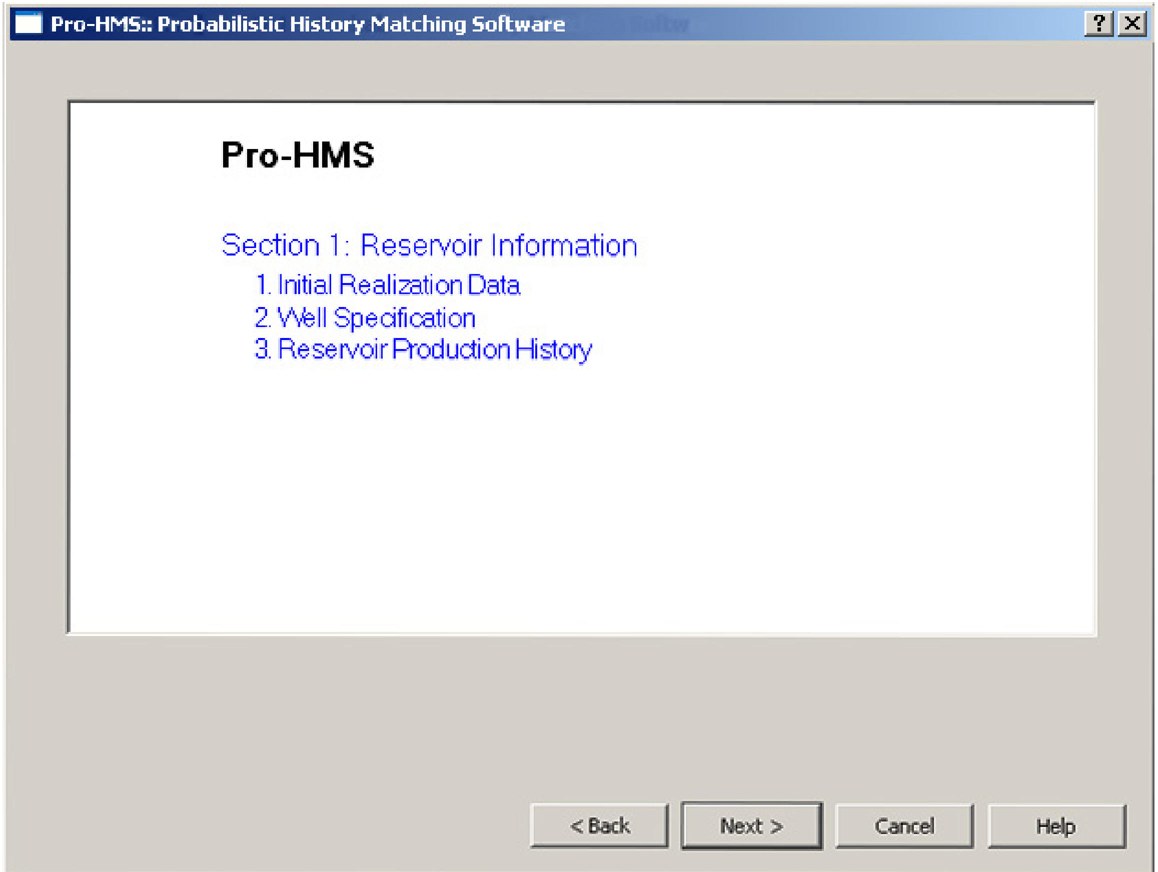
Pro-HMS Screen Manual



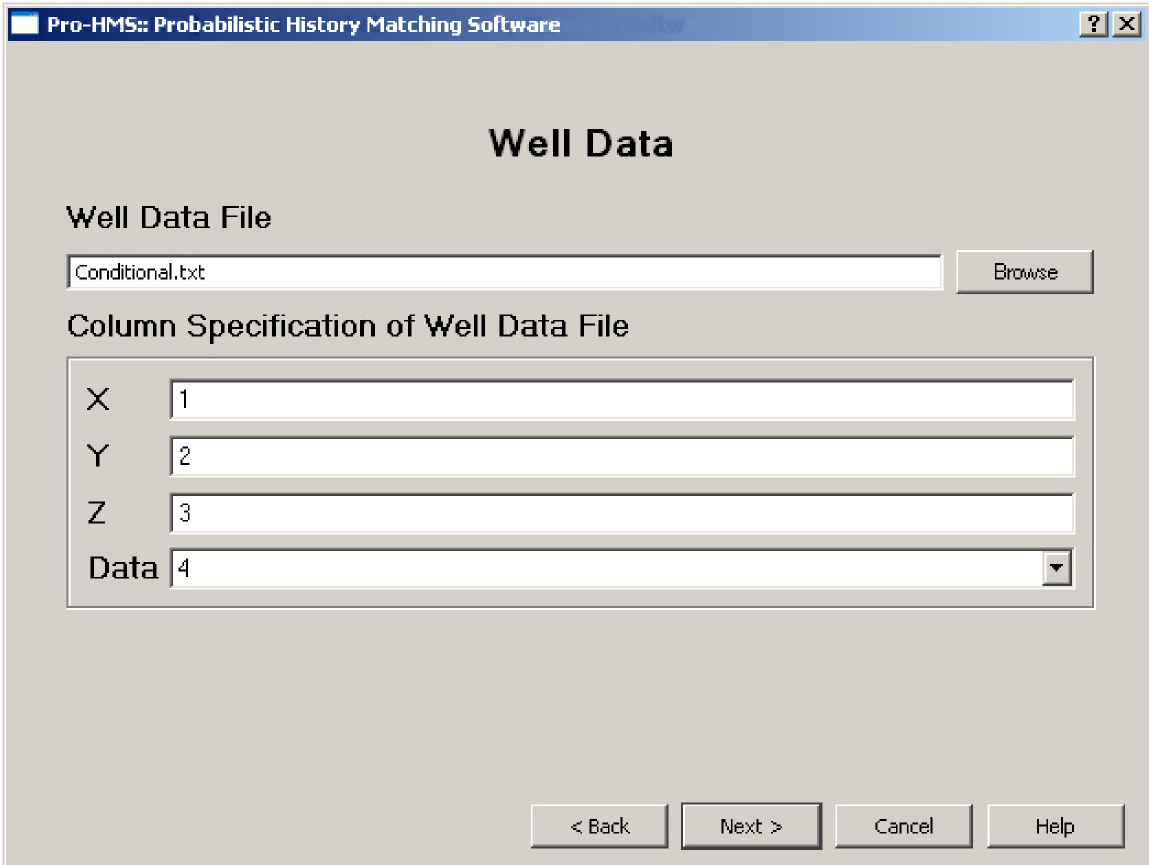
This is the opening screen when Pro-HMS is launched. Pro-HMS is composed is of five sections:

1. Reservoir Information: Basic data such as grid size, dimensions etc. are specified.
2. Initial Realizations Information: Ensemble size, conditioning data etc. are specified.
3. Sub-domain Information: Option to calculate or specify sub-domains, method for calculating sub-domains can be specified here.
4. Perturbation Information: Indicator simulation parameters, updating parameter specification.
5. Flow Information: Simulator type, data file, flow variables to match etc.

There is an option in this screen which is 'Load Input File'. If you have parameter files generated by this software already, that option is clicked to load the files instead of going through all the screens.

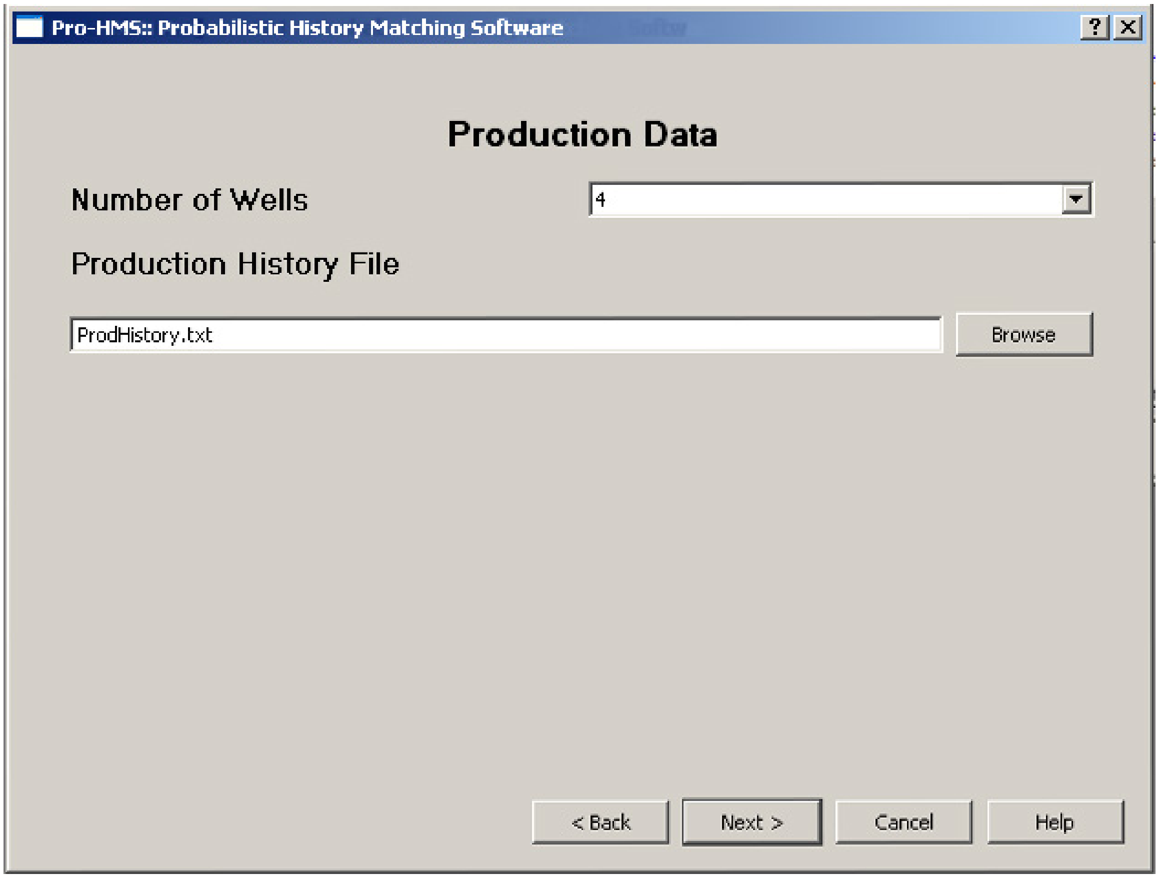


Section 1 is for reservoir information of the reservoir that is to be history matched. The starting guess (initial realization), the number and location of wells and the details of the reservoir production history are specified here.

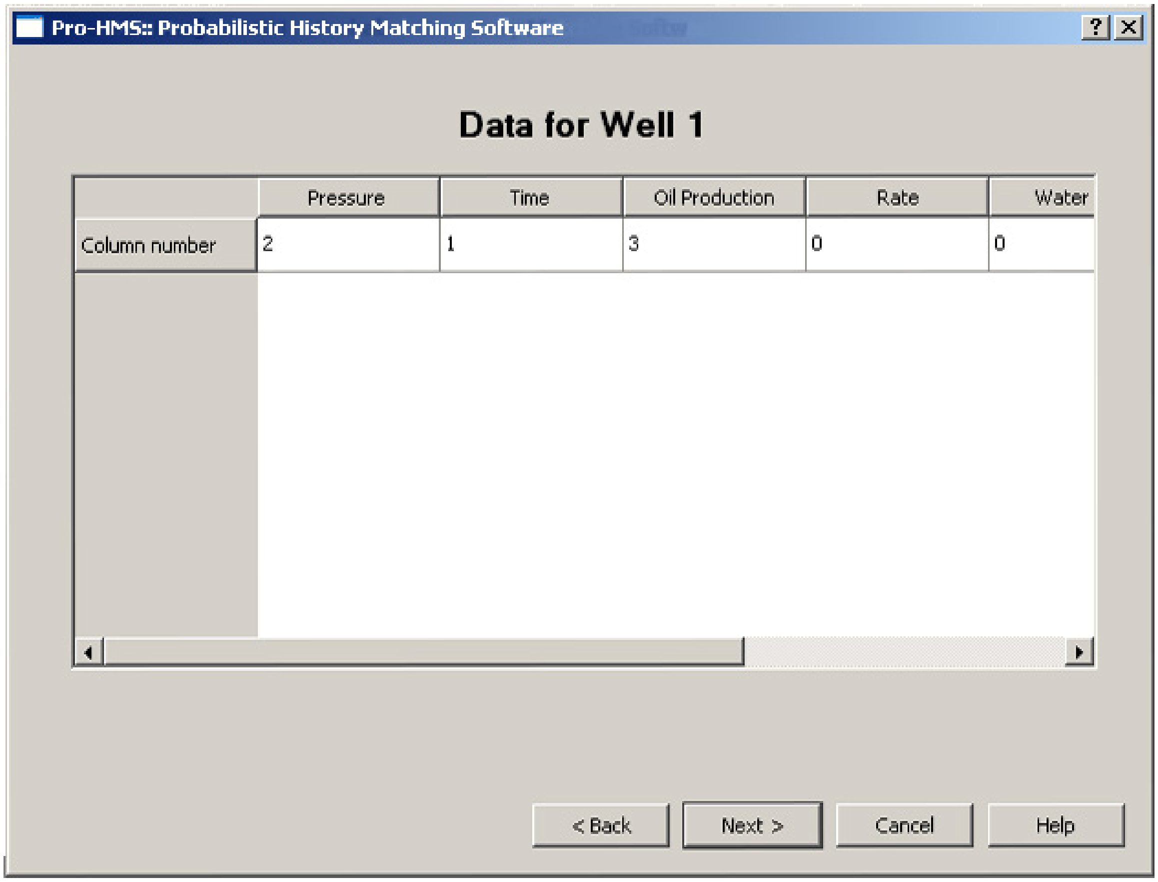


The name of the file with the conditioning "hard" data is specified. The conditional hard data are used when Pro-HMS does sequential indicator simulation.

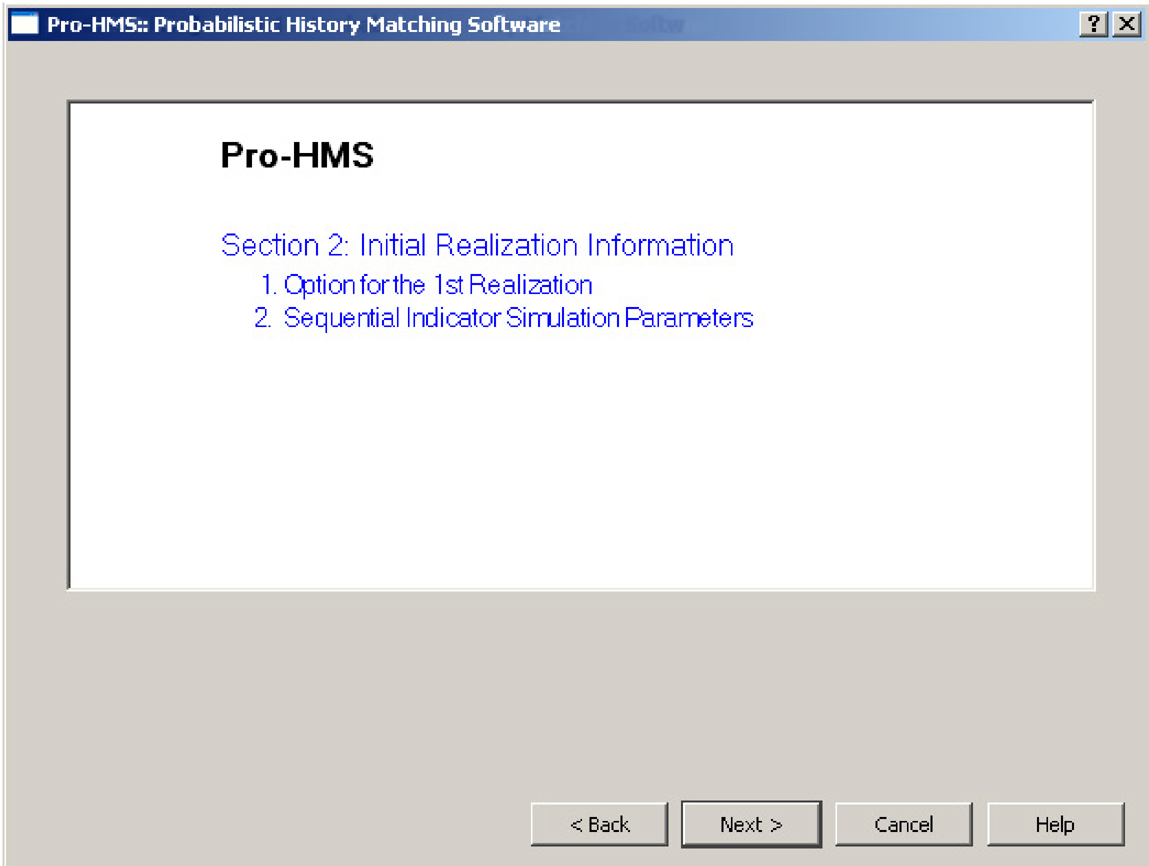
Click 'Browse' button to locate the appropriate well data file. Indicate column numbers for data location. The variable means the column number for the data. The format of the file should be the coordinates of the data location in columns and corresponding data value also in a column. (See instruction of 'Description of Input/Output Files')



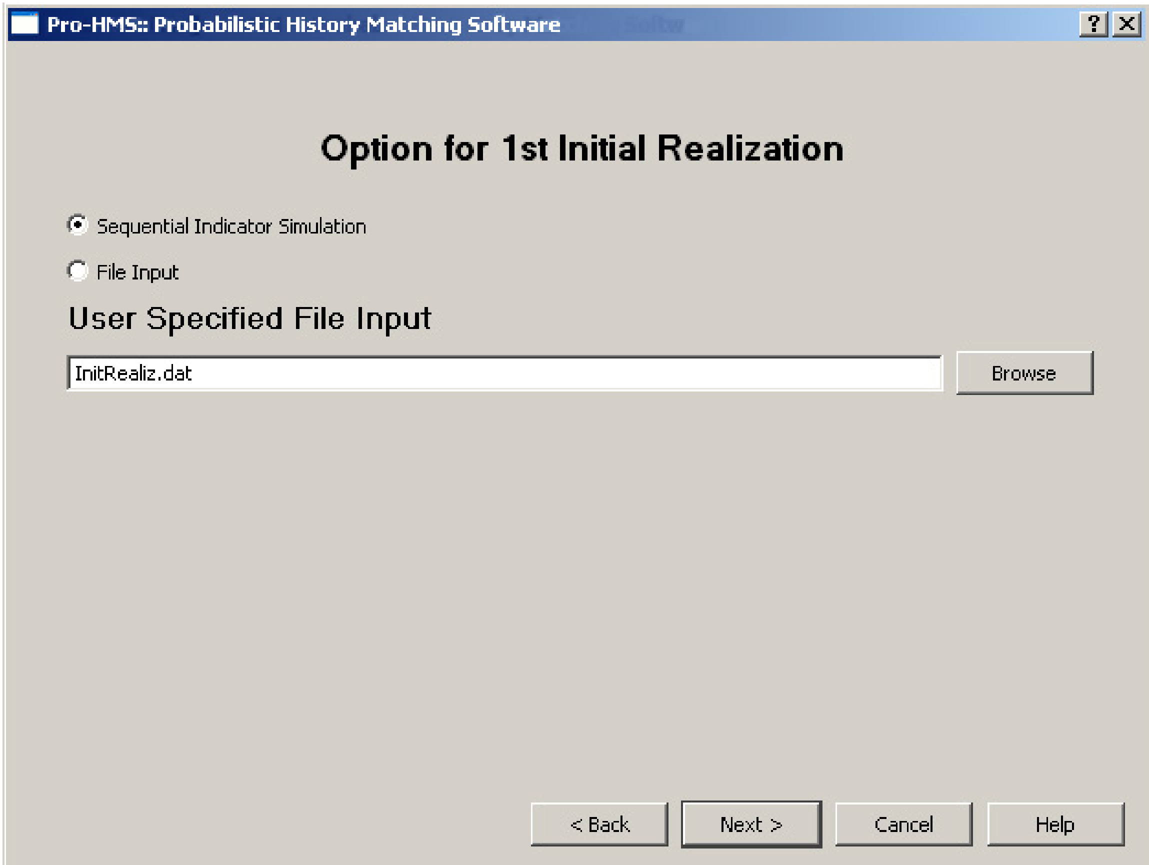
In this screen, the production history (that you want to match with) file and the number of wells whose history are to be matched are specified.



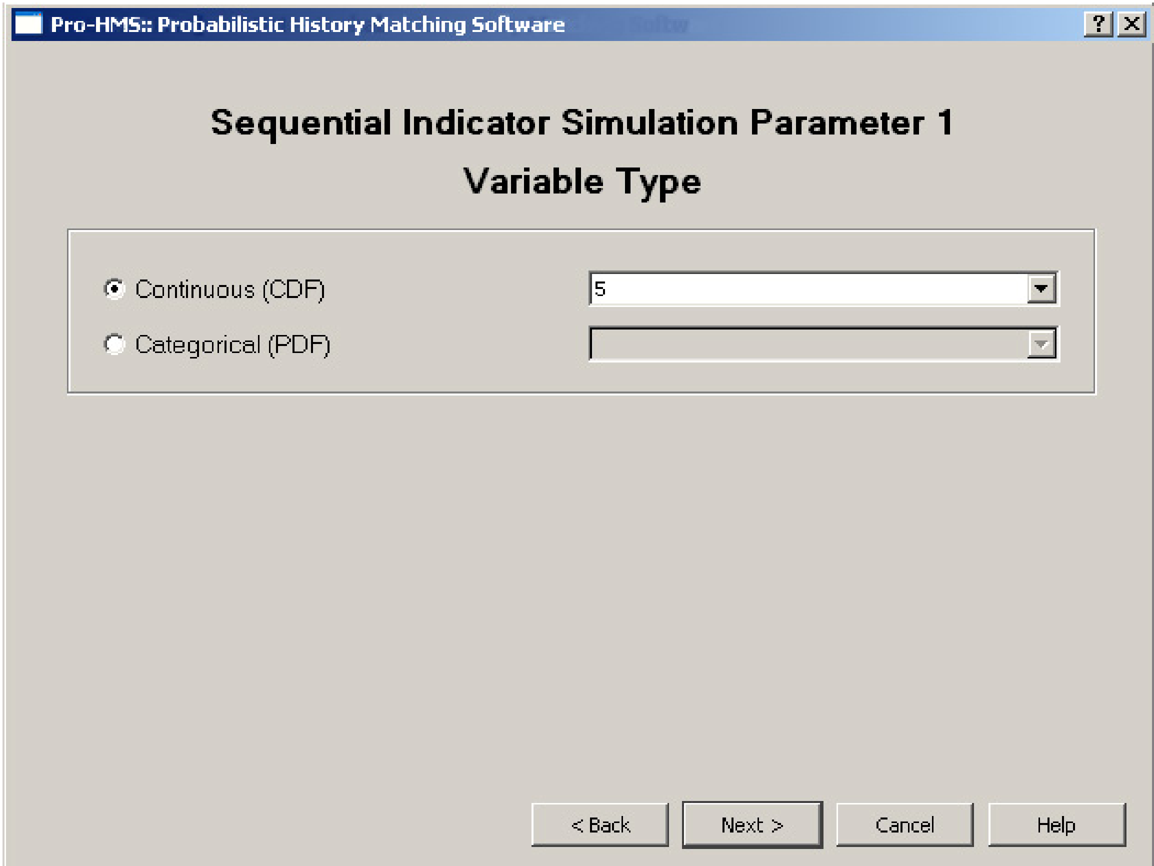
The column for the production variables for a particular well is specified. There are as many screens as the number of wells. The column numbers for well pressure data, production time and well production rates are specified.



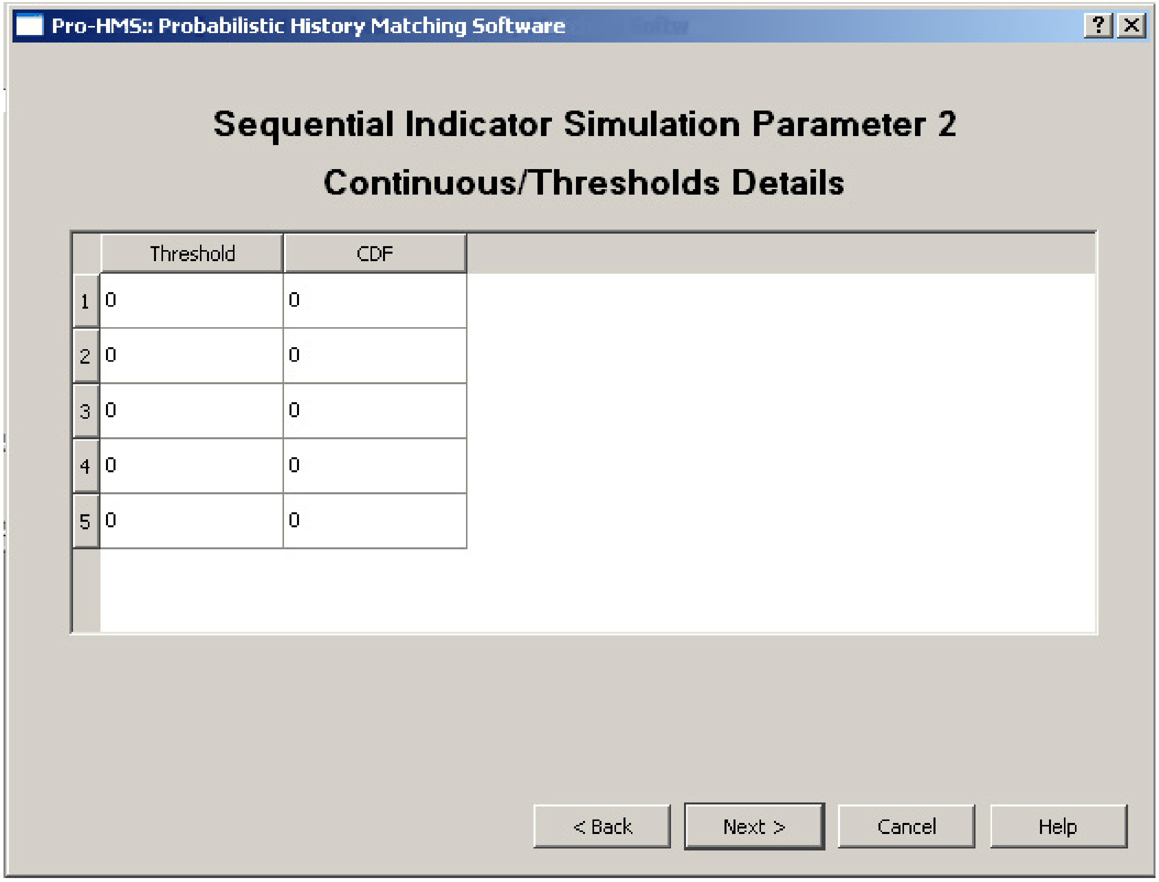
This screen is the introduction to section 2. Section 2 provides information for the initial realization. To obtain an initial guess for the history matching process, sequential indicator simulation is performed. The inputs for that initial SISIM realization are specified in this section.



An option for the initial realization is provided. Either a sequential indicator simulation generated internally within the program can be used or an initial guess can be read in from an external file. If the latter option is selected, the format of the external file should be compatible with the simulator's include files.



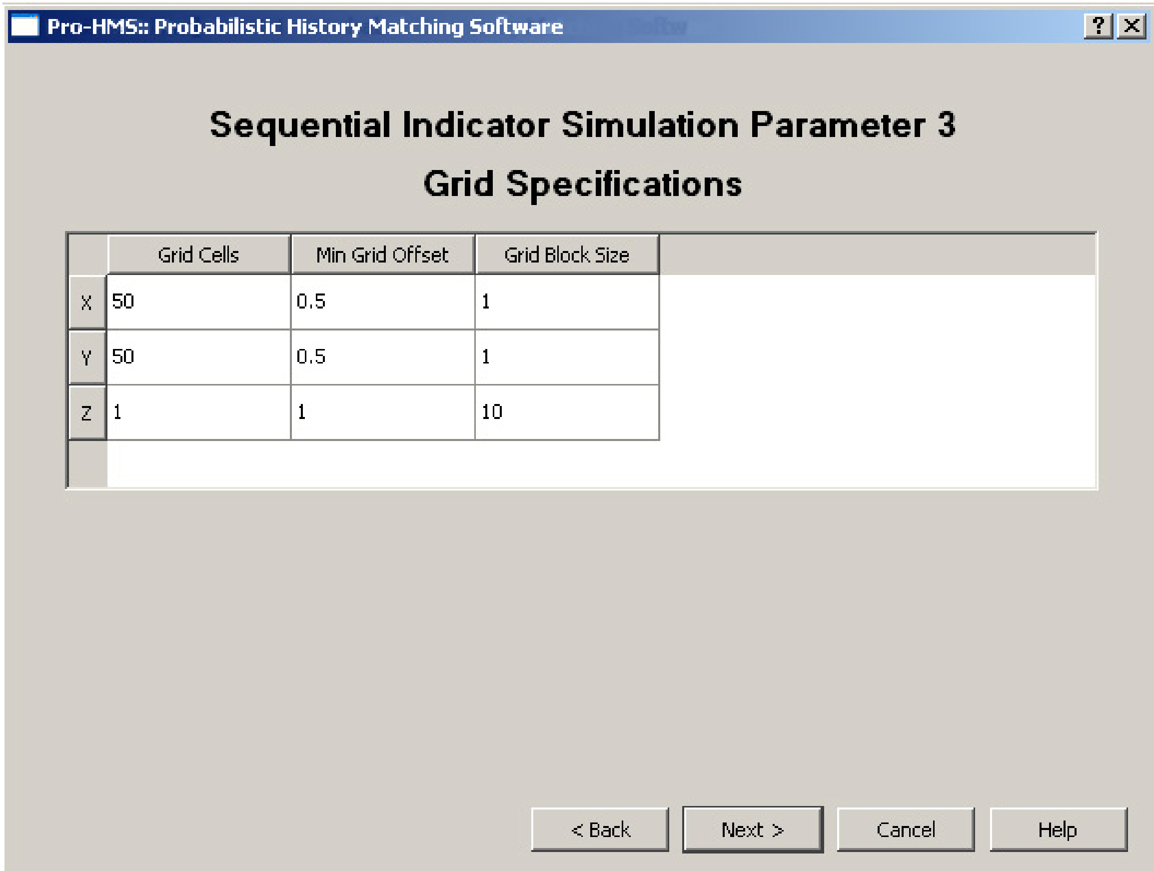
The first screen for sequential indicator simulation - the variable type is selected first. If the variable type is continuous, the number of thresholds has to be specified. If the variable type is categorical, the number of categories has to be specified.



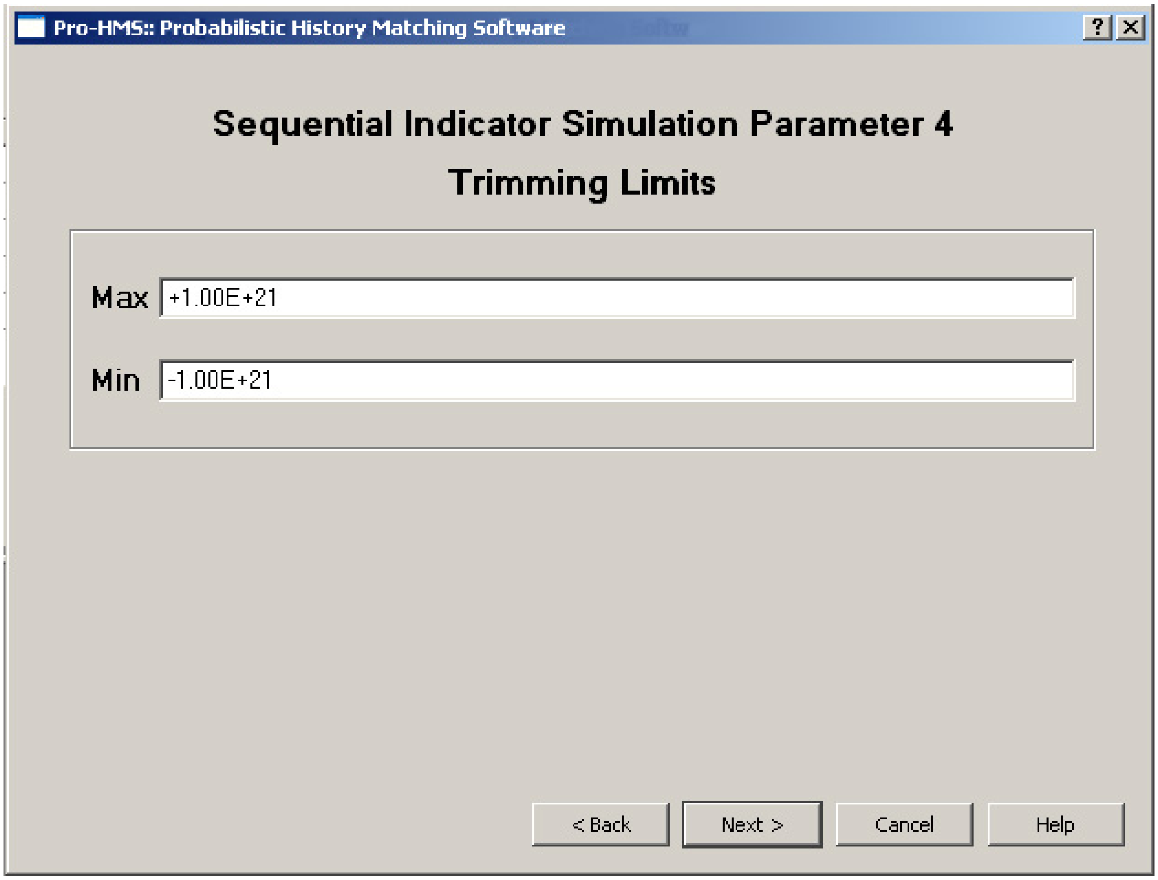
The second screen for sequential indicator simulation:

If the variable type is specified to be continuous then the indicator threshold values have to be specified in the first column and the corresponding cumulative density function (CDF) values in the second column.

If the variable type has been chosen as categorical in the previous screen, the indicator categories have to be specified in the first column and the corresponding probability density function (PDF) values in the second column.

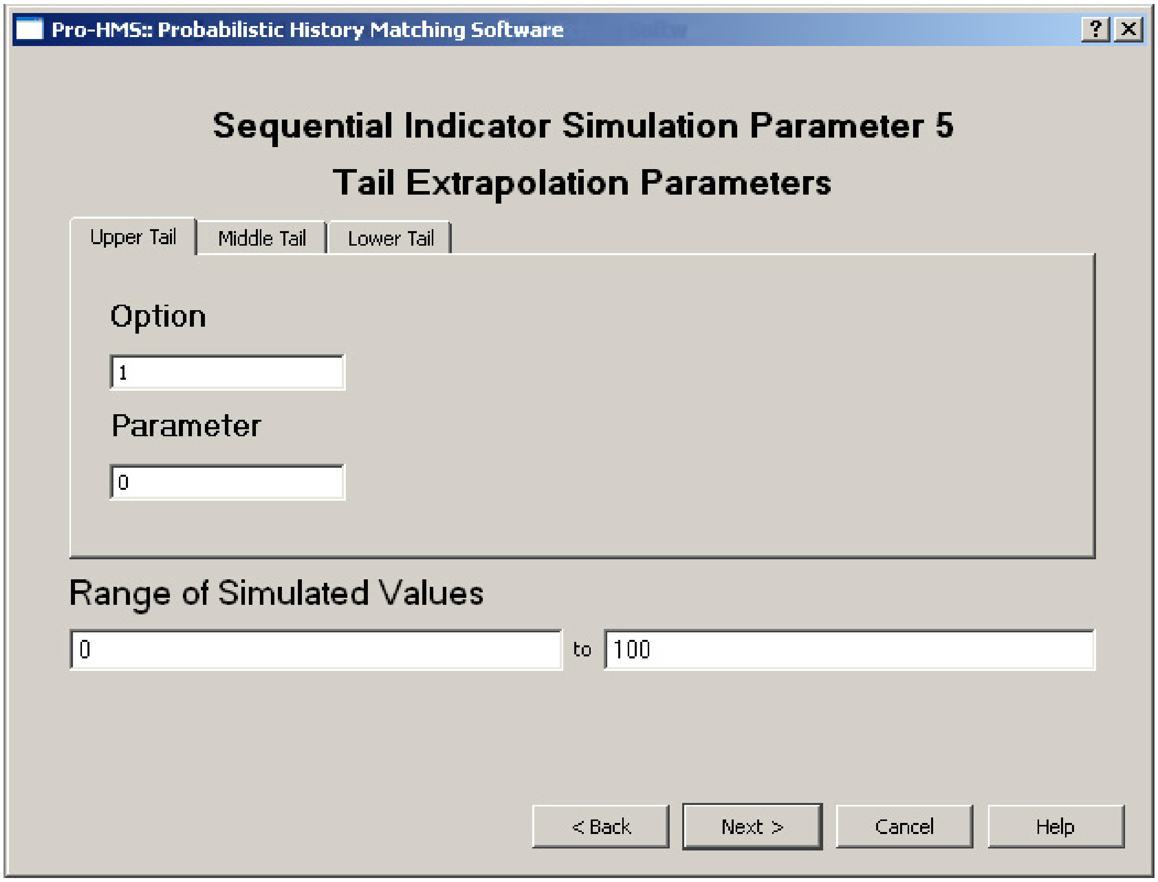


The size of the reservoir is specified in this third screen - the first column is the number of grid cells, the second column is the minimum grid offset, and the last column is the cell size.



The fourth screen of the sequential indicator simulation:

The trimming limits for the conditioning data are specified. There are maximum and minimum data values that are used in the simulation.



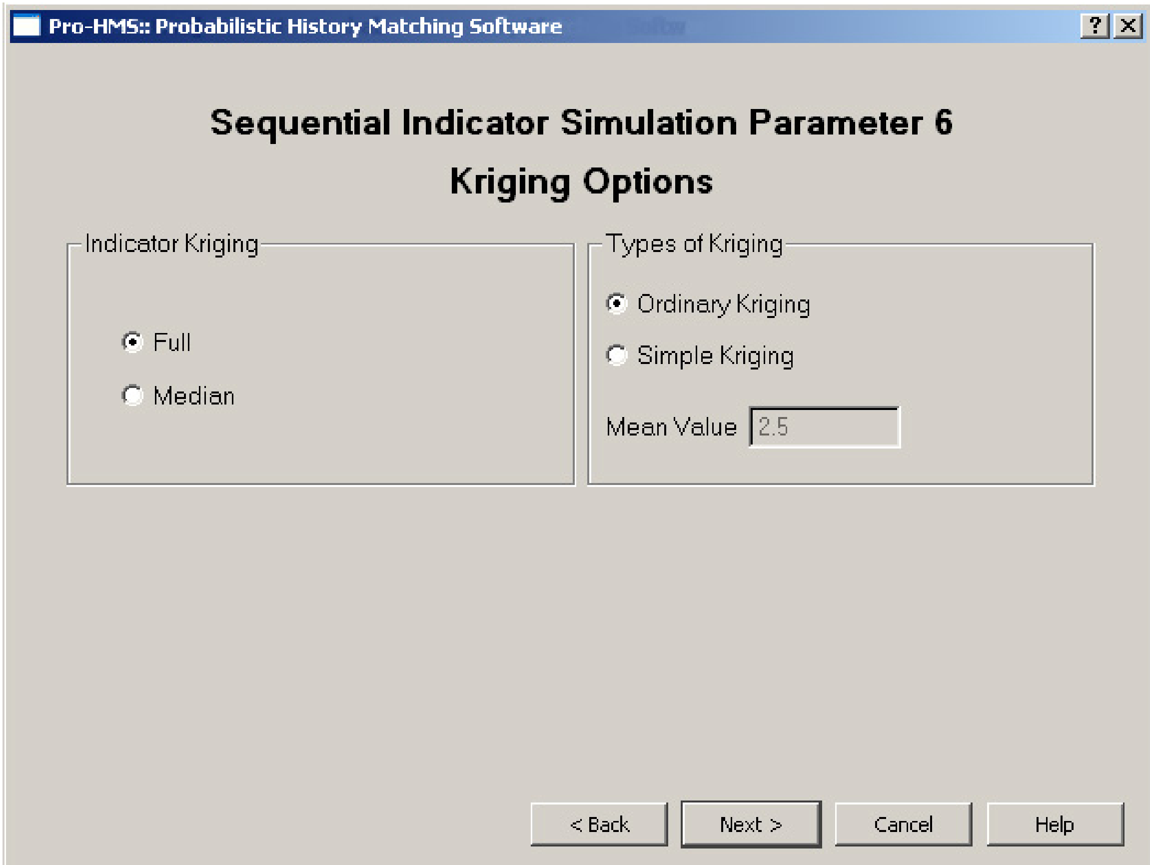
The preference for tail extrapolation of the distribution is specified here. Options for the interpolation of the upper and lower tail of the distribution as well as the middle part of the distribution have to be specified.

Upper tail:
If the option is chosen to be '1', then Pro-HMS implements linear interpolation to the upper limit.
If the option is chosen to be '2', then it implements power model interpolation.
If the option is chosen to be '3', then it implements linear interpolation between tabulated quantiles
If the option is chosen to be '4', then it implements hyperbolic model extrapolation. This option '4' is only available for continuous variables.
If the option is selected to be 2 or 4, the power law parameters are specified under the parameter heading.

Middle tail::
Option 1 implements linear interpolation
Option 2 implements power model interpolation
Option 3 linear interpolation between tabulated quantile values (only for continuous variables).

Lower Tail: specify the extrapolation in the lower tail:
lower tail=1 implements linear interpolation to the lower limit
lower tail=2 power model interpolation to the lower limit
lower tail=3 linear interpolation between tabulated quantiles (only for continuous variables).

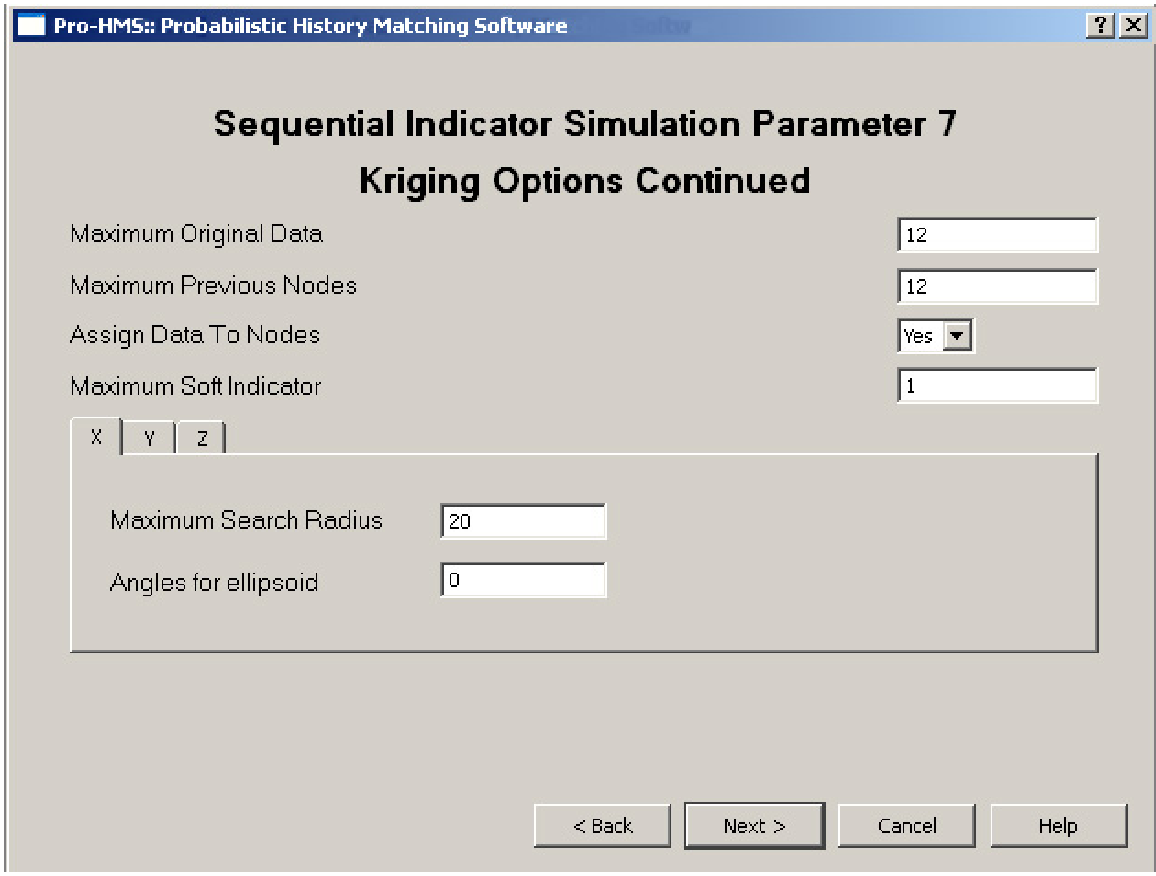
'Range of Simulated Values' is the range of permeability data in the simulated initial model. The numbers that can be input are any real numbers.



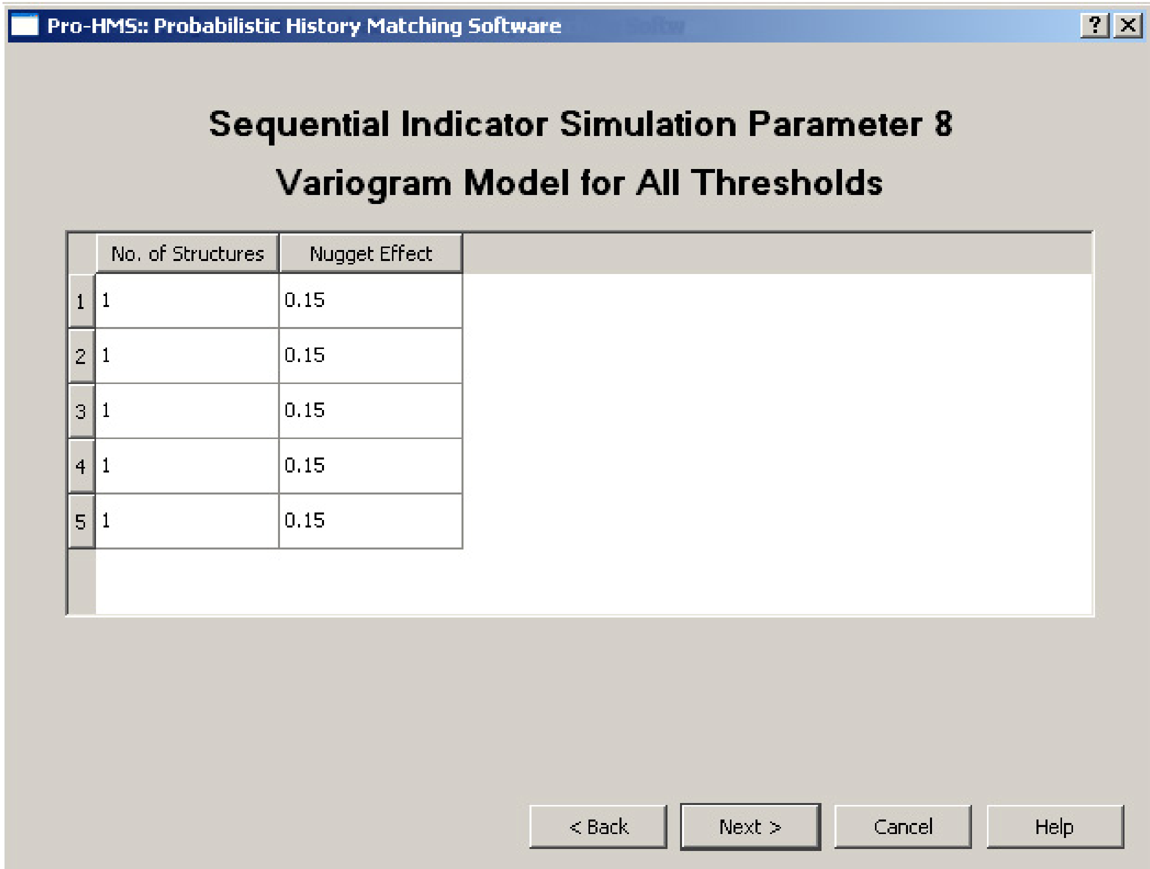
The sixth screen of sequential indicator simulation:

If 'Full' indicator kriging option is selected, then the variogram models corresponding to each threshold/category is used to establish the local conditional probability corresponding to that threshold. On the other hand, if 'Median' indicator kriging is opted, then the median approximation is used, i.e., a single variogram (corresponding to the median threshold) is used for all categories/threshold.

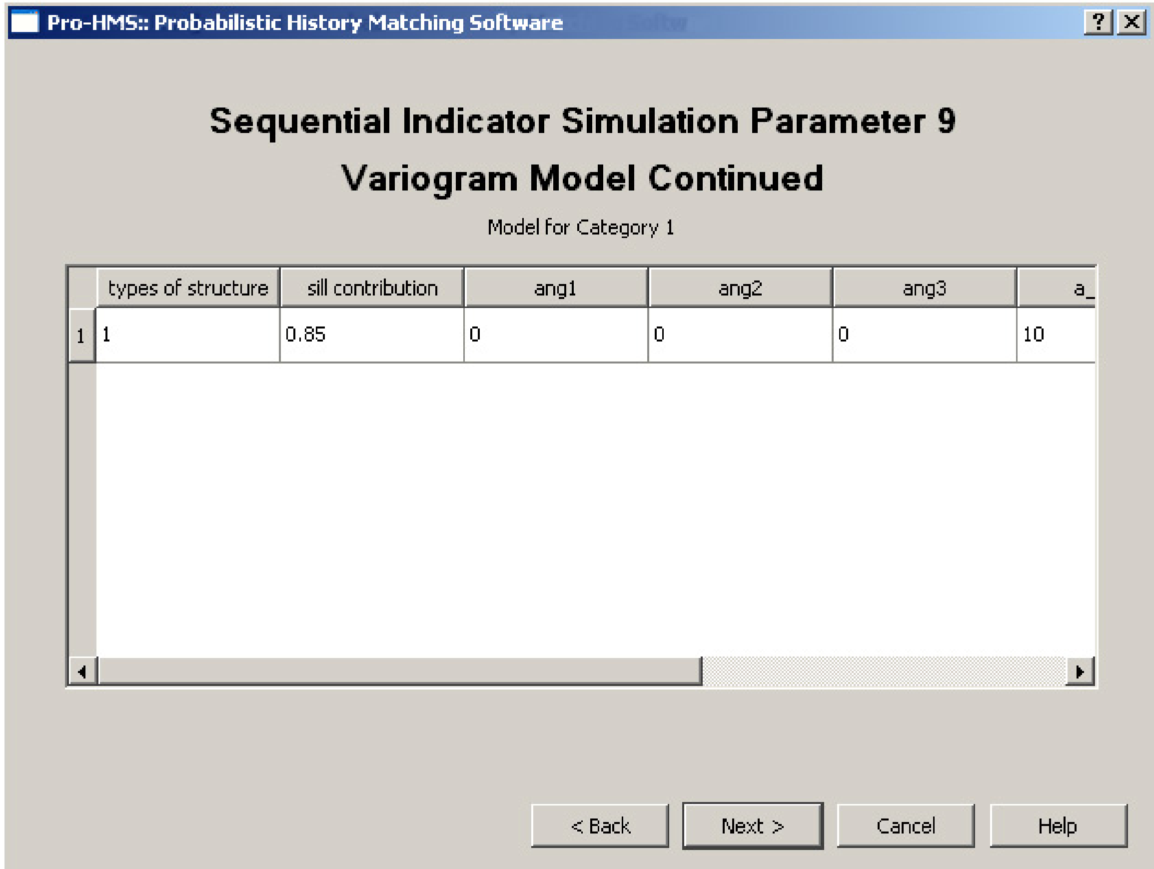
The bottom box is for the choice of kriging. Either ordinary or simple kriging can be selected. If simple kriging is selected, the global mean value has to be entered.



'**Maximum original data**' is the maximum number of the original well data (conditional data) to be used to simulate the value at a grid node in sequential indicator simulation.
'Maximum previous nodes' is the maximum number of previously simulated nodes to be used for constructing the local probability distribution at a node.
'Assign data to nodes' is an option for assigning data at a node.
If 'Yes' is selected, then the data are relocated to grid nodes and a spiral search is used; the parameters of Maximum original data are not taken into consideration.
If 'no' is selected, then the data and previously simulated grid nodes are searched separately: the data are searched with a super block search and the previously simulated nodes are searched with a spiral search.
'Maximum number of soft data': it restricts the number of soft data when an exhaustive secondary variable informs all grid nodes. This option is basically maximum number of soft data at node locations that will be used for the simulation of a node.
'Maximum Search Radius' specifies the size of the search ellipsoid in the horizontal direction and vertical direction.
'Angles for ellipsoid' is the angle parameters that describe the orientation of the search ellipsoid.



This screen is for variogram specifications. The number of structures and nugget effect for each threshold is specified.

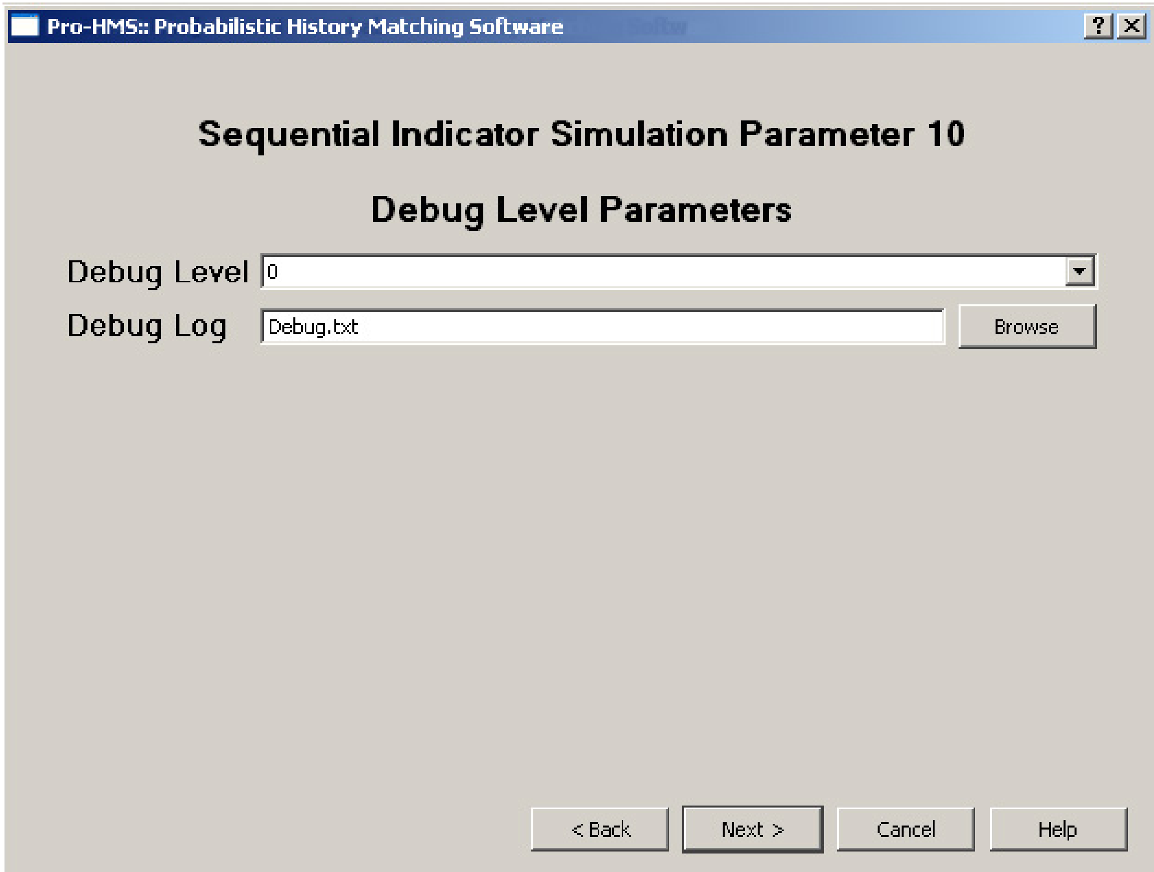


One such screen appears for each threshold. The screen will have as many rows as 'Number of Structure' indicated in the previous screen. Each screen requires to be filled up with the following information.

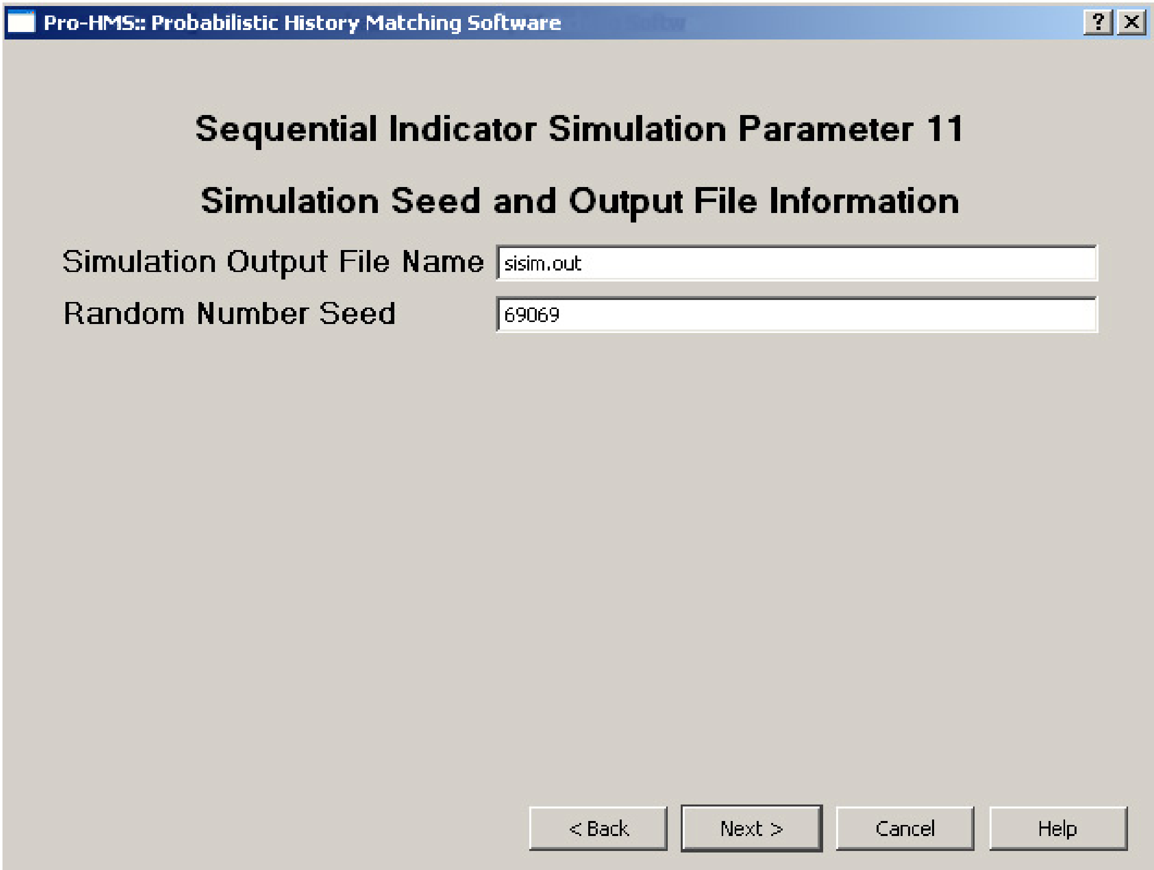
types of structure: 1 is for Spherical model, 2 is for Exponential model, 3 is Gaussian Model, and 4 is for Power model.
sill contribution: sill contribution of each structure
ang1,ang2,ang3: the angles defining the geometric anisotropy
aa_hmax: the maximum horizontal range
aa_hmin: the minimum horizontal range
aa_vert: the vertical range.

Note: Each semivariogram model refers to the corresponding indicator transform. A Gaussian variogram with a small nugget constant is not a legitimate variogram model for a discontinuous indicator function. There is no need to standardize the parameters to a sill of one since only the relative shape affects the kriging weights.

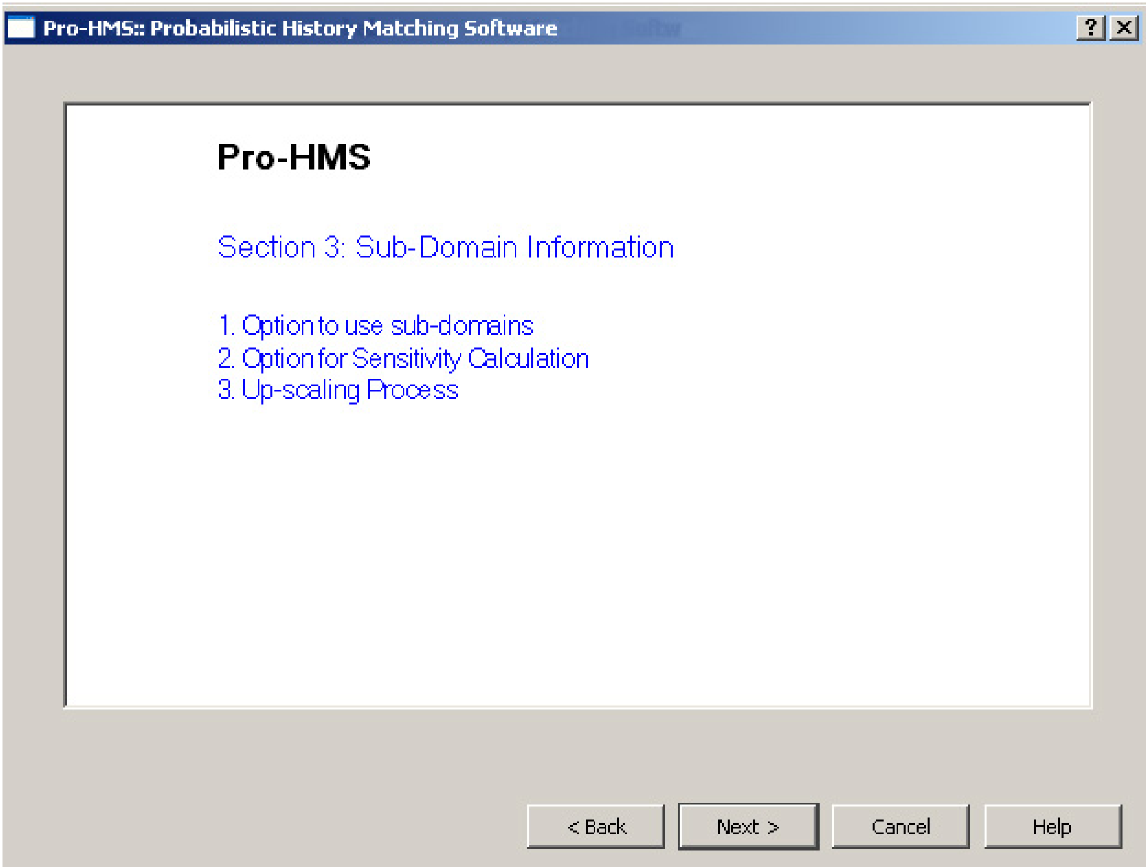
In the case of median indicator kriging, only the variogram corresponding to the median threshold needs to be specified.



A debug level between 0 and 3 can be specified. The higher the debugging level (the larger number) is the more information will be provided (e.g. kriging matrices, weights assigned to data at each simulation location). The ‘Debug Log’ output file in the folder can be checked for the debug outputs.

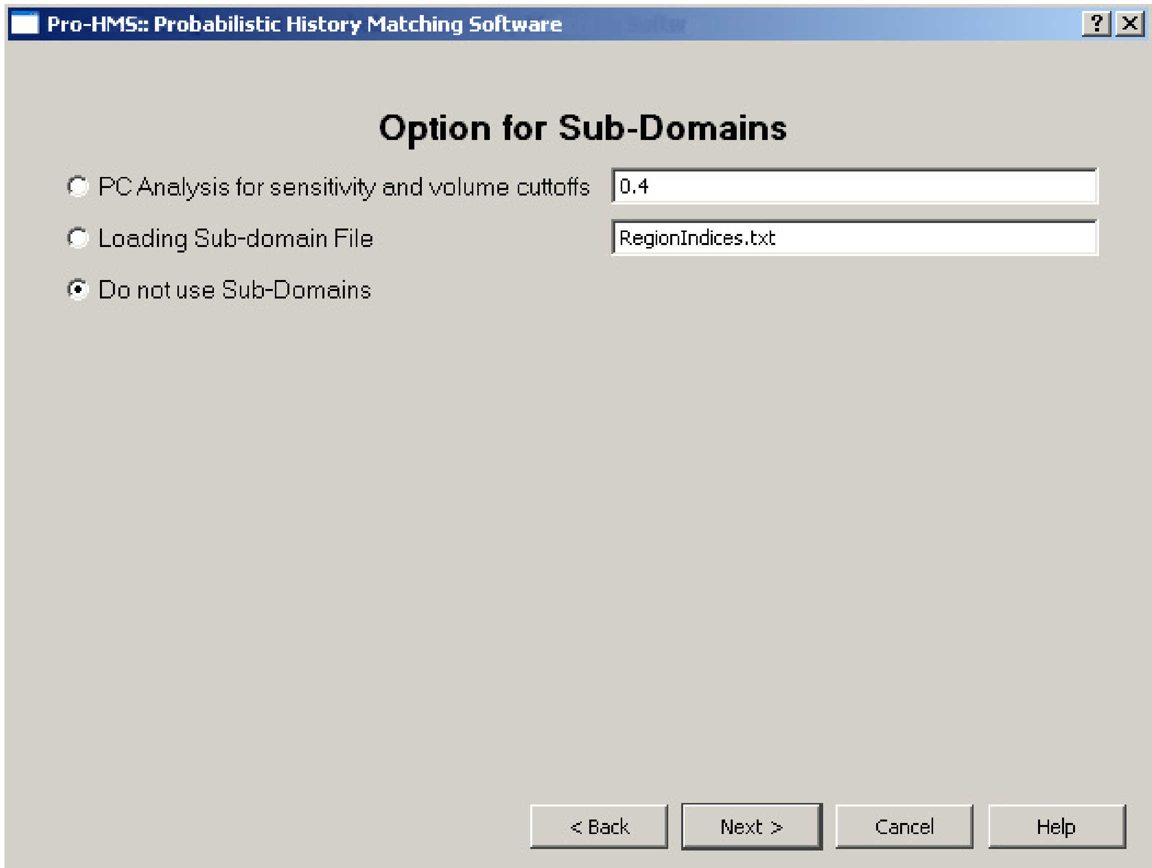


The name of output file generated by sequential indicator simulation has to be specified. The random number seed for ordering the random path through the nodes as well as for drawing values from the local probability distributions has to be specified. The seed should be a large odd integer.



Section 3 is for specifying sub-domain calculation options. The user has the option to use sub-domains or not for the perturbation step.

If the option for calculating/using sub-domains is selected, then there is another option for calculating sub-domains using Principal Component Analysis of sensitivity values or for loading a file with region indices calculated by the user. If the option for calculating the regions is selected, then the upscaling parameters have to be specified.



The two options for specifying sub-domains to the program exist.

'PC Analysis for sensitivity and volume cutoffs' is selected if the probability perturbation is to be done within sub-domains. This option performs internally within Pro-HMS a principal component decomposition of the sensitivity matrix, and then applies volume cutoffs for the sub-domains that are entered by the user in fraction.

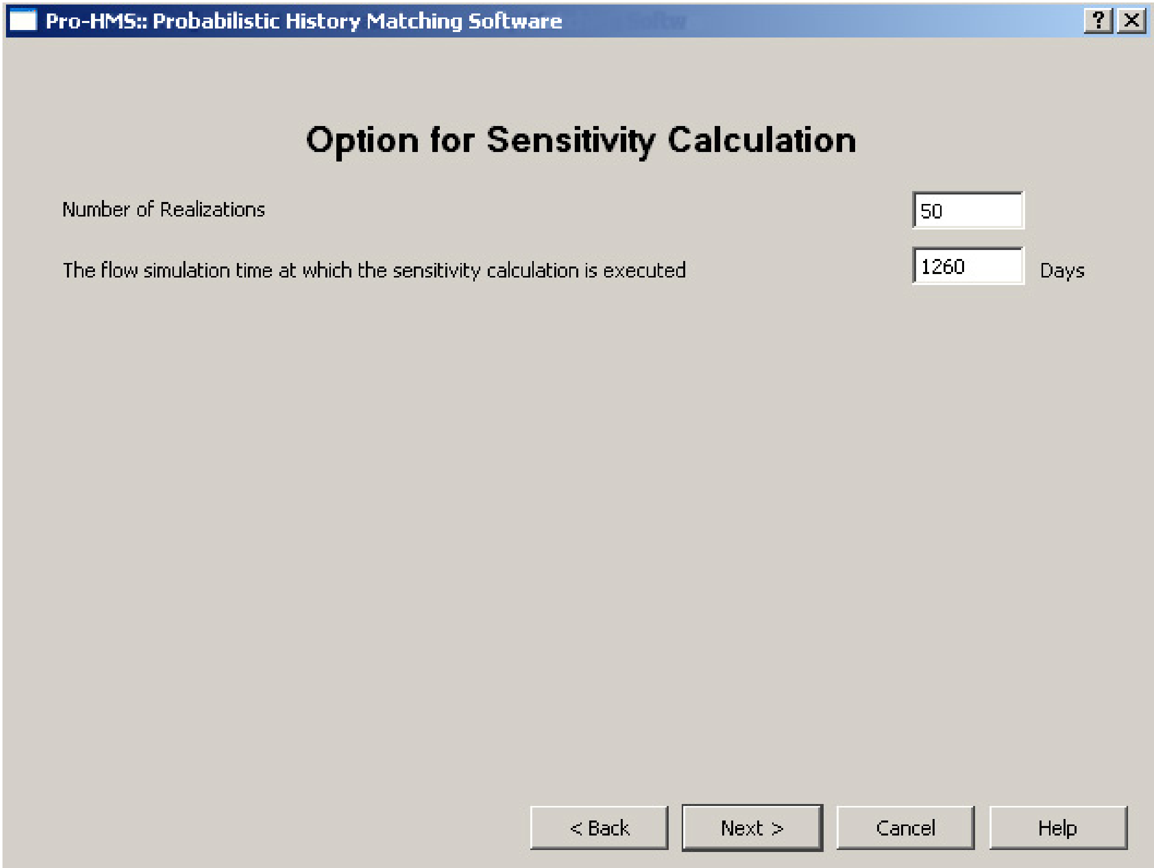
If the 'File for Sub-domains' is selected, a file with user generated sub-domain indices has to be loaded. That file has to have one column only with the region indices. The name always to be 'RegionIndices.txt'.

Example format of 'RegionIndices.txt':
-9999
1
1
1
1
2
-3
5
2
2

2
2
-9999
-9999
-9999
-9999

For more information on this format, refer to ‘Description of Input/Output Files’ (or FileDescription.pdf). Also, refer to the file format of ‘hmissim.par’, which is related to this file, for the correct file format.

If you don’t want to use sub-domain perturbation but desire to perturb all regions of the reservoir, the option ‘Do not use Sub-domains’ is selected.

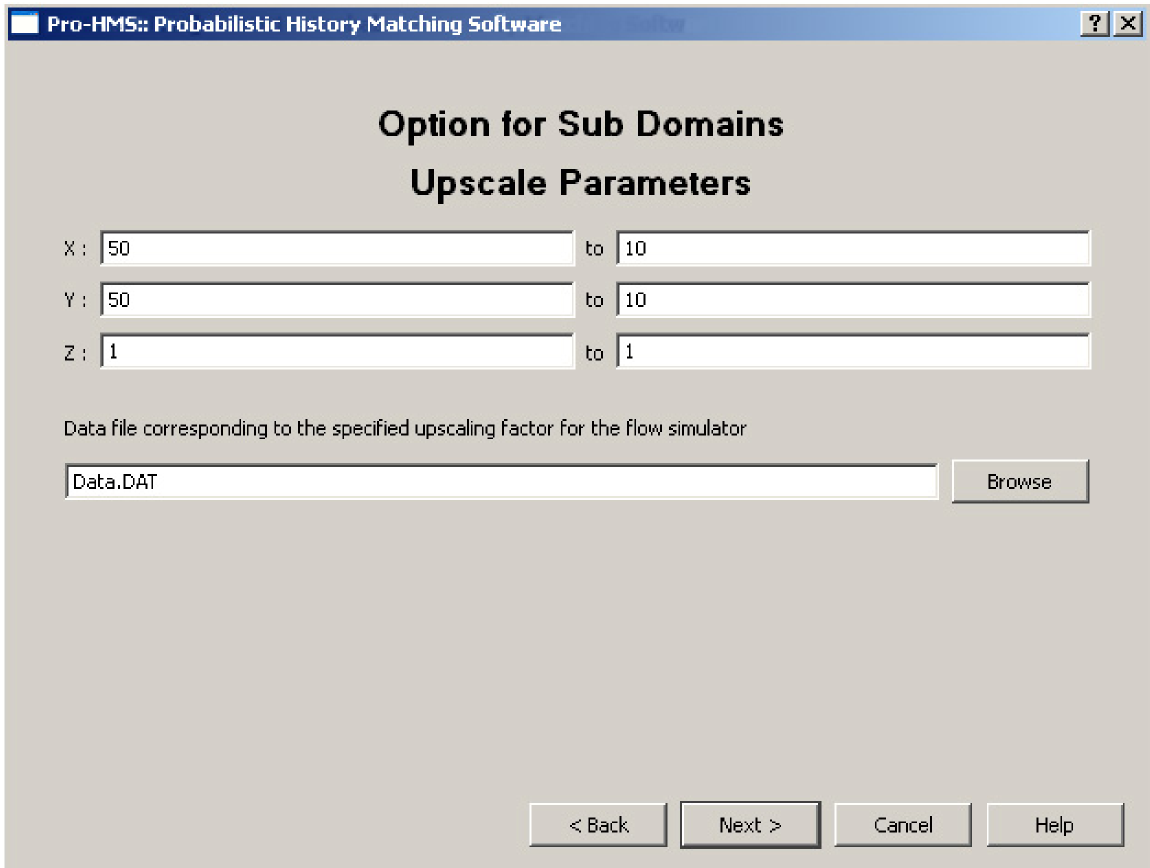


To internally compute the sub-domains, Pro-HMS uses an ensemble of realizations generated by sequential indicator simulation. The number of realizations in the ensemble has to be specified.

Note: The characteristics of the sub-domains could be quite different depending on the size of the ensemble used to calculate the sensitivities. Engineering judgment and prior knowledge of the nature of geological uncertainty have to be used to determine the optimal size of the ensemble.

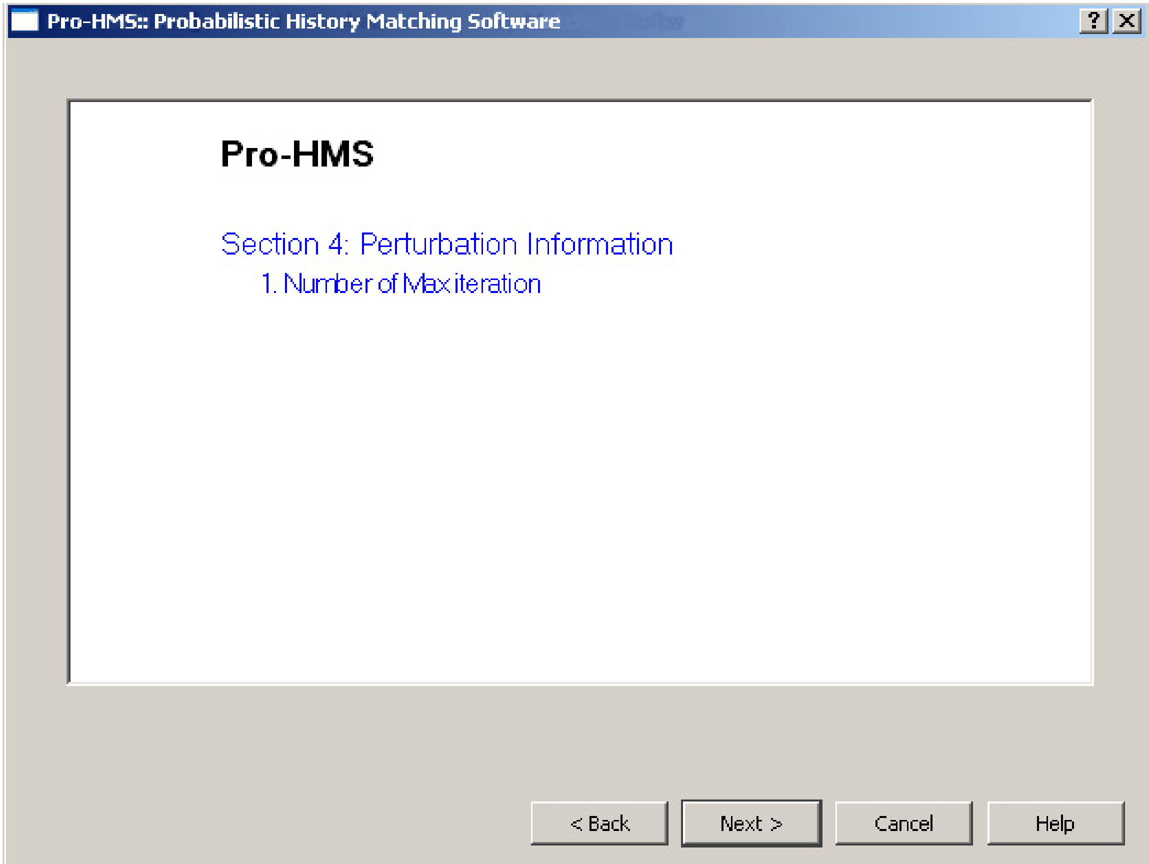
The second option is the duration of the flow simulation to be performed for delineating the zones. For instance, is the production data is available for a duration of 1000days, the user may specify that the sub-domains be delineated on the basis of the full duration of production.

Note: Depending on the nature of the problem and sensitivities of the history match to reservoir parameters, the sub-domains may be delineated on the basis of flow simulations for a reduced duration.

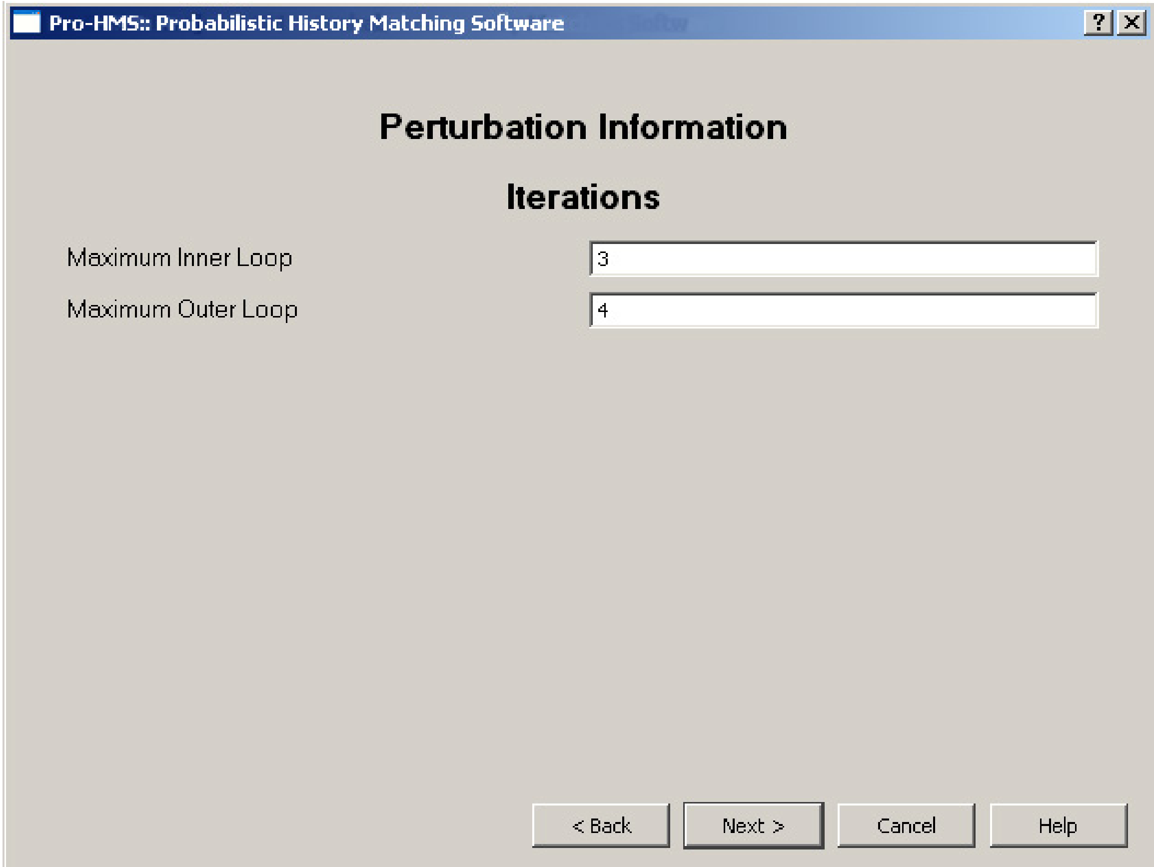


In order to speed up the domain delineation process, the sensitivities are computed using the upscaled ensemble of realizations. The original dimensions of the ensemble realizations as well as the upscaled dimensions are specified.

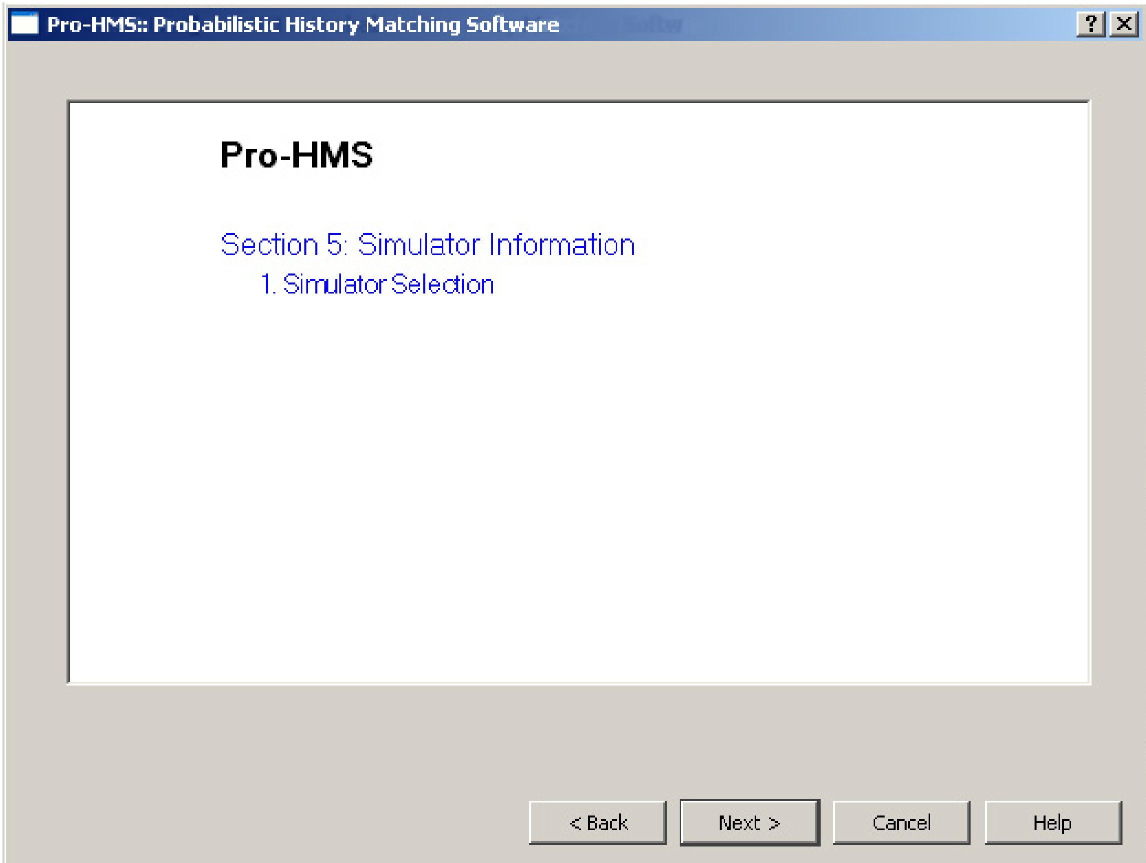
The flow simulation data file is also specified consistent with the upscaled grid specification.



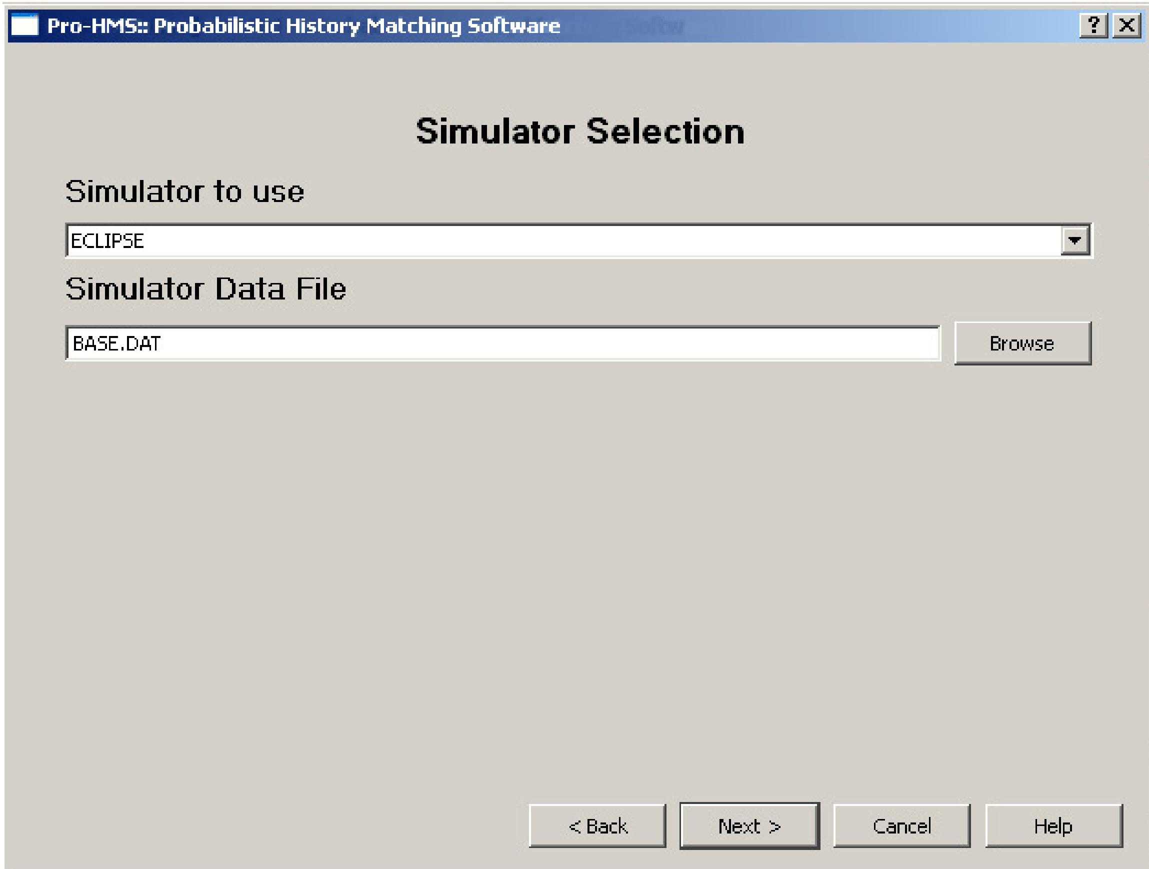
Probability perturbation parameters and options such as the number of inner and outer iterations to be used by the Markov chain process are specified in this section.



The maximum number of inner and outer iterations for the optimization procedure has to be specified. The inner loop is the Dekker-brent loop and the outer loop is for global convergence. The default (based on several trial cases) is 3 for maximum inner iterations and 4 for maximum outer iterations.

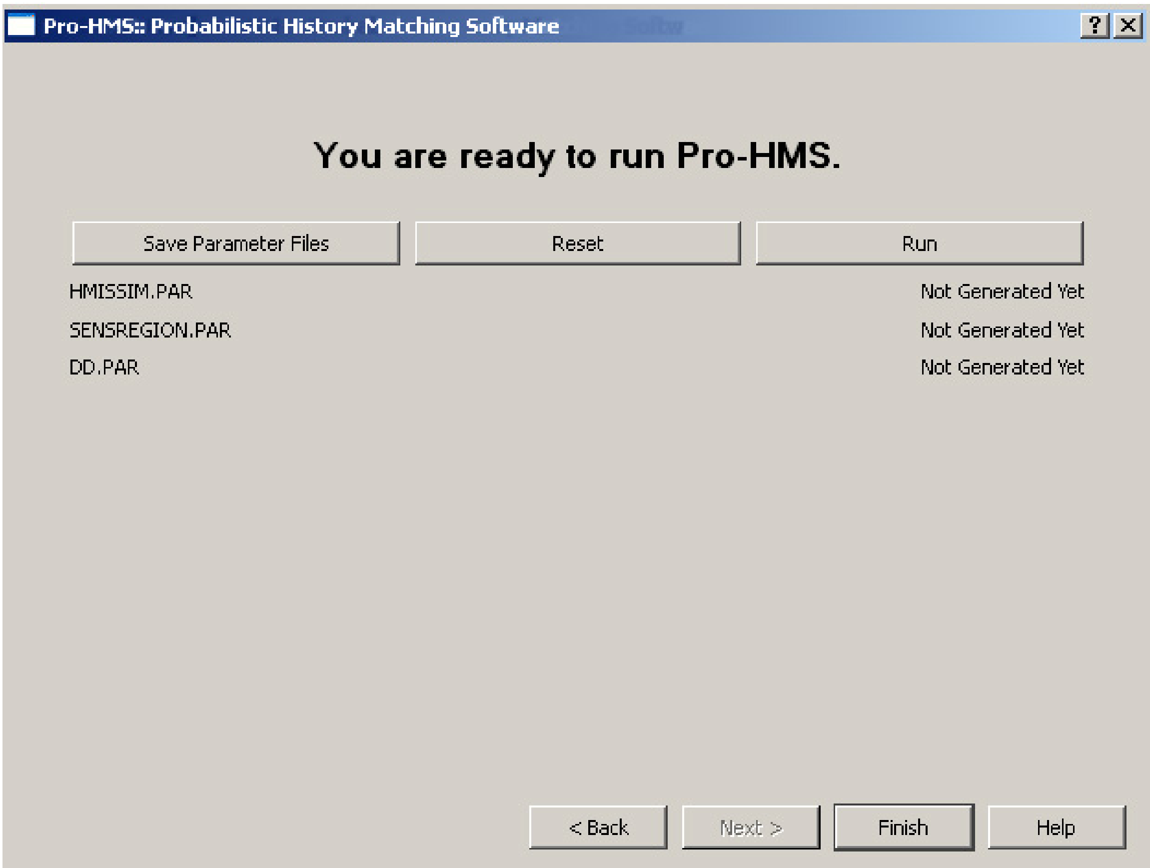


The simulator to be used for the history matching process as well as the simulation data file is specified in Section 5.



Any commercial simulator can be specified provided there is no requirement to execute the simulator only through a dedicated GUI. The corresponding simulation data file has to be input. The selection of simulation parameters to best represent field conditions is at the discretion of the user. The grid dimensions of the reservoir in the data file should be consistent with the grid specification for the indicator simulation models.

Note: A different simulator data file has to be specified for the sub-domain delineation procedure that has the upscaled grid specification.



Click 'Save Parameter Files' button to save the information input using the GUI. Click 'Run' to start the history matching process.