

INTEGRATION OF GEOMORPHIC EXPERIMENT DATA IN  
SURFACE-BASED MODELING: FROM  
CHARACTERIZATION TO SIMULATION

A DISSERTATION  
SUBMITTED TO THE INTERDISCIPLINARY PROGRAM OF  
EARTH, ENERGY, AND ENVIRONMENTAL SCIENCES  
AND THE COMMITTEE ON GRADUATE STUDIES  
OF STANFORD UNIVERSITY  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

Siyao Xu  
March 2014

© 2014 by Siyao Xu. All Rights Reserved.

Re-distributed by Stanford University under license with the author.



This work is licensed under a Creative Commons Attribution-Noncommercial 3.0 United States License.

<http://creativecommons.org/licenses/by-nc/3.0/us/>

This dissertation is online at: <http://purl.stanford.edu/rj323qf2670>

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

**Tapan Mukerji, Primary Adviser**

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

**Jef Caers**

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

**George Hilley**

Approved for the Stanford University Committee on Graduate Studies.

**Patricia J. Gumport, Vice Provost for Graduate Education**

*This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.*



# ABSTRACT

Dynamic process distributing geobodies, defined as geological dynamic processes in this dissertation, are the determinants of spatial heterogeneity of petrofacies, which further affect reservoir performances. But a dynamic process is normally not considered by conventional static reservoir modeling techniques, e.g. two-point statistics, multiple point statistics, and object-based methods. The surface-based method is the only static reservoir modeling technique that explicitly considers a geological dynamic process. Moreover, the surface-based method is an open and flexible framework that is capable of integrating various techniques to generate realistic model realizations. However, current implementations of geological dynamic processes in surface-based models rely on imprecise conceptual rules, which normally involve a large number of empirical coefficients and subjective decisions to quantify qualitative concepts. This does not appear to be an ideal treatment, since workloads in reservoir uncertainty estimation are increased along with empirical coefficients and subjective decisions. The demand of an improved implementation of geological dynamic processes that is more proper for the Monte Carlo uncertainty estimation led to the research in this dissertation. Contributions of this dissertation include a treatment to extract erosion rules from the records of an experiment, a solution to identify a geomorphic experiment to a real depositional system, and a new strategy to implement a geological dynamic process in surface-based models using geomorphic experimental data.

The shale drape coverage in deepwater environments is a key parameter to the sand body connectivity in a reservoir. Proper modeling of shale drape coverage requires understanding of erosion geometries, which must be extracted from changes of intermediate topography of the depositional process. However, intermediate topography is normally not maintained in the stratigraphy of real depositional systems, thus understanding on the erosion geometry is always limited. To study this issue, we developed a workflow based on a geomorphic experiment, where the intermediate status was available in the records. An experiment of a delta basin with recorded intermediate topography was used to demonstrate this workflow. Sequential correlated patterns of erosion and deposition were extracted from intermediate topographies. A surface-based model of lobes and distributary channels was built based on input statistics of experimental erosion-deposition geometries. Since the geomorphic experiment was at a different scale than any real scale depositional system, we proposed to measure the cumulative distribution functions of dimensionless ratios to characterize correlations between erosion depth and deposition thickness in the experiments.

Static reservoir models in the oil industry are always required to be specified to the real scale system of a reservoir. Thus, the selected experiment to provide information for a surface-based model is required to be as similar to the system of the reservoir as possible. However, no hydrodynamic or stratigraphic method has been developed to estimate similarities between an experiment and a specific real depositional system. A solution is provided to identify one experiment that is most similar to a given real system from a set of optional experiments. The solution estimates a similarity between lobe stacking patterns of two systems, which are characterized by the cumulative distribution functions of pairwise lobate proximity measurements. The similarity is estimated with a bootstrap two-sample hypothesis test on the two cumulative distribution functions. Since lobate bodies in experiments can be identified hierarchically from small scales to large scales depending on decisions of the interpreter, the solution also includes an automatic method to quantify lobe hierarchies and to choose lobate stacking patterns at

various scales of interpretation. This solution is applied to estimate the similarity between two delta fan experiments and two published real systems.

Based on statistical similarity analysis, an experimental lobe pattern was identified as the pattern with the highest similarity to a real system. A natural assumption is that the forward process forming identified experimental lobate pattern is also similar to the process forming the given real lobate pattern. In this sense, an experimental lobe pattern is a prior scenario of the spatial-temporal lobe distributing process and is input to the surface-based model. A new surface-based model for a lobate environment was developed, in which the lobe migration mechanism was implemented based on the correlated random walk. The input lobe pattern was treated as a sequence, characterized by statistics of migration distance and migration orientation shifting angle. Both control the random walk behavior of lobe migration. Only three cumulative distribution functions from the input pattern and two empirical coefficients were introduced by the random-walk-based lobe migration mechanism. Compared to tens of empirical coefficients, the new mechanism was more appropriate for uncertainty estimations using a Monte Carlo method. Demonstrations also show that realizations of the new model were hierarchically similar to the input lobe sequence analogous to the similarity between realizations of multiple point statistics and the training image.

# ACKNOWLEDGEMENTS

Turning back and recalling the past four years, I would like to express the most sincere gratitude to my advisor Professor Tapan Mukerji, who has been supporting and guiding me through my time in Stanford. When I came here as a general geoscientist and software engineer, Tapan tolerated my naïve questions with his great patience and led me to the correct path. Tapan is knowledgeable. Every time an obstacle appeared in my research, a conversation with Tapan was the best way to organize my chaotic mind, refine the problem, and locate a good solution. He has also been a good friend when I have had difficulties in life.

I would also like to express my gratitude to Professor Jef Caers, whose active mind is one of the richest sources of ideas. Jef's talk and comments in every seminar introduced his insightful opinions on research, guiding me toward meaningful and promising directions. My work would not be in the current form without Jef.

I would like to appreciate all those I have been talking to through the duration of my Ph.D., including Tim McHargue, George Hilley, and Gary Mavko. Tim's precious experiences from the very frontier of Surface-based modeling and critical comments from George and Gary were the base pillars of successful research.

My appreciation also goes to Professor Chris Paola and his students: Antoinette Abeyta, Sarah Baumgardner, and Dan Cazanacli. Without Chris' data and valuable

responses, our ideas at SCRF would not be implemented. It was my honor to witness and contribute to the beginning of a successful collaboration between SCRF and SAFL.

I also benefitted quite a lot from my three internships at Shell. Names I must mention are Omer Alpak, J-C Noirot, Matt Wolinsky, Long Jin, Tianhong Chen, and Paul Gelderblom. Any work and conversation with you enhanced my understanding about the industry, which was the required background for my research.

I am thankful to my colleagues and friends in SCRF: Andre Jung, Orhun Aydin, Lewis Li (please forgive me for not listing all of them here), from whom I have always received both help and fun. We have built a perfect community for work and life. It is my honor to be a member of SCRF.

I also appreciate my best friend Yang Liu, who was my unofficial mentor. I have learned much from him in research and in stories of this industry, and I will always remember other fun memories we have had, for example, those tickets.

I would like to say thanks to Professor Milton E. Harvey, who granted me selfless help through my two years in Ohio. I will never forget our discussions, covering topics from research to baseball and from politics to food.

I have always been thankful to live and study within the community of Department of Energy Resource Engineering, where you can always find helpful friends.

I dedicate this dissertation to Dr. Ning, my sweetie. I will bear in mind that Christmas, when the primary idea of this work and our wedding came true in Redwood City, the most romantic place in the Bay Area.

This dissertation is also dedicated to my mum, dad, and grandma.

Siyao Xu  
Stanford University  
March, 2014

# TABLE OF CONTENTS

<b>ABSTRACT .....</b>	<b>V</b>
<b>ACKNOWLEDGEMENTS.....</b>	<b>VIII</b>
<b>TABLE OF CONTENTS.....</b>	<b>X</b>
<b>LIST OF TABLES .....</b>	<b>XII</b>
<b>LIST OF FIGURES .....</b>	<b>XIII</b>
<b>CHAPTER 1 INTRODUCTION .....</b>	<b>1</b>
1.1 REVIEW OF STATIC RESERVOIR MODELING ALGORITHMS .....	3
1.2 CHOOSING MODELING ALGORITHMS.....	5
1.3 SURFACE-BASED MODELING AND CHALLENGES .....	8
1.4 THE PROPOSED APPROACH.....	12
1.5 THESIS OUTLINE .....	14
<b>CHAPTER 2 EROSION RULES IN SURFACE-BASED MODELING: GEOMORPHIC EXPERIMENTS AS THE KNOWLEDGE DATABASE .....</b>	<b>15</b>
2.1 INTRODUCTION .....	15
2.1.1 <i>Model Setting: Deepwater Turbidite System</i> .....	15
2.1.2 <i>Erosion in Static Reservoir Modeling</i> .....	17
2.2 METHODOLOGY .....	21
2.2.1 <i>Surface-Based Modeling</i> .....	21
2.2.2 <i>Building a Distributary Channel-Lobe Systems</i> .....	24
2.2.3 <i>Geomorphic Experiments</i> .....	35
2.3 SIMULATION RESULTS .....	43
2.4 CHAPTER SUMMARY.....	47
<b>CHAPTER 3 THE STATISTICAL EXTERNAL SIMILARITY BETWEEN LOBATE ENVIRONMENTS: LINKING EXPERIMENTS TO REAL-SCALE SYSTEMS .....</b>	<b>48</b>
3.1 INTRODUCTION .....	48
3.2 METHODOLOGY .....	50

3.2.1	<i>Experimental Data and Preprocessing</i> .....	51
3.2.2	<i>Stage One – Automatic Lobe Hierarchy Quantification</i> .....	54
3.2.3	<i>Stage Two – Testing Pattern External Similarity</i> .....	67
3.3	APPLICATION.....	75
3.4	CHAPTER SUMMARY .....	83
<b>CHAPTER 4 HIERARCHICAL SIMILARITY OF LOBES: SIMULATION AND COMPARISON.....</b>		<b>85</b>
4.1	INTRODUCTION.....	85
4.2	A LOBE MIGRATION MODEL WITH INPUT LOBE SEQUENCE.....	88
4.2.1	<i>Root Mean Square Distances: The Scaling Factor for Stacking Patterns</i> .....	89
4.2.2	<i>The Bifactor Modeling Scheme</i> .....	90
4.2.3	<i>The Input Stacking Pattern Factor</i> .....	92
4.2.4	<i>The Modeling Algorithms</i> .....	97
4.2.5	<i>Example</i> .....	99
4.3	HIERARCHICAL SIMILARITY .....	105
4.3.1	<i>Quantitative Measure of Hierarchical Similarity</i> .....	106
4.3.2	<i>Implicit Hierarchical Control in the CRW-based Mechanism</i> .....	112
4.3.3	<i>Application of Hierarchical Similarity Control</i> .....	117
4.4	CHAPTER SUMMARY .....	119
<b>CHAPTER 5 CONCLUSIONS AND FUTURE WORK .....</b>		<b>121</b>
5.1	CONCLUSIONS .....	121
5.2	RECOMMENDATIONS FOR FUTURE WORK .....	124
<b>BIBLIOGRAPHY .....</b>		<b>129</b>

# LIST OF TABLES

Table 2.1: Three categories of patterns are interpreted from the sediment event plot 1) channels, where the width of erosion is equal or greater than that of deposition; 2) lobe with erosion, where the width of deposition is greater than that of erosion; 3) lobe without erosion. ....	42
Table 2.2: Dimensionless ratios are used to characterize the relationship between erosion-deposition geometries identified from the sediment event plot. Two probability functions respectively for channel and for lobe with erosion are interpreted. The probability of erosion occurs with lobes is 0.29. ....	43
Table 2.3: Primary simulation parameters. ....	44
Table 3.1: P-values of hypothesis testing between samples. Sample 1 and 2 are detected to be similar, which Sample 3 is different from 1 and 2. ....	74
Table 4.1: Values of primary parameters in the simulation. ....	101

# LIST OF FIGURES

Figure 1.1: Various static reservoir modeling techniques. Along with the improvement of geological realism from two-point statistics to process-based method, the difficulty of conditioning models increases. After Bertoncello 2011. ....	7
Figure 1.2: A combined modeling algorithm. Different algorithms are used at proper stages to maximize their advantages. After Michael et al. 2010.....	8
Figure 2.1: The demonstration of a deepwater turbidite system. The model setting of this work is in the lower fan section of this system (red square). ....	16
Figure 2.2: A demonstration of the hierarchical structure of lobate systems. The nomenclature follows Pr dat et al. 2010.....	17
Figure 2.3: The repetitive appearance of fine grained interlobe layers. After Pr dat, Hodgson, and Flint 2009. ....	19
Figure 2.4: The eroded shale drapes from an outcrop. After Alpak, Barton, and Castineira 2011. .	20
Figure 2.5: The proposed strategy of using geomorphic experiment data in surface-based models. Experimental data are used to extract geometry and depositional rules in Step 4. Steps 1-3 refers to Bertoncello 2011.....	20
Figure 2.6: The general forward workflow for surface-based models. After Caers, 2012. ....	23
Figure 2.7: The template boundary of a distributary channel-lobe geobody interpreted from overhead photos. The channel is interpreted to be constant width and the lobe is interpreted to be an oval. Four centerline control points with geometric or geological interpretation are defined to determine the morphology of a specific geobody. ....	25
Figure 2.8: Parameters controlling template boundary for a lobe. ....	26

Figure 2.9: The channel trend map generated by normalized perpendicular distance from centerline to the boundary. ....27

Figure 2.10: The lobe trend map generated by normalized distance from MPP to the lobe boundary. ....27

Figure 2.11: The concept of positive-negative surfaces for erosion in surface-based models: a) a new lobe is generated and placed onto the previous topographic surface; b) due to erosion, a portion of the new lobe is below the previous topography; c) positive surface: the portion of the new lobe above the previous topography; negative surface: the portion of the new lobe below the previous topography. ....28

Figure 2.12: Two pairs of positive-negative surfaces. The dimensionless ratio of maximum erosion depth versus maximum deposition thickness is  $\frac{3}{7}$ . ....29

Figure 2.13: Parameters to calculate positive-negative surfaces for a) channel and b) lobe using trend map and trend-to-thickness functions in Eq. (2.5) – (2.8). ....30

Figure 2.14: Locations of the centerline control points of a geobody are determined following the order 1) channel source 2) MPP 3) channel turning 4) lobe distal point. ....33

Figure 2.15: The channel source is assumed to be uniformly distributed within a segment around the midpoint of an edge of the modeling grid. ....33

Figure 2.16: The depositional rules for the second control point: lobe Most Proximal Point (MPP). The MPP is drawn from a combined conceptual two-dimensional probability distribution on the modeling grid. The distribution is the weighted mean of three components: a) the next MPP (black dot) tends to be close to the source (star), so for each point in the grid, the probability is proportional to the reciprocal of its distance to a predefined basin source; b) the next MPP (black dot) tends to be close to a previous MPP (blue dot); c) the next MPP (black dot) tends to appear in the region with fewer deposits; d) the final two-dimensional probability map is a weighted mean of the three components. ....34

Figure 2.17: The depositional rule for channel turning. Once the channel source A and lobe most proximal point C are determined, a preset window is placed around the midpoint of Segment  $\overline{AC}$ . The conceptual probability within the window is determined by the normalized reciprocal distance from the previous channel turning G to every cell within the window. The new channel turning B is sampled within the window. ....34

Figure 2.18: The experiment setting for DB-03. The experiment was performed in a tank of five meters by five meters, in which the infeed point was set at a corner of the tank. A prerun

was performed to generate the initial topography, which was a symmetric delta with the radius of 2.5 meters apart from the infeed point. .... 37

Figure 2.19: The intermediate topography of the experiment was measured along three lines at 1.5 meters, 1.75 meters, and 2 meters apart from the infeed point, respectively. .... 37

Figure 2.20: A fraction of the final stratigraphy measured on one topography line, in which the intermediate conditions are partially removed by following events. .... 38

Figure 2.21: The erosion and deposition are identified by computing differences between two successive records. Erosion is defined by reduced elevation between two successive elevation records, while deposition is defined by increased elevation..... 38

Figure 2.22: Examples of generating a sediment event plot from intermediate topography records: a) to identify the relative elevation change regions between two successive intermediate topography records (the upper plot). The relative change geometry is plotted as its temporal order (the lower plot); b) to identify the series of geometries of the elevation change and plot the all geometries along the vertical axis as the order of occurrence. .... 41

Figure 2.23: The sediment event plot of 100 records on the topography line at 1.75 meters apart from the infeed point..... 42

Figure 2.24: The intermediate steps of the simulation for the first four lobes. Two cross sections are used to monitor stratigraphic changes at every step. Erosion does not occur in Lobe 1 while Lobe 2 – Lobe 4 all cause erosion. .... 45

Figure 2.25: A realization of 30 lobes. Two cross sections along and perpendicular to the basin flow orientation are also shown. .... 46

Figure 2.26: A fractions of the experimental and simulated sequential pattern plots. Geometry of erosion and deposition are different, since no three-dimensional data are available to calibrate the trend-to-thickness function. The occurrence of paired negative-positive surfaces for the channel and lobes (with and without erosion) are observed in the simulated result, as expected. .... 46

Figure 2.27: The QQ plots prove the reproduction of cumulative distribution functions of the dimensionless ratios in the model realization..... 47

Figure 3.1: Corrected overhead photos of an experiment taken at regular intervals, in which only one active channel was recorded within the successive photos. Courtesy by Prof. Chris Paola, San Anthony Falls National Lab. .... 53

Figure 3.2: The lobe connected to the latest channel was interpreted on every photo, and a series of lobes are obtained..... 53

Figure 3.3: Parameterization of lobe stacking patterns: a) the source is defined as the MPP on the lobe boundary to the basin source; b) lobe orientation is defined by the unit vector pointing from the source point to the geometric centroid of a lobe; c) polygonal proximity is measured by the Hausdorff distance between two lobes; d) shape proximity is measured through procrustes analysis. ....56

Figure 3.4: Synthetic demonstration of the temporal constraint: a) based on rules in the interpretation, only agglomerations on temporarily successive lobes are legitimate in clustering; b) the sequential cumulated distance from Lobe  $t_1$  to Lobe  $t_3$  is the distance following the sequential route from  $t_1$  to  $t_2$  to  $t_3$ , promising that  $d_{13}$  is always greater than  $d_{12}$  and  $d_{23}$ . .... 62

Figure 3.5: The hierarchical clustering result is visualized by a dendrogram. In the synthetic demonstration, distance  $d_{12} < d_{23} < d_{13}$ , the dendrogram visualizes a lower scale group  $(t_1, t_2)$  and a higher scale group  $((t_1, t_2), t_3)$ . .... 62

Figure 3.6: The dendrogram of a complex dataset. Visualization of large datasets with dendrograms is inconvenient. .... 63

Figure 3.7: The reachability plot as a dendrogram alternative to visualize a large dataset. The height of the bars represent the distance from one observation to all the left neighbors in the plot. The first bar is set to be the maximum height within this clustering. .... 63

Figure 3.8: The threshold intersects with bars greater than the chosen scale of interpretation within a reachability plot. Lobes between two intersections form a lobe at the chosen scale... 64

Figure 3.9: A sample lobe sequence of 36 lobes. Color represents lobe location within the sequence, ranging from 1 to 36. .... 65

Figure 3.10: a) The dendrogram of the sample lobes, the 36 sample lobes are organized by the order of being clustered in the clustering algorithm, and listed along the horizontal axis; b) The reachability plot of the sample lobe; lobes are listed with the same order as in the dendrogram. .... 66

Figure 3.11: Clusters on the exemplary sample of lobes with a selected threshold. Lobes in each group form a lobe at the scale of interpretation. .... 67

Figure 3.12: a) Exemplary set of 60 lobes, which is a super set of the sample lobe set in Fig. 3.9. b) The point process of MPP of lobes in Fig. 3.12 a. .... 69

Figure 3.13:  $G$  functions (normalized using procrustes analysis) of the spatial pattern in Fig. 3.12. .... 70

Figure 3.14: a) Lobe Sample 1 of 13 lobes; b) Sample 2 of 16 lobes; c) Sample 3 of 13 lobes. Sample 1 and Sample 2 are from the same depositional system. .... 73

Figure 3.15: <i>G</i> functions for all three samples, the samples are normalized using procrustes analysis so length measurements from different systems are comparable.....	74
Figure 3.16: The workflow of scale-dependent similarity analysis. The result is a p-value versus a scale plot for each parameter.....	77
Figure 3.17: Lobe patterns used in the application: a) stacking pattern of Exp. A; b) stacking pattern of Exp. B; c) stacking pattern of Kutai basin; d) stacking pattern of Amazon fan. Arrows mark out the overall flow orientations for experiments and the interpreted overall flow orientation for the real-scale cases..	77
Figure 3.18: Hierarchies of experiments: a) hierarchy of Exp. A; b) hierarchy of Exp. B.....	78
Figure 3.19: P-values versus scales of interpretation plots with respect to the Kutai basin. The distance measurements are rescaled with RMSD of the Kutai basin. Peaks of p-values are marked out by the black dots and a narrow range for scales of interpretation with the highest p-values for all four parameters (shaded regions) is identified. a) MPP; b) Orientation; c) Polygonal distance; d) Shape.....	80
Figure 3.20: P-values versus scales of interpretation plots with respect to the Amazon fan. The distance measurements are rescaled with RMSD of the Amazon fan. Peaks of p-values are marked out by the black dots, but the peaks vary in a wide range of scales. a) MPP; b) Orientation; c) Polygonal distance; d) Shape.....	81
Figure 3.21: Stacking patterns from Exp. B at peaks of p-values with respect to the Kutai basin: a) patterns at the scale of interpretation with the highest p-value for the MPP (38 lobes) at the scale of 13.8km; b) patterns at the scale of interpretation with the highest p-value for $\theta$ (35 lobes) at the scale of 15.1km for; c) patterns at the scale of interpretation with the highest p-value for polygonal distance (30 lobes) at the scale of 15.8km; d) patterns at the scale of interpretation with the highest p-value for shape (28 lobes) at the scale of 16.8km. ....	82
Figure 3.22: Stacking patterns from Exp. B at peaks of p-values with respect to the Amazon basin: a) patterns at the scale of interpretation with the highest p-value for the MPP (22 lobes) at the scale of 62km; b) patterns at the scale of interpretation with the highest p-value for $\theta$ (26 lobes) at the scale of 47km for; c) patterns at the scale of interpretation with the highest p-value for polygonal distance (30 lobes) at the scale of 42km; d) patterns at the scale of interpretation with the highest p-value for shape (65 lobes) at the scale of 14km. ....	82
Figure 4.1: a) The plane view of an experimental stacking pattern with the high p-values in Chapter 3, including 38 lobes. b) The sequence view of the experimental stacking pattern, which is explicitly input to the algorithm in this chapter.....	89

Figure 4.2: Root Mean Square Distance (RMSD) as the scaling factor for stacking patterns. RMSD is determined by the averaged squared distance from the geometric centroid to the interpreted basin boundary. It is a measure of the scale of a stacking pattern and is used as the scaling vector to normalize stacking patterns. All length measurements and coordinates from the experiment are normalized by the RMSD of the experimental pattern, and the simulation is performed in the dimensionless space. The RMSD of the field basin can be used to scale the realizations to the basin scale. Refer to Eq. (4.1) – (4.3)..... 90

Figure 4.3: MPP migration and lobe relative (to the basin flow orientation) shifting angle are modeled in this chapter. Scale versus stacking pattern plots of the area distance and the shape are noisy, thus the use of them as inputs requires further studies. a)  $\Delta r$ , the stepwise migrating distance of MPP; b)  $\Delta\theta_{MPP}$ , the stepwise MPP migrating orientation, measured from 0 to  $2\pi$ ; c)  $\Delta\theta_L$ , the stepwise lobe relative (to the basin flow orientation) orientation, measured from 0 to  $\pi$ ..... 94

Figure 4.4: The idea of Correlated Random Walk (CRW). CRW simulates spatial point sequences. The point at time  $t + 1$  is determined by deviations from the point at  $t$  by a moving distance  $\Delta r$  and an angle  $\Delta\theta_{MPP}$  deviating from the previous MPP moving orientation.  $\Delta r$  and  $\Delta\theta_{MPP}$  are modeled as stochastic variables with cumulative distribution functions (CDFs). The choice of CDFs depends on the problem. CRW is applied to model MPP migration in this work and controlled by empirical CDFs from the input lobe sequence..... 95

Figure 4.5: The randomness in CRW. The migration distance  $\Delta r$  is usually modeled by a weibull distribution and  $\Delta\theta_{MPP}$  is modeled by a wrapped cauchy distribution. Both distributions provide parameters to control the means and shapes. Distributions with higher probability to large  $\Delta r$  and  $\Delta\theta_{MPP}$  lead to more randomly distributed point sequences, while higher probability for small  $\Delta r$  and  $\Delta\theta_{MPP}$  lead to sequences consistently migrating in the same direction. .... 97

Figure 4.6: The implementation of the combined lobe migration mechanism: a) the bifactor model combining MPP migration information from the input lobe sequence and the conceptual compensational stacking rule in response to the intermediate topography in the simulation, refers to Algorithm 4.1; b) determining the lobe orientation, refers to Algorithm 4.2..... 99

Figure 4.7: Model grid setting for the test simulation. The basin source sample within a small section around the middle point of an edge (refer to Table 4.1) and the basin flow orientation

were set to be perpendicular to the edge of the basin source toward the modeling grid. ....	101
Figure 4.8: ECDFs from the inputs lobe migration sequence. ....	102
Figure 4.9: Samples of intermediate topographies of a realization with 100 lobes. Several lobe complexes following the compensational stacking rule are observed in the simulation. ....	103
Figure 4.10: One realization of 100 lobes and the proximal, medial, and distal depositional sequence plot. Lateral shifting and spatial-temporal clustering are observed. ....	104
Figure 4.11: The reproduction of input ECDFs of $\Delta r$ , $\Delta\theta_{MPP}$ , and $\Delta\theta_{Lobe}$ . Relative MPP migrating orientation $\Delta\theta_{MPP}$ is not reproduced due to the combination with in-situ compensation stacking rules. ....	105
Figure 4.12: MPP migration sequence of the input stacking pattern: a) the plain view; b) the perspective view. ....	109
Figure 4.13: MPP migration sequence of the realization: a) the plain view; b) the perspective view. ....	109
Figure 4.14: The sequential hierarchies of input and realization. The sequential hierarchy is quantified by the dendrogram with the sequential cumulated distance. The hierarchical similarity is defined by the similarity between two dendrograms. a) The input stacking pattern; b) The realization. ....	110
Figure 4.15 A dendrogram is a weighted binary tree with three components: topology, label, and weights. a) topology: the bifurcation structure between leaves of the tree; b) label: links each leaf to a branch; c) weights: the height at which two branches merge, implying the distance between two clusters represented by the two branches. ....	111
Figure 4.16: Cophenetic distance: distance at which two leaves merge. In the example, the cophenetic distance from Leaf 1 to Leaf 3 is 2.5. ....	111
Figure 4.17: ECDFs of the cophenetic distance from the input lobe sequence and from the simulated realization in Fig. 4.14. ....	112
Figure 4.18: The linkage function, a strategy of calculating the intercluster distance in the agglomerative hierarchical clustering, determines the ECDF of the cophenetic distance. An imprecise relationship exists between the ECDF of the stepwise migration distance and the ECDF of the cophenetic distance from the dendrogram. ....	115
Figure 4.19: The denrograms generated with different linkage functions using the input lobe sequence in Fig. 4.12 and the realization in Fig. 4.13. The visual comparison can easily identify the difference between the dendrograms with different linkage function on the same set of data. However, no significant difference is observed between dendrograms of the	

input and realizations with the same linkage function. The L1-norm differences between the ECDFs of the cophenetic distances of the input and realization dendrograms are calculated for each linkage function: 3758.4 for complete linkage, 1566.6 for average linkage, and 899.3 for single linkage, which is consistent with the theoretic analysis. .... 116

Figure 4.20: First two component multidimensional scaling plot of hierarchies of realizations, demonstrating the effect of an input lobe sequence on constraining the MPP hierarchies of model realizations, which is difficult with the model built in Chapter 2. .... 118

# Chapter 1

## INTRODUCTION

Static reservoir modeling is an information integration procedure that aims to produce geologically reasonable model realizations for dynamic reservoir simulations and, as a result, can be used to estimate reservoir performance uncertainty. The information includes soft and hard data plus qualitative geological knowledge. Hard data are logs and cores measured around wells and provide precise information. Soft data are seismic surveys performed at large regions with coarser spatial and temporal resolution, such that the provided information is considered averaged and smoothed over a unit volume of rocks. Qualitative geological knowledge is provided by field surveys on outcrops analogous to the depositional environment of a reservoir, containing direct field measurements and interpretations based on measurements. The direct measurements include structural and stratigraphic architectures and petrophysical properties. Interpretations provide conceptual models of the spatial distribution of petrofacies in the studied depositional environments or even the ‘dynamic’ depositional process driving the spatial distribution of petrofacies. A static model used for hydrocarbon exploration and development is required to reasonably integrate all

available information such that any features affecting the subsurface flow movement and hydrocarbon production are realistically reproduced.

However, all of the three types of information are limited. The hard data are at very fine scales compared to the reservoir, such that only a tiny fraction of the subsurface is revealed. The scale of the seismic survey is too coarse to describe stratigraphies at the scale of reservoirs. Interpretation may vary from one interpreter to another, creating uncertainty in estimations. In the context of uncertainty estimation and history matching, static modeling algorithms are supposed to be simple and computationally efficient. A simple and fast static model reduces the workload of uncertainty estimations and history matching. In short, the core of static reservoir modeling is to build models as geologically realistic as possible while the model simplicity and computational efficiency are considered. Various static modeling algorithms have been devised in earlier publications, some of which populate petrophysical properties in modeling domains that account for the statistical spatial correlation from the data. Others attempt to use the real underlying physics of depositional processes. However, the choice of modeling techniques is problem-dependent. The focus of this thesis is the surface-based method, which attempts to explicitly emulate realistic spatial distribution in interpreted geometric bodies of petrofacies and concurrently avoid inefficiencies in solving hydrodynamic equations. In Section 1.1, we will review major techniques of static reservoir modeling. Then, in Section 1.2, the general surface-based framework of using different algorithms to generate satisfying static models is described. Finally, in Section 1.3, challenges in surface-based method are discussed, leading to the following chapters of this dissertation.

## 1.1 Review of Static Reservoir Modeling Algorithms

Various techniques have been developed for static reservoir modeling in recent decades. A general trend in the development of static modeling algorithms is the increment of explicit geology accounting in the algorithms and the accompanied increment in algorithm complexity.

Reservoir engineers and geophysicists developed algorithms in the two-point geostatistics toolbox, focusing on the spatial distribution of petrophysical properties important for dynamic reservoir simulation. Spatial correlation of petrophysical properties is characterized by the variogram (Goovaerts 1997), which is, in turn, used to control the simulation. Variogram is the quantitative form of geology in two-point geostatistics and the unit to be simulated is a property value in a cell, so it is impossible to model complex geometries. Two-point geostatistics is simple and efficient, thus it has been widely utilized in the industry. Multiple-point geostatistics (Strebelle 2002; Zhang, Switzer, and Journel 2006; Arpat and Caers 2007) is an improved technique developed by reservoir engineers as well, aiming at simulating more complex heterogeneities. Multiple-point statistics also works on property values in cells. As an alternative to variograms and correlations between pairwise values, multiple-point statistics uses training images to characterize the prior geological information and the cell value correlation is estimated using several surrounding points. A training image is originally a sketch of conceptual spatial distributions of petrofacies. Realizations of multiple-point statistics are visually plausible to the training image, while conditioning remains easy due to the cell-based nature of the algorithm. Advancement in multiple-point statistics can generate realizations with nonstationary training images (Honarkhah and Caers 2012).

The idea of the object-based method (Shmaryan and Deutsch 1999; Deutsch and Wang 1996; Deutsch 2002) is closer to the geologist's perspective than that of the engineer. The method works at the scale of interpreted geometric objects of petrofacies, or geobodies. Geobodies are placed stochastically into the modeling grid with spatial

constraints, such as adjacency of different geobodies or trends of the spatial distribution of geobodies, until a predefined criterion is satisfied. Since geobodies constitute explicit geometric information, complex patterns are produced by the object-based method. However, conditioning of object-based models to hard data is limited to trial-and-error, which is extremely computationally expensive.

Process-based models attempt to solve the problem from the perspective of physical laws. Hydrodynamic equations are used to describe the underlying physics of depositional processes. The process-based model exceeds other methods in terms of realism. However, since the equations are solved numerically over a fine grid in order to provide geological details, process-based models are the most time consuming and not directly applicable to reservoir modeling. Moreover, there is no information to estimate physical parameters of a paleoflow, such as discharged, sediment concentrations that occurred millions of years ago etc., thus realizations of process-based models are normally used as three dimensional general references to improve geological understanding that is not available in real depositional systems.

Similar to object-based modeling, surface-based modeling (Pyrzcz, Catuneanu, and Deutsch 2005; McHargue et al. 2011; Pyrcz and Strebelle 2006; Pyrcz et al. 2012; Michael et al. 2010; Miller et al. 2008) also explicitly models the geometric information of petrofacies to generate complex realizations. Nevertheless, more complex depositional rules are utilized to mimic deposition-erosion processes in a forward scheme, such that the final stratigraphy is plausible to the results of process-based methods. Since no equations are solved, complex stratigraphy can be generated at much lower computational costs with surface-based methods. The surface-based model is a forward modeling method, thus an inversion-based conditioning algorithm (Bertoncello 2011; Bertoncello et al. 2013), which is time-consuming, is required.

## 1.2 Choosing Modeling Algorithms

As indicated in previous sections, a static modeling algorithm is discriminated from others by differences in its strategies to integrate soft data, hard data, and qualitative knowledge. Generally speaking, the realism of static models improves from statistical methods to process-based methods along with the increments of qualitative geological knowledge explicitly considered. For example, geology is quantified by relatively loose controls in variograms or training images at the pixel level, and the underlying hydrodynamics are modeled with equations in more realistic process-based models (Fig. 1.1). Each algorithm has its own advantages and disadvantages, and the choice of algorithm depends on the problem in practical cases.

Usually, there are rules of thumbs to guide the decision making for modelers, which aim at identifying a proper algorithm for a problem. Two-point statistics is the simplest method in the theory, implementation, and usage, thus it is still a must in all commercial geological modeling software packages. Since the method requires a relatively large number of wells and cannot generate complex nonlinear geometries, it is more appropriate to model reservoirs in a relatively small domain within the depositional basin where no complicated nonlinear geometries are detected. Multiple-point statistics produces more complex geometries while conditioning to wells and seismic data remains easy. Therefore, the method can model a wider domain in the depositional system as long as training images are available. Training images significantly increase information available for building a static reservoir model, so multiple-point statistics does not require a large number of wells to generate satisfactory realizations. Object-based and surface-based methods are used in problems where even qualitative descriptions of patterns of the final deposited petrofacies are difficult to obtain. Both methods attempt to utilize more conceptual knowledge of depositional environments, such as geometries of petrofacies, to compensate for the lack of field data. Object-based methods place spatial adjacency constraints on the appearance of different types of geometries, while surface-based methods constrain placement and

morphology geometries with dynamic depositional rules in a forward simulation scheme. Constraints at the scale of geometric information improves the realism of model realizations, yet the simplicity of conditioning is sacrificed. Compared to geostatistical methods, which are constrained at the scale of the pixel, conditioning of object-based and surface-based model realizations is time consuming. Thus they are usually used for cases with a very limited number of wells to condition or are used to generate training images for multiple point statistics. Although both object-based and surface-based methods are based on geometric information, the surface-based method aims at imitating the sequential movement of geobodies forming stratigraphy driven by the underlying hydrodynamic laws. The process-based methods produce very plausible realizations by explicitly solving hydrodynamic equations. However, using hydrodynamic equations dramatically increases the number of parameters of the algorithm and the computational demand. Moreover, most parameters of hydrodynamic equations are physical measures of paleofluid and paleotopographic conditions, which are ordinarily impossible to measure or estimate in the field. Considering that hundreds of hours are required to construct an unconditioned realization with one set of values of parameters, a process-based method is not appropriate for the direct generation of conditioned static reservoir models, where wide ranges of parameter values must be explored for a Monte Carlo uncertainty estimation. As a matter of fact, process-based methods are valuable references from which three dimensional geometry and the relationship between the physical parameters and geometries of petrofacies are measurable. It also improves geological knowledge on depositional environments that are limited in real analogues, such as deepwater systems.

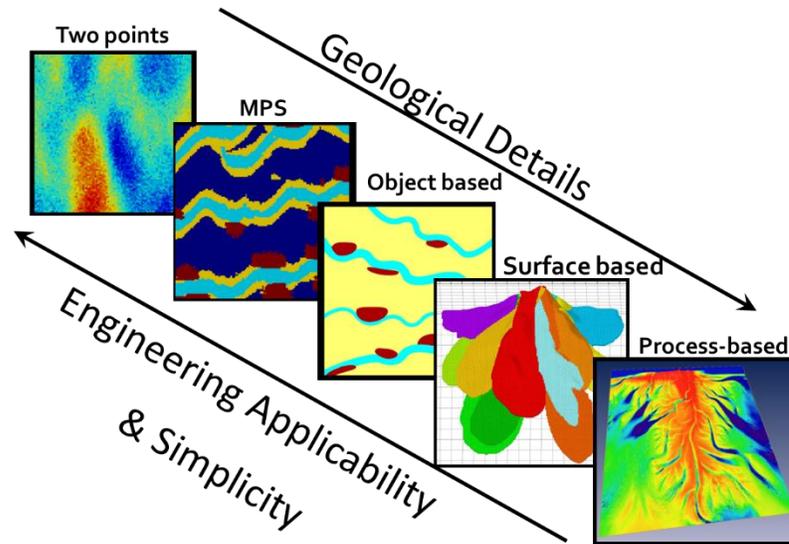


Figure 1.1: Various static reservoir modeling techniques. Along with the improvement of geological realism from two-point statistics to process-based method, the difficulty of conditioning models increases. After Bertonecello 2011.

Besides choosing a single modeling algorithm, a hybrid algorithm is developed to make use of advantages of various methods with application to modeling of distributary channel-lobe systems (Michael et al. 2010). Essentially, the hybrid modeling algorithm belongs to the surface-based category, in which each surface is generated with the object-based method and multiple-point statistics, demonstrating that the surface-based method is a powerful and flexible framework to assemble surfaces generated by various techniques. Fig. 1.2 shows the basic methodology of the hybrid algorithm. Since the process-based method is computationally intensive, only a limited number of realizations are obtained within the period of a modeling project. Thus, a single realization of the process-based method is treated as a knowledge database, from which geometric parametric parameters, e.g. length and width of lobes, are interpreted. Probability distribution functions of interpreted geometric parameters are input to the algorithm. The algorithm starts from sampling a length and a width from input statistics, with which a lobe is generated in the manner of geobodies and placed in the modeling

grid with stochastic rules based on qualitative geological knowledge. The realization with a single geobody is then input to multiple-point statistics as the training image to generate a lobe conditioned to soft and hard data while the lobe length and width of the training image remain the same. The output lobe of multiple-point statistics becomes one surface for the surface-based framework. Iterating the steps generate a series of surfaces, in which a series of lobe objects are conditioned to wells. Finally, a three-dimensional stratigraphic model is obtained by stacking the surface sequence.

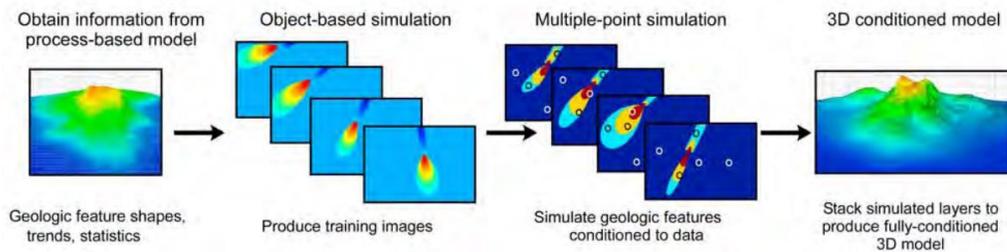


Figure 1.2: A hybrid modeling algorithm. Different algorithms are used at proper stages to maximize their advantages. After Michael et al. 2010.

### 1.3 Surface-Based Modeling and Challenges

From the perspective of pursuing an optimal balance between geological realism and the algorithmic simplicity, the hybrid algorithm is a promising demonstration of the future of static reservoir modeling. The surface-based method plays a crucial role in this framework, which is the key to combining different algorithms. First, the surface-based method fills the gap between the underlying physics and a static reservoir model. As noted above, the process-based method attempts to describe the depositional process by means of hydrodynamic laws. Thus, parameters of the process-based method are physical quantities such as flow discharges, sediment grain composition, or compaction of flow beds, etc. Yet those parameters are not the interest of petroleum geologists and reservoir engineers. Instead of physical parameters, which are impossible to measure for any specific reservoirs, geometric and petrophysical parameters are meaningful in static modeling because they are the possible forms of information from wells, seismic

and geological surveys. Conventional statistical methods, including two-point, multiple-point, and object-based, describe the problem in a manner favorable to reservoir engineering. Parameters of these methods are generally the chosen statistics of either petrophysical properties or the geometries of petrofacies, which is straightforward and simple. However, none of these methods are forward models; therefore, some process-related features sensitive to reservoir production, such as erosion on the thin flow barrier layers (Li 2008; Li and Caers 2011), have to be modeled in additional steps. The surface-based method fills the gap in the sense of using forward simulation with geometric parameters. Second, the surface-based method is a flexible framework to integrate different techniques. As long as a sequence of surfaces are generated and each surface is the function of previous surfaces, any technique can be used to generate the surfaces. In the hybrid modeling example in Section 1.2, multiple-point statistics is chosen to generate a surface since it focuses on improving the performance in conditioning realizations to wells. If placement and in-situ morphology of geometries are controlled by more complex depositional rules based on a conceptual depositional model, the algorithm can also be applied to a statistical hypothesis test on the conceptual model with additional real data, and thus can improve geological understandings (McHargue et al. 2011; Sylvester, Pirmez, and Cantelli 2011). For cases where the conceptual depositional model of geobodies are evaluated to be reliable, realizations of surface-based models can also take the place of process-based models to provide training images in the hybrid modeling algorithm.

Determining the placement and in-situ morphology of geometry with rules from conceptual depositional models creates a new and unique use for the surface-based method, namely the capability of performing a sensitivity study on the parameters of the geological process in reservoir performance. More specifically, a new concept, the geological dynamic process, is proposed in this dissertation, which is a valuable direction for surface-based methods as well. Analogous to the forward dynamic process of hydrodynamic flows used in the process-based method, a geological dynamic process treats the spatial distribution of geometries as a forward and dynamic process. Assuming

a stochastic relationship exists between the location and in-situ morphology of a new geometric body and the intermediate topography, the geological dynamic process is described by the stochastic model. Distinguished from studying the physical process at the scale of hydrodynamic measurements, the geological process studies the dynamics at the scale of interpreted geometric bodies. A geological process is essentially driven by the hydrodynamic process; however, the randomness takes a more significant role than at the scale of hydrodynamics. The randomness is first due to events beyond the scope of hydrodynamics, such as the occurrence of avulsion, the number and volume of sediment events etc. The interpretational uncertainty based on experts' judgments is the second source of randomness, e.g. the shape measures of channel, lobe, bar, etc.

Since conceptual models in geology, including the spatial distribution model of petrofacies (training images) and the dynamic model of depositional process (depositional rules), are the primary source of uncertainty in estimating the reservoir performance, the strategy of modeling a geological dynamic process should be guided by the context of uncertainty estimation and history matching. In a new proposed strategy known as rapid updating of geological models (Caers 2013), the prior geological model, such as training images in multiple-point statistics, will be updated while new wells are inconsistent with the original interpretation. If a surface-based model is applied in the rapid updating, values of parameters controlling the depositional rules will be updated analogously to updating training images in multiple point statistics. Using the lobate environment as an example, parameters controlling the intensity of stacking patterns of lobes should be updated consistently to the additional data. More specifically, the new interpretation of the lobe stacking pattern may be intensely stacking rather than randomly distributed as in the original interpretation. Hence, the design of a model should provide parameters for directly controlling lobe migration, such that the lobe stacking pattern can vary from randomly distributed to intensely clustered simply by varying values of these parameters. The model design is also supposed to be implemented with the simplest techniques possible, such that the uncertainty estimation for a simple model is easier with the Monte Carlo method.

In current surface-based methods, the geological process relies on conceptual depositional models. Examples of conceptual rules include trend maps of appearance (Li and Caers 2011; Bertoncello et al. 2013), migration rates in response to slope (Sylvester, Pirmez, and Cantelli 2011), or frequency of aggradation (McHargue et al. 2011). Conceptual rules can generate realistic stratigraphy; however, the treatment is not ideal for uncertainty estimation, which demands more precise controls. The demand of a more precise method to implement a depositional dynamic process in surface-based models led to the work in this thesis. The primary challenge to implement a precise depositional process is the lack of data. First, records of intermediate topography are required to understand the relationships between intermediate topographies and geometries. However, a large fraction of intermediate topographies have been removed by sediment events, since the erosional process occurs along with the depositional process such that intermediate topographies are not maintained in outcrops in fields. Second, to calibrate a precise algorithm for the dynamic process of geometric bodies would require a sequence of sufficient amounts of two-dimensional or three-dimensional geometric bodies, such that proper parameters and coefficients can be determined for the model. However, sufficient data from field surveys are always a luxury. In this dissertation, a new type of data, geomorphological experiments, are proposed as a knowledge database for the surface-base method, and a quantitative solution is devised to select proper information for a specific real scale depositional system from a few optional geomorphic experiments. Quantitative characterizations and algorithms for a set of geological concepts, such as lobe stacking patterns, lobe hierarchies are introduced, which are not only the basis of building a precise lobe migration model but also useful tools for quantitative studies between lobate experiments for experimental stratigraphers. Finally, a surface-based model for the lobate environment is designed, in which the hierarchy of lobes in a realization honors both the hierarchy of an input sequence of lobes and the conceptual compensational stacking rule. We also introduce a measure to quantify the dissimilarity between lobe hierarchies and demonstrate that hierarchies of lobes in realizations are statistically

similar to the input lobe stacking pattern, which is an effective strategy in the application of uncertainty estimations and history matching.

## **1.4 The Proposed Approach**

Comprehensively controlled small systems, including process-based models and geomorphic experiments, are proposed as a feasible source of obtaining the intermediate topographic information, because measurements can be taken at any spatial and temporal resolution as long as the equipment permits. The geomorphic experiment was chosen to extract erosion rules for surface-based models in this dissertation because more complicated and diversified processes are observed in geomorphic experiments than in process-based models. An experiment of a delta basin, performed in a tank of five meters by five meters, was chosen, in which coarse and fine sediments with flows were injected from a corner to form the delta fan. The intermediate topographic information was recorded in two forms, overhead photos providing two-dimensional information and elevations along three lines perpendicular to the injecting flow orientation. Sequential patterns of erosion and deposition were extracted from the intermediate topographies. A surface-based model of lobes and distributary channels were constructed, in which each depositional event included a negative surface representing erosion and a positive surface representing deposition. Since this geomorphic experiment was not designed to be an analog to any real scale depositional system and length measurements in the experiment were different from real scale systems, the distribution functions of dimensionless ratios of erosion depth versus deposition thickness were measured to control the relation between negative and positive surfaces. The depositional rules were extracted and implemented in the form of conventional conceptual models.

Models in applications of a static reservoir modeling are always required to be analogous to the real system of a reservoir. Yet in conventional applications of experimental stratigraphy, experimental data are usually used to improve understanding

of general sedimentological knowledge, and no method has been developed to specify an experiment to a real depositional system. Specifying an experiment to a real system by hydrodynamic laws is infeasible because physical quantities of paleoflow are not available for any reservoir. Additionally, conventional stratigraphic methods, based on petrophysical properties, also do not work because sediments in experimental systems are usually simplified to two or three categories (e.g. fine and coarse) and not comparable to the real system. Assuming that patterns of sediments in a small system can be similar to that in a large system, defined as statistical external similarity (Paola et al. 2009), a solution is provided to quantify similarities between sediment lobate stacking patterns in the experiment and in the real system, so one experiment that is the most similar to a specific real system is identifiable from a few points. The identified experiment can be used to calibrate algorithms applicable to modeling the given real system. The proposed solution quantifies stacking patterns of lobes with cumulative distribution functions of pairwise lobe proximity measurements, and the similarity between lobe stacking patterns is estimated with bootstrap two-sample hypothesis tests based on a L1-norm difference measure between two cumulative distribution functions. Since lobate bodies in experiments can be identified hierarchically from small scales to large scales, depending on the decision of interpreter, the solution also includes a treatment to quantify lobe hierarchies and to choose the most appropriate scale at which the interpretation should be taken. This solution is demonstrated by comparing two delta fan experiments to two published real systems, and the method successfully identified lobe patterns from one experiment at scales with significantly high similarity to a real system.

Once a lobe pattern is identified, it is used as input to the surface-based model. Another surface-based model is developed, in which a new type of lobe migration mechanism is implemented based on a correlated random walk (CRW). When used as an input, the lobe pattern is treated as a sequence. The input migration sequence is characterized by cumulative distribution functions of migration distance and migration orientation, shifting the angle to control the random walk behavior of the lobe migration.

The simple lobe migration mechanism is controlled only by three distribution functions from the input pattern and two coefficients, which is proper for uncertainty estimations. A quantitative measure of the hierarchical similarity between two lobe patterns was proposed to estimate the reproduction of lobe hierarchy in model realizations. It was proven that lobe hierarchies are implicitly controlled in the new mechanism and all realizations are hierarchically similar to the inputs.

## **1.5 Thesis Outline**

This thesis consists of five chapters. The method of extracting erosion-deposition patterns from experimental intermediate topographic information is demonstrated in Chapter 2. In this chapter, steps of building a surface-based model based on extracted statistics of dimensionless erosion-deposition geometric information are also introduced.

Chapter 3 explains the methodology of identifying a proper experiment and the proper scale of interpretation in this experiment. First, an automatic algorithm to select lobe patterns versus scale is developed based on agglomerative hierarchical clustering. Second, a two-sample bootstrap hypothesis test based on the L1-norm measure of dissimilarity between distribution functions is devised to test the similarity between lobe patterns, which is tailored to small size samples such as lobes. Finally, the method is applied to real datasets of two experimental and two real lobe systems.

Chapter 4 demonstrates a simple strategy of modeling the lobe migration mechanism based on a CRW that is controlled by statistics of the input lobe pattern. The hierarchical similarity measure based on cophenetic distance in the hierarchical clustering result is proposed to validate reproduction of lobe hierarchies in realizations of the new model.

Chapter 5 discusses contributions of this dissertation and suggests potential future research.

# Chapter 2

## **EROSION RULES IN SURFACE-BASED MODELING: GEOMORPHIC EXPERIMENTS AS THE KNOWLEDGE DATABASE**

### **2.1 Introduction**

#### **2.1.1 Model Setting: Deepwater Turbidite System**

The model in this chapter is set to be distributary channels and lobes in the lower fans of a deepwater turbidite system. Deepwater turbidite systems transport sediments from the edge of a shelf through the ocean slope down to the ocean basin. Based on differences of the depositional processes and the resulting stratigraphy, a turbidite system is divided into three sections: upper, middle, and lower fans (Fig. 2.1) from the shelf toward the basin. The upper section features submarine canyons, where sediment events are initiated by a ‘landslide’ on the edge of shelf. The inertial energy of sediments keep increasing due to gravitational acceleration. While additional sediments are

entrained from the bed into the turbidity flow, additional clear water is entrained from the ambient water as well. The intensity of entrainment increases along with the flow energy. Since the flow energy keeps increasing due to the high slope in submarine canyons, this section is dominated by erosion. The middle section starts from the toe of the ocean slope, where accelerated gravitational flow hits the topographic surface with a high gradient in elevation. This section includes well developed channels filled with coarse-grained sediments. In the lower fan section, the turbidity flow weakens until its death due to the low slope. The primary feature of the lower fan is the hierarchical structure of lobes and distributary channels (Fig. 2.2). Lobes include relatively large sand bodies along with shale margins. Along with the thin, interlobe, and shale layers, hydrocarbon reservoirs may exist in the lower fan. Thus the lower fan is of interest to the petroleum industry (Deptuck et al. 2008; Droz et al. 2003; Garland et al. 1999; Gervais et al. 2006; Groenenberg et al. 2010; A. H. Saller et al. 2004; Sullivan et al. 2000; Van Wagoner et al. 2013). Due to the limited understanding of lobe spatial distributions and the architecture of stratigraphy in a deepwater turbidite system, an uncertainty estimation with static reservoir models is an important step at the appraisal stage.

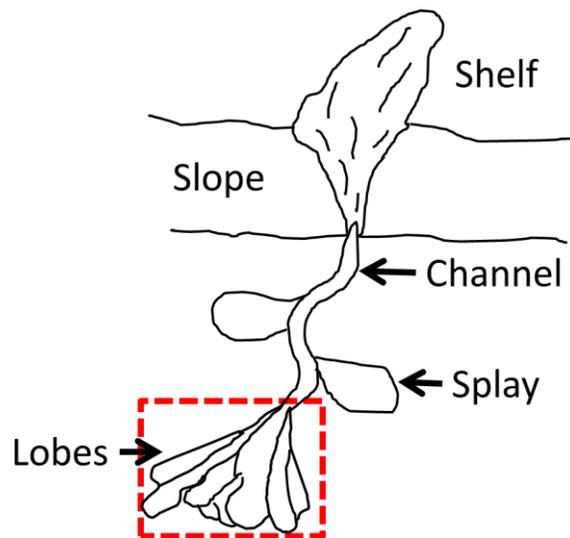


Figure 2.1: The demonstration of a deepwater turbidite system. The model setting of this work is in the lower fan section of this system (red square).

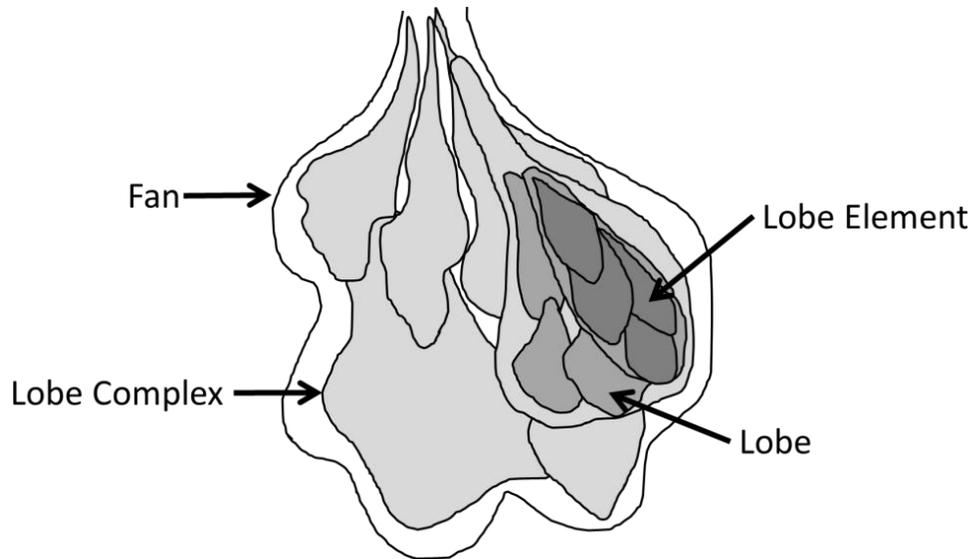


Figure 2.2: A demonstration of the hierarchical structure of lobate systems. The nomenclature follows Prat et al. 2010.

### 2.1.2 Erosion in Static Reservoir Modeling

The interlobe fine grained layer (Fig. 2.3) formed by either allogenic or autogenic processes in the lobate environment is a significant feature of deepwater systems. Fine-grained layers are observed interbedded with sand bodies of lobes in the stratigraphy. Since the fine grain layers are normally assumed impermeable and laterally continuous all over the basin area, the existence of a shale layer between two lobe sand bodies works as a flow barrier in hydrocarbon reservoirs. The thin and fine grain layers may be eroded by successive sediment events, forming holes on the thin flow barriers, and causing direct contact between the sand bodies of two lobes. Therefore erosion may significantly change the production performance of a reservoir (Fig. 2.4).

Modeling erosion on fine grain layers is challenging. Erosion is conceptually modeled as holes in the subseismic scale fine-grained layers. The spatial distribution of eroded holes is a complicated nonlinear function of both successive sediment events and intermediate topographies. However, obtaining sufficient data on paleoflow and

paleointermediate topography from field studies or seismic surveys is difficult. First, only limited vertical two dimensional data of erosion are available in outcrops, which is insufficient to build three dimensional erosion geometry and three dimensional geobodies. Second, the subseismic scale fine-grained layers, ranging from tens of centimeters to several meters, are not detected by seismic surveys; therefore, it is impossible to quantitatively characterize the spatial distribution of eroded holes from the seismics. Finally, the intermediate paleotopography is not maintained in any outcrops; therefore, the relationship between spatial distribution and erosion cannot be quantitatively characterized. Since the measurement of intermediate topographies in active real depositional systems is also infeasible, a new source of information, where intermediate topographies are easily available, must be used. One possible source is the process-based model. Intermediate topographies can be recorded at any temporal interval of the process-based model, so erosion geometry and spatial distribution of eroded holes can be quantitatively estimated. However, the geomorphic experiment is used in this dissertation because more complicated patterns of sediments are observed in geomorphic experiments than in process-based models.

Using the surface-based method is appropriate for modeling spatial distribution of erosion on fine-grained layers. Compared to conventional statistical methods, the surface-based method is a forward model, which is naturally proper to emulate the sequential forward process that distributes lobes in the basin. Since the surface-based model works at the scale of the interpreted geometry of petrofacies, complex nonstationary stratigraphy can be easily generated and the uncertainty of geometric parameters can more easily be reduced than the physical parameters in process-based models. Moreover, proper conceptual depositional rules in surface-based models are computationally efficient, while the realism of model realizations is as plausible as process-based methods.

The surface-based method still requires understanding of three-dimensional erosion geometry and intermediate topographies, which are characterized using the

geomorphic experimental data in this chapter (Fig. 2.5). Generally speaking, geomorphic experiments are simplified, small depositional systems controlled and operated by experimental stratigraphers. Since the experiments are comprehensively controlled, the initial conditions and the intermediate states of the processes are recorded; therefore, the geometry of erosion can be calculated. A method of using intermediate topographic information from a geomorphic experiment was developed. Here, the overall workflow of surface-based modeling followed with a detailed description of building the model is introduced. A geomorphic experiment of a delta fan system, along with a treatment of visualizing and characterizing erosion and deposition geometry, are also introduced. The erosion geometric information is measured with dimensionless ratios so it is rescalable to a real system. Erosion in the surface-based model is controlled by statistics of the dimensionless ratios.

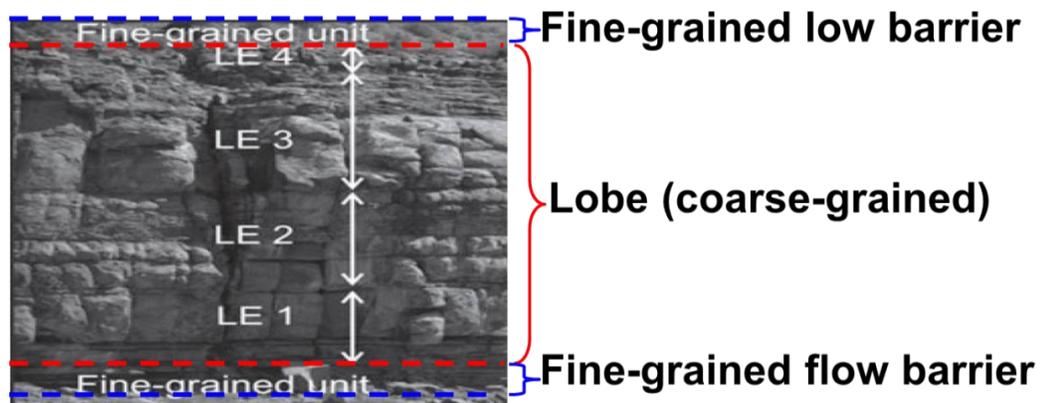


Figure 2.3: The repetitive appearance of fine grained interlobe layers. After Pr dat, Hodgson, and Flint 2009.

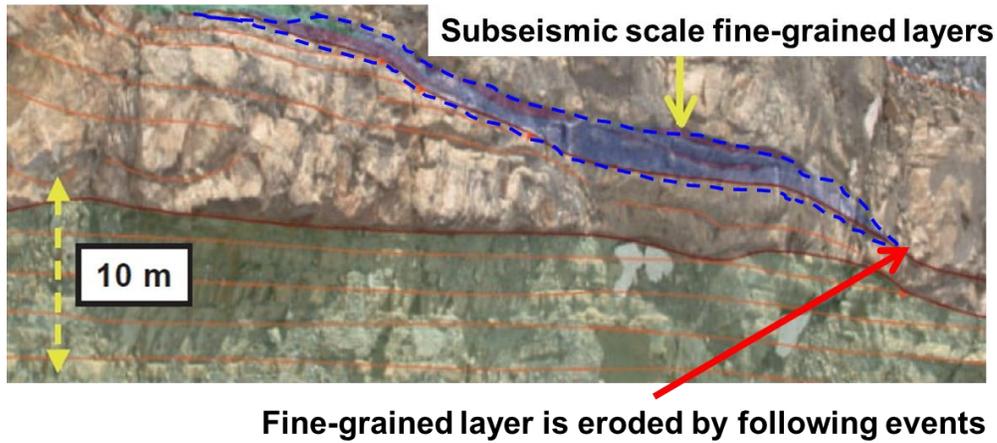


Figure 2.4: The eroded shale drapes from an outcrop. After Alpak, Barton, and Castineira 2011.

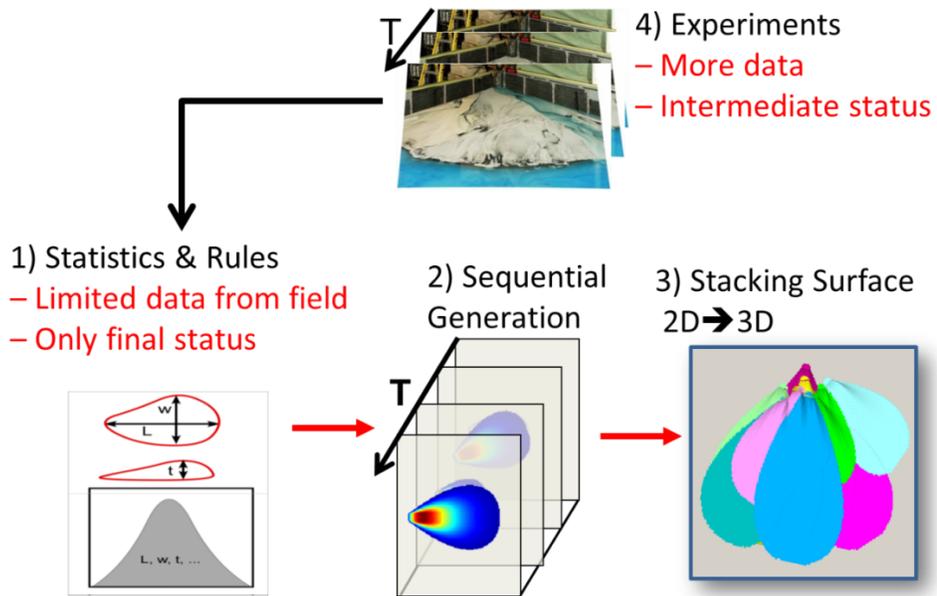


Figure 2.5: The proposed strategy of using geomorphic experiment data in surface-based models. Experimental data are used to extract geometry and depositional rules in Step 4. Steps 1-3 refers to Bertonecello 2011.

## **2.2 Methodology**

In this section, a general workflow of a surface-based model (Section 2.2.1) and the details of generating geometry and making the depositional rules (Section 2.2.2) are presented. Finally, a treatment of characterizing erosion and deposition rules for the surface-based model using delta fan geomorphic experiment data is presented. The simulation results demonstrate that the surface-based model, coupled with proper geomorphic experiment data, is a feasible direction for improving the surface-based method (Section 2.2.3).

### **2.2.1 Surface-Based Modeling**

As noted in Section 1.2, the basic idea of surface-based modeling is to imitate with conceptual rules the geological depositional process driven by the underlying physical laws of the turbidity current. The geological depositional process is defined as the forward dynamic process driving the spatial distribution of geobodies. Different from the process-based model through the perspective of physics, such as turbidity current and sediment grains etc., the geological dynamic process considers the problem through geobodies interpreted by geologists from petrofacies. The term ‘geobody’ refers to lobes and channels in the lobate environment, which are at a higher spatial and temporal scale than flow and particle. A geobody is the interpreted geometry of sediments related to regular geometry formed by flow and sediment; therefore it reflects underlying physics as well. Depending on the depositional environments and the scale of interest, the geobody can represent any geometry within the hierarchy. In the case of this dissertation, the lobe is the level of our interest in the hierarchy of lobate geometries. Because uncertainties are involved with the interpretation of geometries and rules, depositional rules should reflect the general physical laws and account for uncertainties of rules as well. Therefore, the design of geological processes should be stochastic rather than deterministic. They are normally obtained from conceptual geological understandings about depositional processes. In most surface-based modeling algorithms, only conceptual depositional rules are implemented because of the lack of

precise quantitative measurements on the intermediate conditions of the geological dynamic processes.

Surface-based models have been developed for different depositional environments (Bertoncello et al. 2013; McHargue et al. 2011; Pyrcz, Catuneanu, and Deutsch 2005; Pyrcz et al. 2012; Pyrcz and Strebelle 2006). Generally, the model is constructed based on a sequential forward workflow (Fig. 2.6). In the first step, a template of geometric boundary is designed to represent the general shape in the depositional system to be modeled. The template boundary is designed to be controlled by parameters with geological interpretation, such as lobe length or width. Values of geometric parameters of a template boundary are controlled by input statistics at each step of the simulation. The morphology of every geobody varies based on depositional rules. In current surface-based models, the morphology of geobodies is controlled by centerline key points. A key point is a geologically or geometrically meaningful point on the centerline of a template boundary, such as the most proximal point of a lobe or a point on a channel centerline. Locations of key points are determined based on the conceptual depositional rules such as compensational stacking. Key points determine a centerline, and a boundary is determined by fitting the template boundary along a centerline. Once a boundary is obtained, a property trend map is generated over the geobody, representing the conceptual model of petrofacies within a geobody, such as thickness, petrofacies, or petrophysical properties. With the trend map and a defined trend-thickness function, the geobody thickness map is calculated. The current topography is updated by adding the latest geobody thickness map on the current topography, which is then used as topography for the next step. This procedure is repeated until a predefined criterion, such as a top surface or a volume from the seismic survey, is satisfied.

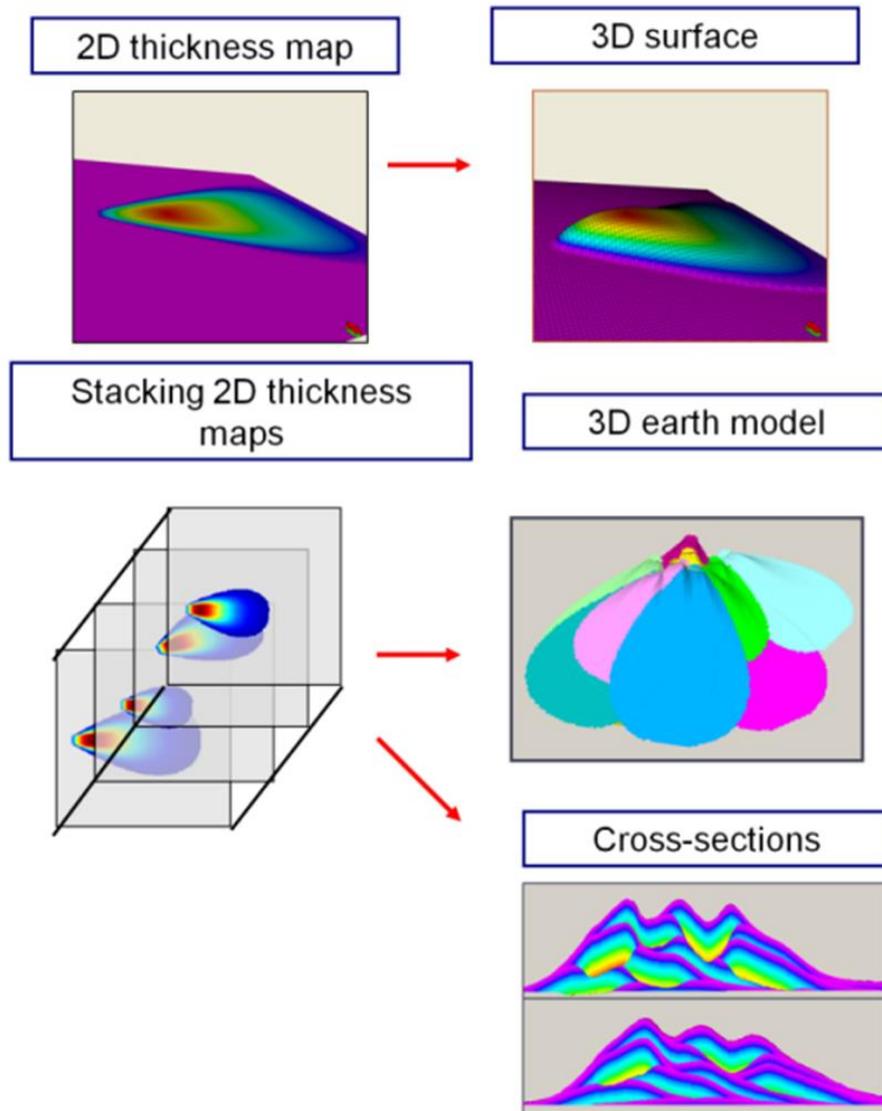


Figure 2.6: The general forward workflow for surface-based models. After Caers, 2012.

### **2.2.2 Building a Distributary Channel-Lobe Systems**

A geobody is generated through a three-step procedure including template boundary, trend maps, and geobody surfaces. The modeler determines the geometry of geobodies and the depositional rules through close collaborations with considerations on algorithm complexity, and purposes of the model. The geometry should be algorithmically flexible to be adjusted according to intermediate topographies.

#### **Template Boundary**

A template boundary is defined as the edge of a geobody in the horizontal plain. Based on overhead photos recorded in the experiments, ovals represent the general boundary of the lobe. The distal half of a lobe is slightly wider than the proximal. The distributary channel is interpreted as a constant width and with light sinuosity (Fig. 2.7). Four centerline control points, all of which are geologically interpretable, are used to vary the in-situ morphology of a geobody based on intermediate topographies. The control points are the channel source, the channel turning, the most proximal point, and the lobe distal point (Fig. 2.7).

The control points are defined as follows.

- The channel source point is defined as the most upstream point of a geobody. Conceptually, the sediment flow starts from the channel source for a geobody.
- The lobe most proximal point (MPP) is defined as the transition point from the channel part to the lobe part, which is the most important point in this model.
- The channel turning point is defined as the location where maximum curvature is present. Since distributary channels in the experiments show low sinuosity, only one turning point is set on the centerline between the channel source and the MPP.

- The lobe distal point is defined as the most downstream point on the centerline of a geobody.

Locations of the centerline control points are determined by depositional rules with a geologically reasonable order (details in the following paragraphs), and the centerline is calculated by fitting a spline curve on the control points.

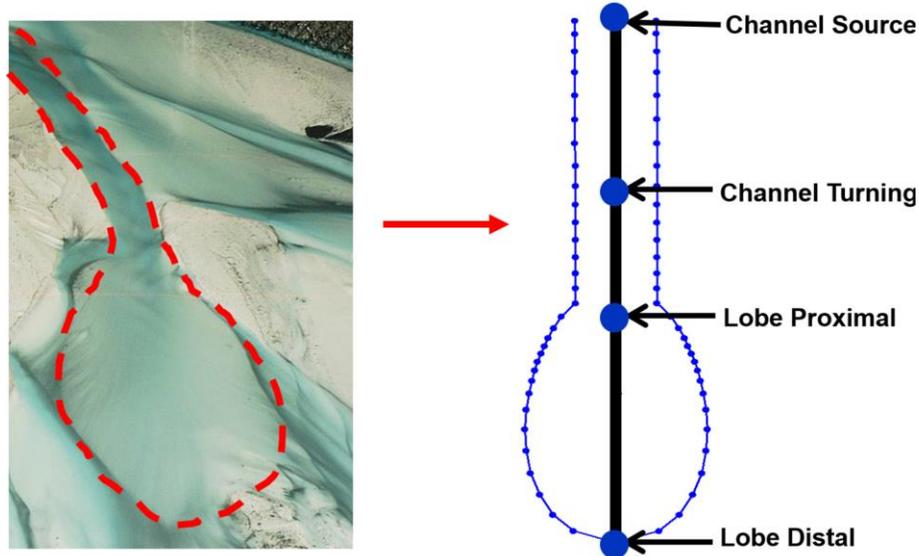


Figure 2.7: The template boundary of a distributary channel-lobe geobody interpreted from overhead photos. The channel is interpreted to be constant width and the lobe is interpreted to be an oval. Four centerline control points with geometric or geological interpretation are defined to determine the morphology of a specific geobody.

The template boundary is calculated along the centerline for the channel part and the lobe part, respectively. Since the centerline morphology will be varied by depositional rules for every specific geobody, geometric equations of the template boundary are designed to be sufficiently robust, such that the boundary points can be calculated along a curved centerline. The lobe boundary is defined by an oval shape equation. Given an arbitrary lobe centerline point  $P_c(x_c, y_c)$  apart from  $MPP(x_0, y_0)$  toward the downstream direction at a normalized distance  $d_p$  along the centerline, a related boundary point  $P_b(x_b, y_b)$  is either of the two points whose perpendicular distance to  $P_c$  is defined by Eq. (2.1). The perpendicular distance from channel

boundary points to channel centerline points are set to be constantly equal to the lobe width at MPP ( $d_p = 0$ ).

$$\frac{w_p}{2} = L * \sqrt{\frac{(f_w)^2 * A}{4B}} \quad (2.1)$$

$$A = 1 - \left[ \frac{\bar{d}_p + \frac{c_2 - 1}{2}}{\frac{1}{2}(1 + c_2)} \right]^2 \quad (2.2)$$

$$B = e^{c_1[\bar{d}_p + \frac{c_2 - 1}{2}]} \quad (2.3)$$

where  $L$  is the input lobe length,  $f_w = \frac{w_{max}}{L}$  is the ratio of lobe width versus lobe length;  $d_p = \frac{\sqrt{(x_p - x_0)^2 + (y_p - y_0)^2}}{L}$  is the normalized distance from point  $P$  to MPP, ranges from 0 at MPP to 1 at Lobe End;  $c_1$  and  $c_2$  are built-in constants determining the oval geometry.

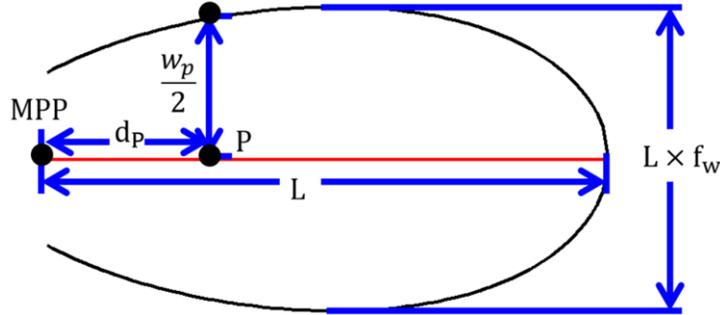


Figure 2.8: Parameters controlling template boundary for a lobe.

## Trend Maps

Besides the two-dimensional boundary, other location-related features of a geobody, such as the spatial distribution of thickness and petrophysical facies, are conceptualized to be functions of the geometry. Trend maps generated from geometry describe conceptual models of spatial distribution. Two trend maps are generated respectively for the lobe and the channel. Since properties of a channel, such as deposit

thickness, generally decrease from the centerline to the boundary, the channel trend map is generated by calculating the normalized perpendicular distance from the centerline to the channel boundary (Fig. 2.9). Lobe thickness varies gradually from MPP to the boundary; therefore, the lobe trend map is generated by calculating the normalized distance from MPP to the lobe boundary (Fig. 2.10).

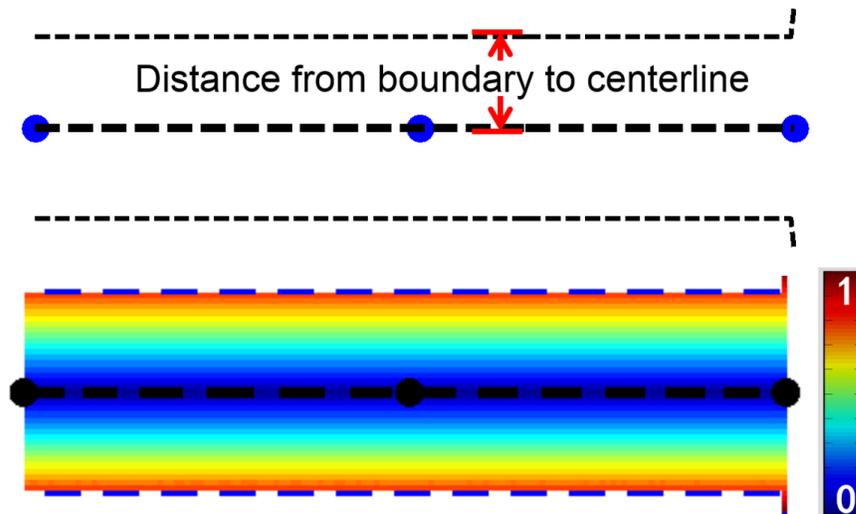


Figure 2.9: The channel trend map generated by normalized perpendicular distance from centerline to the boundary.

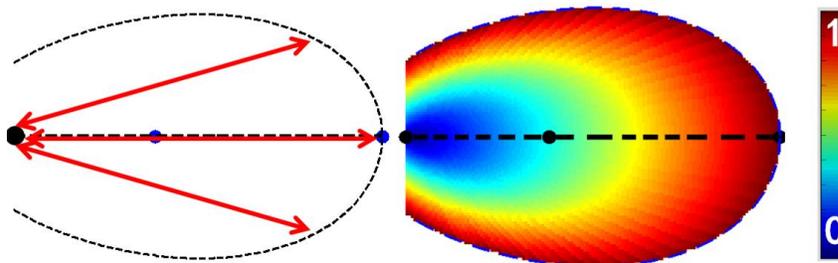


Figure 2.10: The lobe trend map generated by normalized distance from MPP to the lobe boundary.

## Geobody Surface

For modeling erosion of lobes in surface-based methods, the concept of positive and negative surface pairs is introduced (Fig. 2.11). If deposition is the only process considered, a new surface is generated by adding a thickness map of the geobody onto previous topography, so the increment of the surface is equal to the thickness of the latest lobe. However, when erosion is considered, sediments may be removed as well as deposited; therefore, a fraction of the new geobody will be below the previous topography and the increment of the topography is less than the lobe thickness. In the concept of positive-negative surface, the portion of a new lobe below the previous topography is modeled as a negative surface and the portion above the previous topography is the positive surface. The net-lobe thickness is calculated as the summation of the absolute values of the negative surface and the positive surface. The surfaces are generated with the same trend map and a trend-to-thickness function. The negative surface depth is correlated to the positive thickness with a dimensionless ratio. Two examples of lobe surfaces are demonstrated in Fig. 2.12.

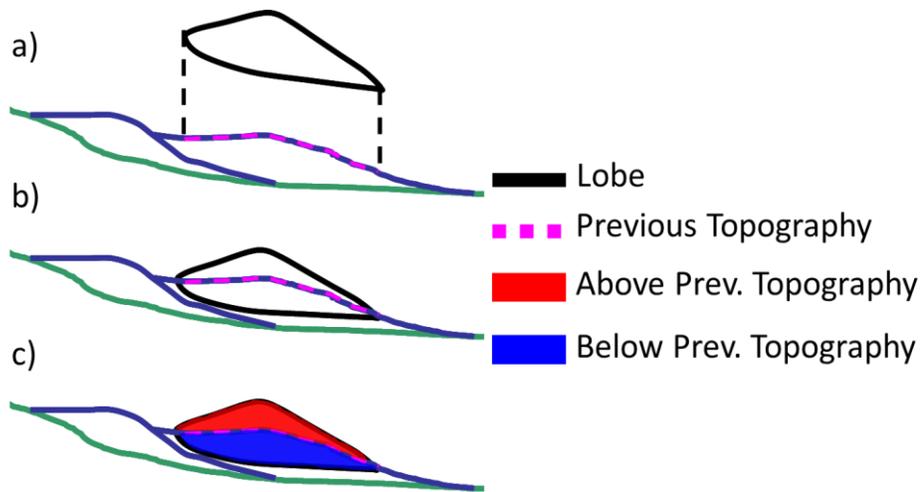


Figure 2.11: The concept of positive-negative surfaces for erosion in surface-based models: a) a new lobe is generated and placed onto the previous topographic surface; b) due to erosion, a portion of the new lobe is below the previous topography; c) positive surface: the portion of the new lobe above the previous topography; negative surface: the portion of the new lobe below the previous topography.

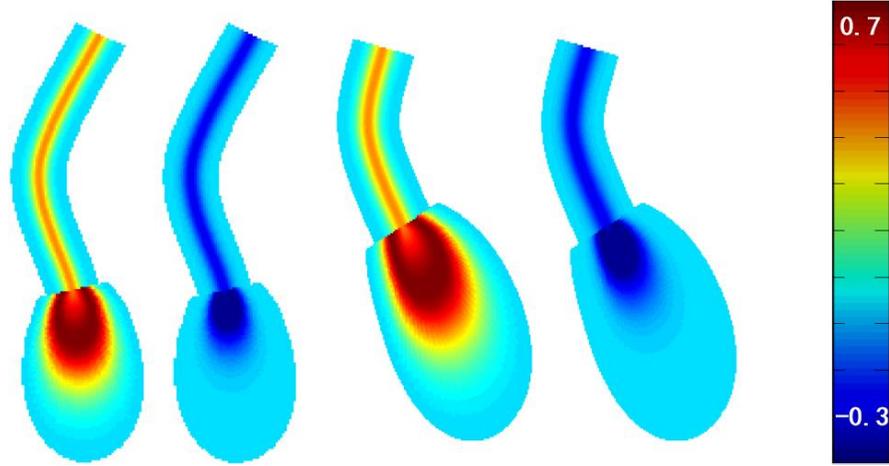


Figure 2.12: Two pairs of positive-negative surfaces. The dimensionless ratio of maximum erosion depth versus maximum deposition thickness is  $\frac{3}{7}$ .

The thickness of the positive surface and the depth of the negative surface are calculated by two sets of trend-to-thickness functions. In Fig. 2.13, a and b demonstrate parameters to calculate channel and lobe thickness maps. Thickness of the positive surface of the channel is computed in Eq. (2.4).

$$h_p = h_{1max} * e^{-\frac{d_p^2}{2s^2}} \quad (2.4)$$

where  $h_p$  is the thickness at an arbitrary point  $P$  within the channel;  $d_p$  is the normalized perpendicular distance from  $P$  to the channel centerline obtained from the channel trend map;  $h_{1max}$  is the maximum thickness of the channel positive surfaces from input; and  $s$  controls the decreasing rate of thickness from the centerline to the boundary. The negative surface can be generated from the positive surface by replacing  $h_{1max}$  with  $h_{2max} = h_{1max} * f_{p2n}$ , where  $f_{p2n}$  is a negative ratio of the channel positive thickness versus the channel negative depth. The lobe surfaces are generated by Eq. (2.5) – (2.8).

$$h_{1P} = h_{1max} * e^C \quad (2.5)$$

$$C = \frac{-(d_P - d_{1Depo})^2}{2s^2 * f_{skew}^2} \quad (2.6)$$

$$f_{skew} = \sqrt{-\frac{d_{1Depo}^2}{2s^2 * D}} \quad (2.7)$$

$$D = \log\left(\frac{h_{channel}}{h_{1max}}\right) \quad (2.8)$$

where  $h_{1P}$  is the thickness in the positive surface at an arbitrary point  $P$  within the lobe area;  $d_P$  is the normalized distance obtained from the lobe trend map;  $h_{1max}$  is the maximum thickness of the lobe positive surfaces from inputs;  $s$  is decreasing ratio from the maximum thickness point to zero on the lobe boundary;  $h_{channel}$  is the channel thickness; and  $d_{1Depo}$  is the distance from MPP to the thickest point of the lobe positive surface along the lobe centerline. The negative lobe surface is computed by multiplying a correlation ratio on the positive lobe surface.

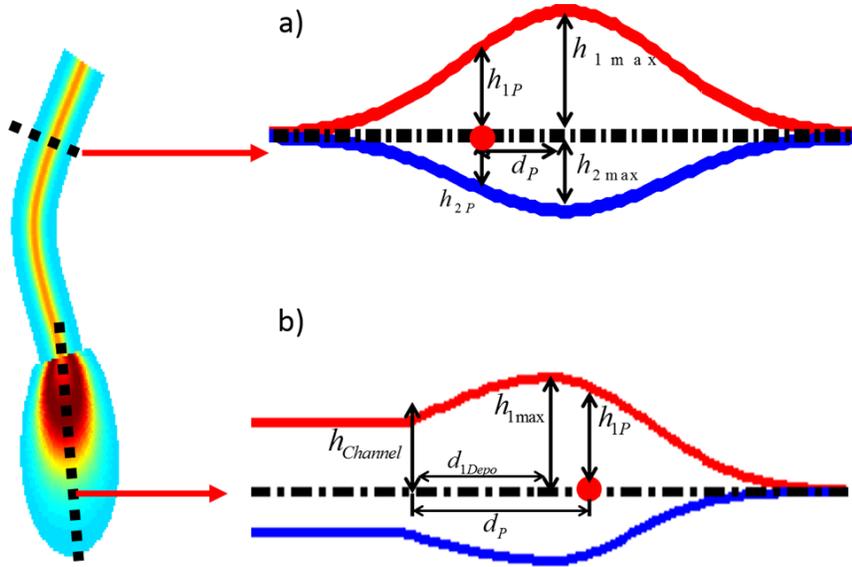


Figure 2.13: Parameters to calculate positive-negative surfaces for a) channel and b) lobe using trend map and trend-to-thickness functions in Eq. (2.5) – (2.8).

## Depositional Rules

The conceptual depositional rules used in this chapter include 1) the order to determining locations of the centerline control points of the geobody, and 2) specific rules to determine every control point. At every step, locations of the four centerline control points should be identified based on the intermediate topography and conceptual rules for each points. The determination of locations of the centerline control points follows the order 1) channel source 2) MPP 3) channel turning point 4) lobe distal point (Fig. 2.14). This order attempts to mimic the physical reasonable behavior of a sediment event. The placement of a new geobody starts by finding the location of the channel source point, which is the initiation point of a sediment event entering the modeling domain. Since behaviors of geobodies within a distributary channel-lobe system are normally featured by the occurrence locations of lobes, the MPP is the second point to be determined in the model. The vector  $\overline{AC}$  (Fig. 2.14) represents the forward direction of a sediment event, thus it is conceptually interpreted as the overall orientation of flows forming this geobody. Because the channel turning point reflects random perturbations around the overall orientation, it is statistically simulated within a preset range surrounding the midpoint of segment  $\overline{AC}$ . Specific rules for every control point are presented as follows:

1) Channel source: the channel source is assumed to be uniformly distributed within a segment located on the edge of the modeling grid (Fig. 2.15).

2) MPP: The new MPP location is determined using an artificial two-dimensional probability distribution covering the whole modeling grid. The two-dimensional distribution consists of three conceptual components. The first component accounts for the conceptual understanding that a lobe is more likely to occur in the same vicinity as the basin source (Fig. 2.16 a). Once a channel source is determined, distance from every cell in the grid to the source is calculated. Probabilities of the source components are modeled as the normalized inversed distance. The second component accounts for

another truth that the location of a new MPP tends to be proximate to the previous MPP (Fig. 2.16 b). Similar to the source component, probabilities of the previous MPP component are computed with the inverted distance to the previous MPP. The last component accounts for the impacts of compensational stacking (Fig. 2.16 c). The compensational rule reflects the physical truth that sediments tend to be deposited in regions with low relief. Probabilities of the compensational rules are converted using the normalized reciprocal of the surface relief. The final artificial two-dimensional probability distribution function is a weighted mean of the three components, representing the relative likelihood of a pixel to be the next MPP.

3) Channel turning: The interpretation of this rule is that a new channel avulsion point tends to be close to the previous avulsion point (Fig. 2.17). Once the channel source and MPP are determined, a preset window is placed around the midpoint of segment  $\overline{AC}$ . Probabilities of cells within the window are determined by the normalized inversed distance from each cell to the previous channel turning.

4) Lobe distal point: Since the power of the sediment flow starts to weaken from the MPP toward the downstream direction, segment  $\overline{BC}$  is assumed to represent the mean direction of the lobe (Fig. 2.14). The exact lobe orientation represented by vector  $\overline{CD}$  is sampled from a predefined gaussian distribution with  $\overline{BC}$  as the mean in order to incorporate some randomness. The coordinate of D is calculated with vector  $\overline{CD}$  and the given lobe length  $L$ .

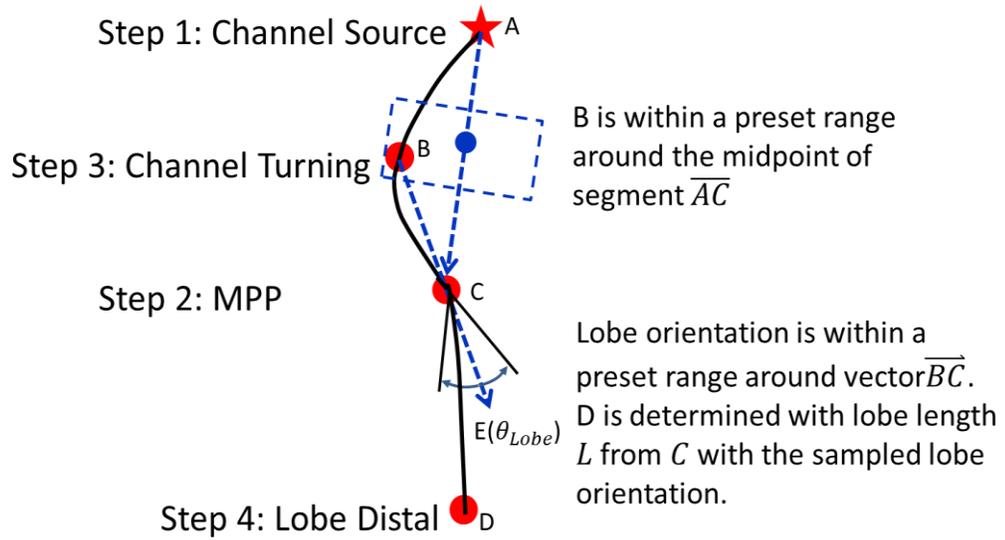


Figure 2.14: Locations of the centerline control points of a geobody are determined following the order 1) channel source 2) MPP 3) channel turning 4) lobe distal point.

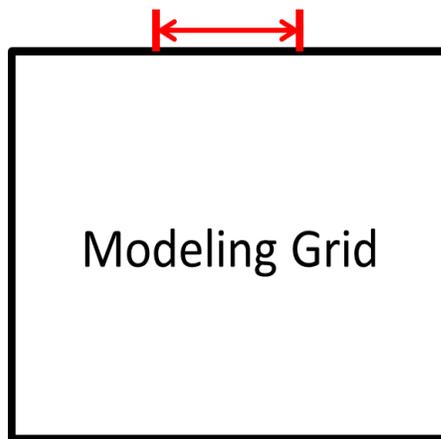


Figure 2.15: The channel source is assumed to be uniformly distributed within a segment around the midpoint of an edge of the modeling grid.

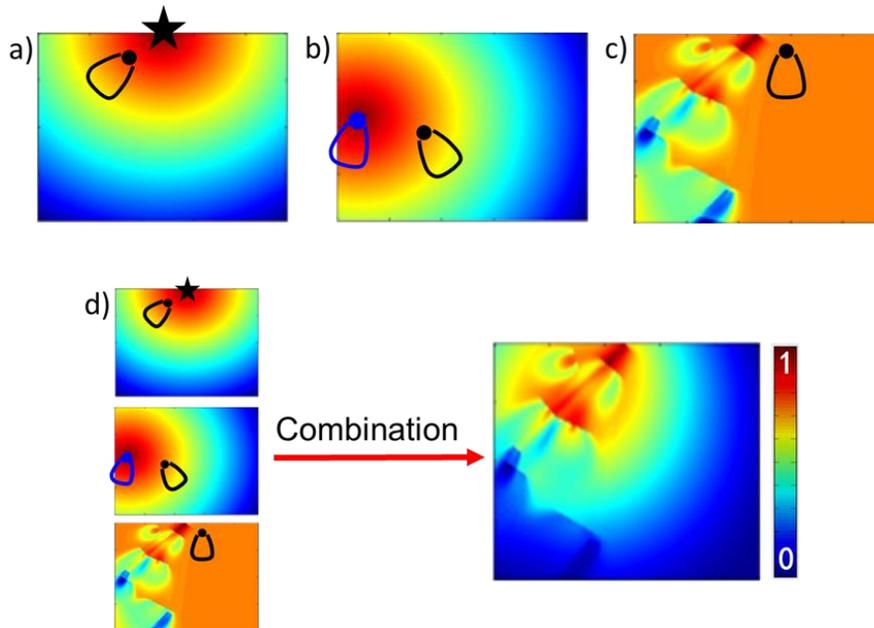


Figure 2.16: The depositional rules for the second control point: lobe Most Proximal Point (MPP). The MPP is drawn from a combined conceptual two-dimensional probability distribution on the modeling grid. The distribution is the weighted mean of three components: a) the next MPP (black dot) tends to be close to the source (star), so for each point in the grid, the probability is proportional to the reciprocal of its distance to a predefined basin source; b) the next MPP (black dot) tends to be close to a previous MPP (blue dot); c) the next MPP (black dot) tends to appear in the region with fewer deposits; d) the final two-dimensional probability map is a weighted mean of the three components.

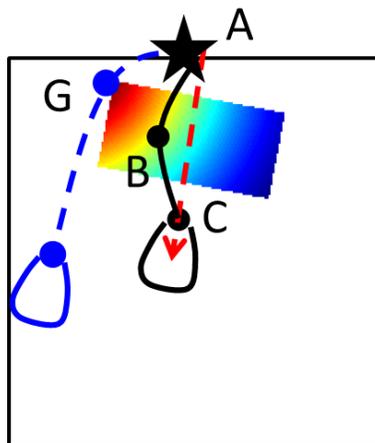


Figure 2.17: The depositional rule for channel turning. Once the channel source A and lobe most proximal point C are determined, a preset window is placed around the midpoint of segment  $\overline{AC}$ . The conceptual probability within the window is determined by the normalized reciprocal distance from the previous channel turning G to every cell within the window. The new channel turning B is sampled within the window.

### **2.2.3 Geomorphic Experiments**

The developed workflow of extracting erosion rules are presented with an experiment of a delta basin, the DB-03 experiment. In this section, the experiment setting is briefly introduced. Then, the method of visualizing and characterizing erosion geometry and patterns from the recorded intermediate topography is presented. Additionally, the correlation between erosion depth and deposition thickness is characterized by a dimensionless ratio, such that the experimental information is applicable to the length scale of real systems, is proposed. The cumulative distribution functions of dimensionless ratios are inputs to the lobate model documented in Section 2.2.2.

#### **Experiment Setting**

The DB-03 experiment was performed and documented by the Sedimentology Group, San Anthony Falls National Laboratory, University of Minnesota. The original purpose of this experiment was to study the creation and preservation of channel-form sand bodies in alluvial systems (Sheets, Paola, and Kelberer 2009). The data from the DB-03 experiment have also been used in studies on the compensational stacking of sedimentary deposits (Straub et al. 2009) and the clustering of sand bodies in fluvial stratigraphy (Hajek, Heller, and Sheets 2010). Only a brief introduction of the experiment is made in this section. For details of the experiment, refer to Sheets, Paola, and Kelberer 2009.

The DB-03 experiment was performed in a tank measuring five meters by five meters and 0.6 meters deep (Fig. 2.18). The sediment infeed point was placed on a corner of the tank. A prerun was performed in which a delta fan was created as the initial topography. The initial topography was a symmetric delta with the radius of 2.5 meters toward the downstream direction from the sediment infeed point. The tank floor was controlled so that the shoreline of the delta was maintained at an approximately constant

location. The injected sediment was a mixture of silica sand (coarse) and anthracite coal (fine).

A subaerial laser scanner recorded intermediate topographies of the delta at intervals of every 2 minutes along three lines perpendicular to the overall infeed flow orientation. The three measurement lines were located at 1.5 meters, 1.75 meters, and 2 meters apart from the infeed point in the downstream direction, respectively (Fig. 2.19). The experiment lasted 30 hours and produced an average of 0.2 meters of stratigraphy. 1180 records of intermediate topographies were available at every line. A fraction of the final stratigraphy is shown in Fig. 2.20. Erosion and deposition are identified by comparing the differences between two successive records (Fig. 2.21). Formal definitions of topography, stratigraphy, erosion, and deposition are presented as follows:

- Topography is defined as the intermediate top surfaces at every  $t$ , represented by a spatial-temporal function  $Z(\vec{x}, t)$ ,  $\vec{x}$  is the spatial coordinate (one dimension along a topographic line); the temporal coordinate is a scalar  $i = 1, 2, 3, \dots, t$ , representing the sequence of records;
- Stratigraphy at  $t$  is defined as all deposited surfaces previous to  $t$ , represented by  $S(\vec{x}, i = 1, 2, \dots, t - 1)$ ;
- Erosion from  $t - \Delta t$  to  $t$  is defined by  $E(\vec{x}, \Delta t)$ , where  $Z(\vec{x}, t) < Z(\vec{x}, t - \Delta t)$ , implying sediment removal caused by the depositional process within the interval  $\Delta t$ ;
- Deposition from  $t - \Delta t$  to  $t$  is defined by  $D(\vec{x}, \Delta t)$ , where  $Z(\vec{x}, t - \Delta t) < Z(\vec{x}, t)$ , implying additional deposits within the interval  $\Delta t$ ;
- If  $Z(\vec{x}, \Delta t) = 0$ , the topography is not modified by a sediment event during  $\Delta t$ ;

The experiment was a subaerial delta but not a deepwater system; however, the treatment is still applicable as long as a similar form of records is measured from a subaqueous experiment.

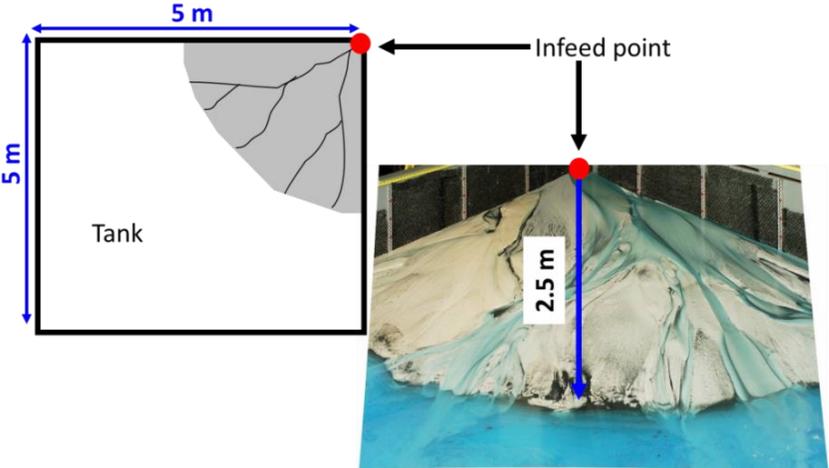


Figure 2.18: The experiment setting for DB-03. The experiment was performed in a tank of five meters by five meters, in which the infeed point was set at a corner of the tank. A prerun was performed to generate the initial topography, which was a symmetric delta with the radius of 2.5 meters apart from the infeed point.

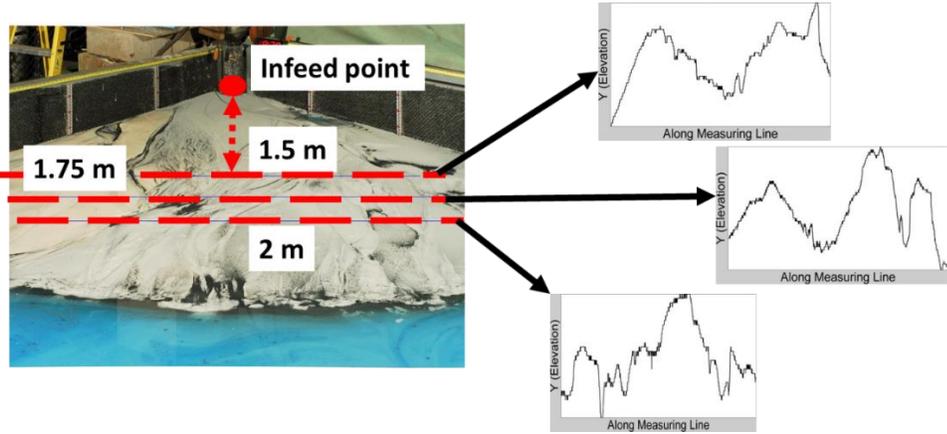


Figure 2.19: The intermediate topography of the experiment was measured along three lines at 1.5 meters, 1.75 meters, and 2 meters apart from the infeed point, respectively.



Figure 2.20: A fraction of the final stratigraphy measured on one topography line, in which the intermediate conditions are partially removed by following events.

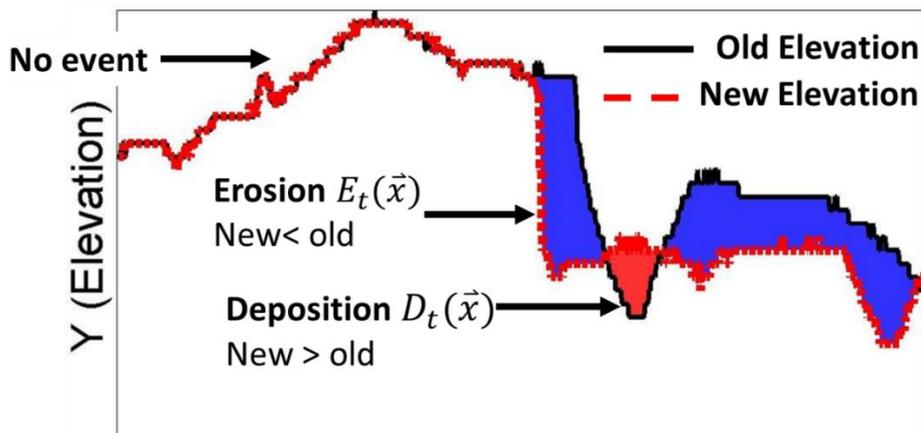


Figure 2.21: The erosion and deposition are identified by computing differences between two successive records. Erosion is defined by reduced elevation between two successive elevation records, while deposition is defined by increased elevation.

## Erosion-Deposition Patterns

For the purpose of applying experimental data to the positive-negative surfaces scheme, illustrated in Fig. 2.11, geometries of erosion and deposition were all measured relative to previous topography. A sediment event plot was devised to visualize the sequential pattern of erosion and deposition geometries. In the sediment event plot, geometries were visualized in the form of net relative elevation changes, and all geometries were plotted along the sequence of occurrence. The plot is demonstrated with an example of two successive records (Fig. 2.22). In the original measurement of elevation along a line, the y-axis represents elevation. An erosion region (blue) is identified at  $T = 1$  (the upper plot in Fig. 2.22 a). Then deposition (red) occurred at  $T = 2$  (the upper plot in Fig. 2.22 b). In the sediment event plot, changes in topography are plotted at the same spatial location (x-axis) as in the original plot, but the geometry of topographic changes for all records are plotted as their order of occurrence (y-axis) (lower plot in Fig. 2.22 a and b). The sediment event plot is a quantified visualization of  $E(\vec{x}, \Delta t)$  and  $D(\vec{x}, \Delta t)$ . An example of the topography evolution plot of 100 records is shown in Fig. 2.23. First, the sediment event plot reveals that erosion usually occurs along with deposition, which is consistent with the scheme of negative-positive surfaces.

Three categories of geometries can be observed in the sediment event plot. In the first category, the width of erosion geometry is greater than the width of deposition, implying the depositional process is erosion-dominated at this stage. The width of erosion is less than the width of deposition in the second category, so the process is deposition-dominated. In the last category, deposition occurs with no erosion. Since three-dimensional geometries of the erosion and deposition are not available, we must assume that different types of geobodies determine different types of geometry pairs. Interpretations for each category are presented in Table 2.1. Since channels are erosion dominated, the first category is interpreted to measure from the channels. Additionally, lobes are deposition-dominated, so the second and third categories are both interpreted

to be lobes. However, some lobes occur with erosion (the second category) and other lobes occur without erosion. The probability of lobes occurring with erosion is 0.29. Measurements of erosion depth or deposition thickness are not directly applicable to real scale systems because length measurements in experiments are different from those in real systems. Based on the observation that erosion and deposition geometry in two successive records are similar, a dimensionless ratio  $r = \frac{h_{Erosion}}{h_{Deposition}}$  characterizes the relationship between the maximum erosion depth and the maximum deposition thickness (Table 2.2). Probability distribution functions of  $r$  for the channel and for the lobe with erosion are input into the surface-based model.

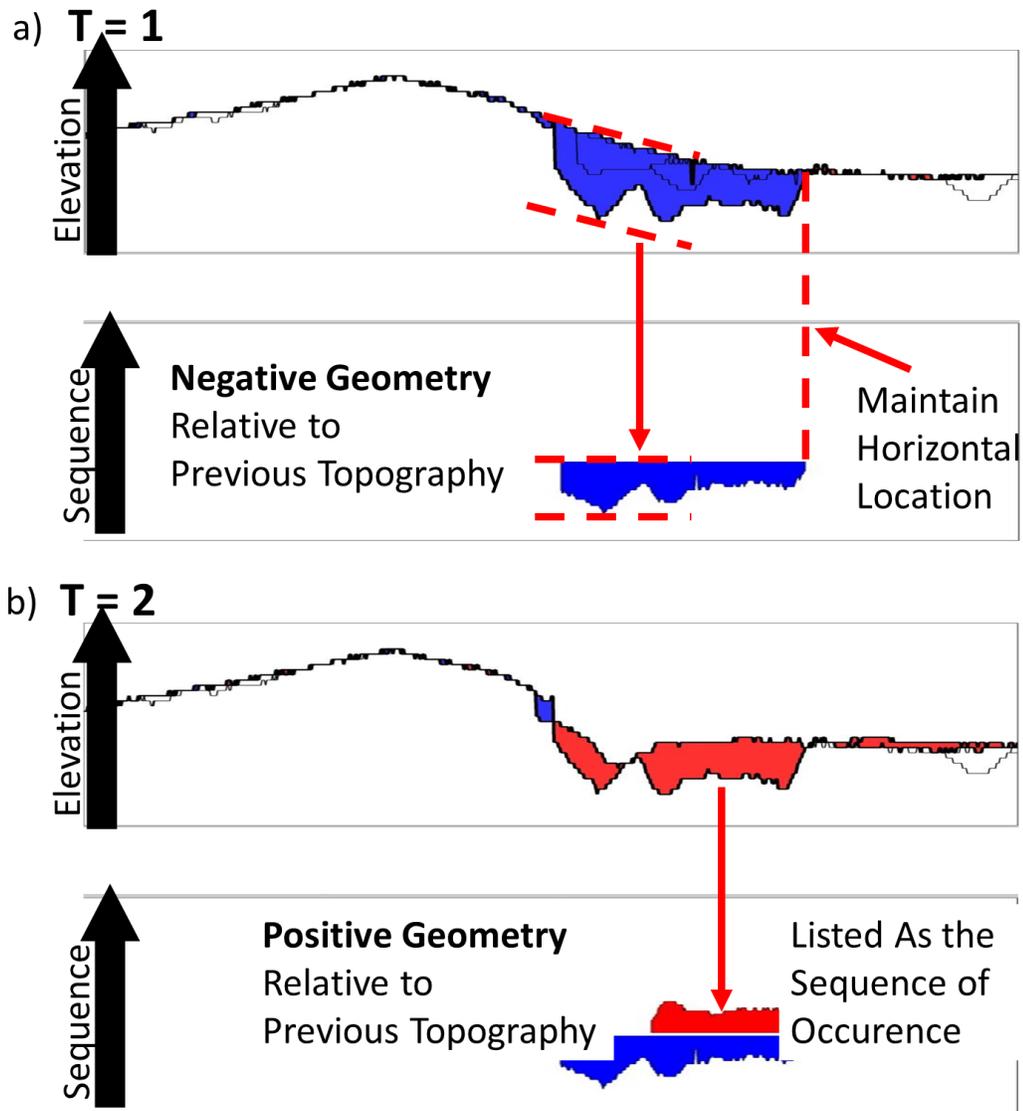


Figure 2.22: Examples of generating a sediment event plot from intermediate topography records: a) to identify the relative elevation change regions between two successive intermediate topography records (the upper plot). The relative change geometry is plotted as its temporal order (the lower plot); b) to identify the series of geometries of the elevation change and plot the all geometries along the vertical axis as the order of occurrence.

## Correlated erosion - deposition pairs

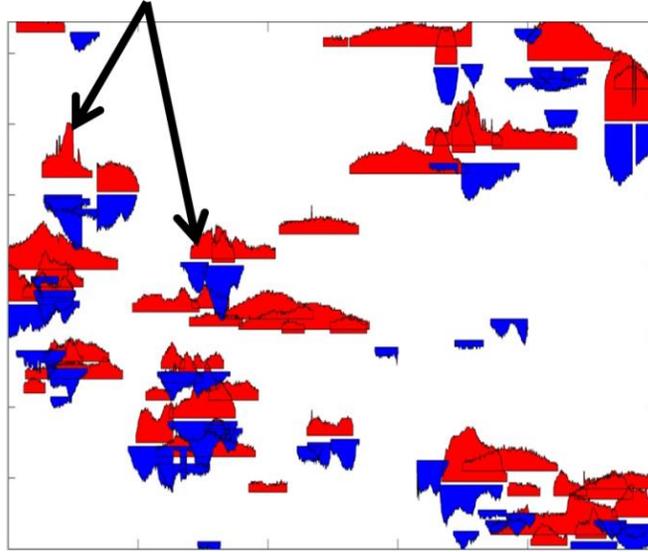


Figure 2.23: The sediment event plot of 100 records on the topography line at 1.75 meters apart from the infeed point.

Channel		$W_{Erosion} \geq W_{Deposition}$
Lobe with erosion		$W_{Erosion} < W_{Deposition}$
Lobe without erosion		$W_{Erosion} = 0$

Table 2.1: Three categories of patterns are interpreted from the sediment event plot 1) channels, where the width of erosion is equal or greater than that of deposition; 2) lobe with erosion, where the width of deposition is greater than that of erosion; 3) lobe without erosion.

	$h_{Deposition}$	
	$h_{Erosion}$	
	$h_{Erosion}$	0.29

Table 2.2: Dimensionless ratios are used to characterize the relationship between erosion-deposition geometries identified from the sediment event plot. Two probability functions respectively for channel and for lobe with erosion are interpreted. The probability of erosion occurs with lobes is 0.29.

### 2.3 Simulation Results

A test simulation of the lobate model is performed to demonstrate the application of erosion rules with input statistics from the DB-03 experiment. As indicated earlier, statistics from the experiment include two distribution functions of the dimensionless ratio  $r$  for the channel and for the lobe with erosion. The model setting and deposition rules were demonstrated in Section 2.2.2. To account for the input statistics, negative-positive surfaces of the channel are generated with  $r_{channel}$  sampled from its probability distribution function. Occurrence of erosion within lobes is controlled by the probability 0.29. If erosion occurs,  $r_{lobe}$  is sampled to generate negative-positive surfaces for the lobe part. If erosion does not occur, a positive surface is generated without a negative surface. Important geometric and simulation parameters are listed in Table 2.3. It is noted that input statistics are extracted from two-dimensional vertical sections, while the realizations are three-dimensional. However, the treatment is still applicable if three-dimensional intermediate topography records are available. The implication of the current treatment is to demonstrate the capability of producing complex stratigraphy

and reproducing the input experimental statistics with surface-based models. The intermediate status of the first four lobes is presented in Fig. 2.24: the first lobe is deposited without erosion, while Lobe 2 – 4 all erode previous topography. Two cross sections perpendicular to the basin flow orientation monitor changes in the stratigraphy made by every event. A model realization of 40 lobes and two cross sections of topography are shown in Fig. 2.25. Cross sections in Fig. 2.24 and 2.25 demonstrate that the surface-based model can produce stratigraphy that is more realistic and more complicated than conventional static reservoir modeling techniques.

A fraction of the sediment event plot from a realization of 100 lobes is compared with that from the experiment (Fig. 2.26). The two plots are different from the perspective of the geometry of negative-positive surfaces, which is reasonable because no three-dimensional geometric information is available to calibrate the trend-to-thickness function in the model. However, the similarity between the two plots is that all three categories of negative-positive surfaces were observed in the simulation as expected. The QQ-plot (Fig. 2.27) compares the simulated and experimental distribution functions of  $r_{channel}$  and  $r_{lobe}$  and demonstrates the reproduction of input statistics of surface-based model. Finally, erosion occurred in 32% of lobes in the realization, which is close to the inputs. Notice that the distribution functions of the  $r_{channel}$  and  $r_{lobe}$  and the probability of lobes with erosion are measured from the three-dimensional realization.

Grid Dimension	300 x 250	
$dx$	1	
Lobe Length ( $L$ )	$50*dx$ to $130*dx$	uniformly distributed
Lobe Width	$0.3*L$ to $0.7*L$	uniformly distributed
Lobe Thickness	$0.002*L$ to $0.008*L$	uniformly distributed
Initial Surface	Flat	

Table 2.3: Primary simulation parameters.

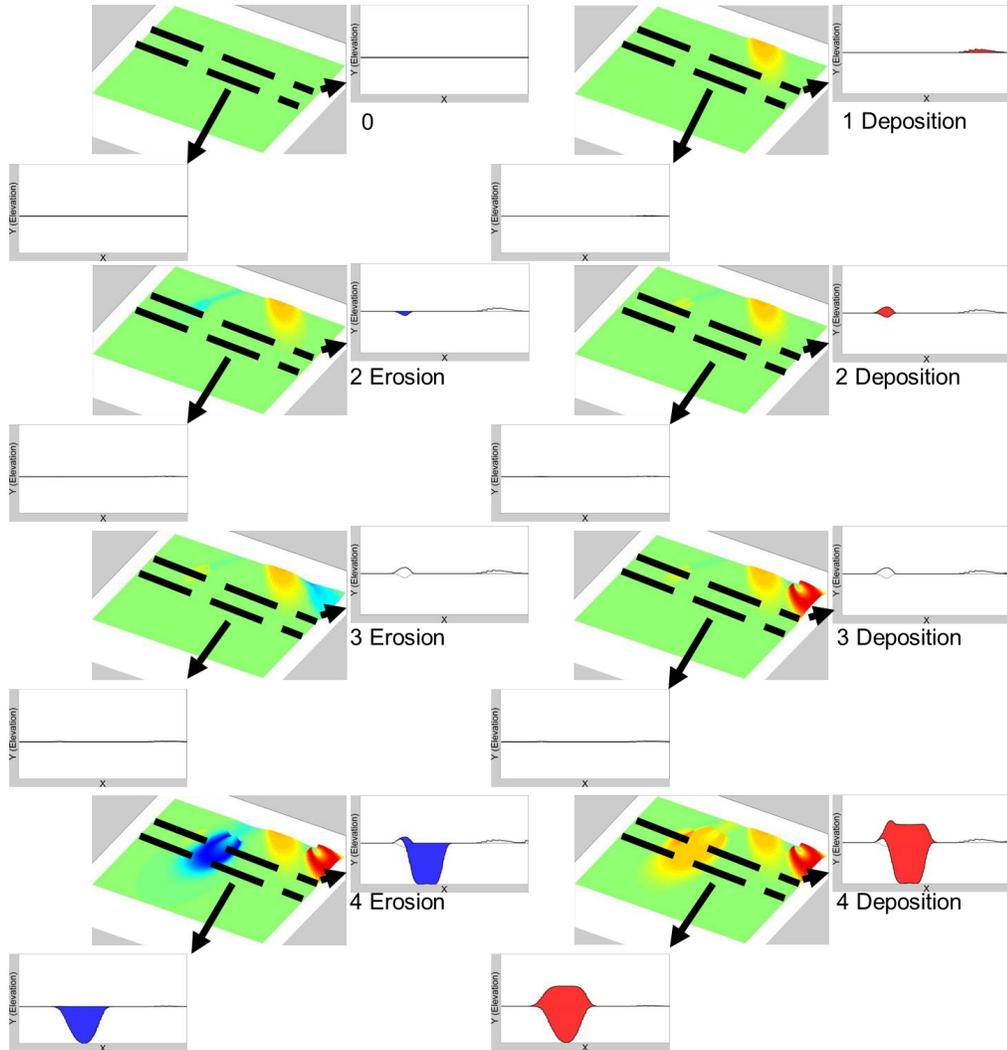


Figure 2.24: The intermediate steps of the simulation for the first four lobes. Two cross sections are used to monitor stratigraphic changes at every step. Erosion does not occur in Lobe 1 while Lobe 2 – Lobe 4 all cause erosion.

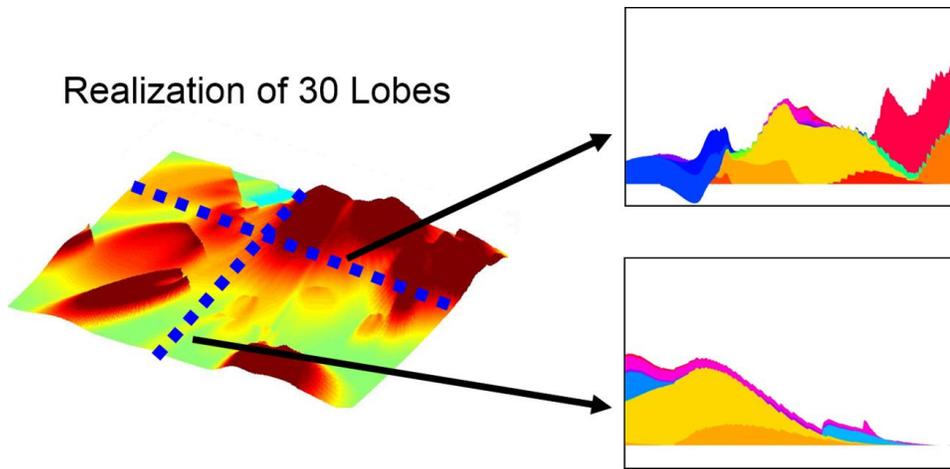


Figure 2.25: A realization of 30 lobes. Two cross sections along and perpendicular to the basin flow orientation are also shown.

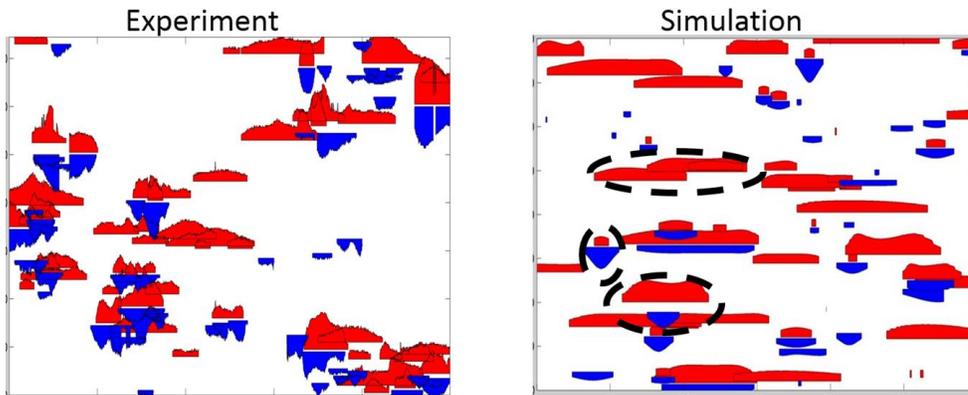


Figure 2.26: A fractions of the experimental and simulated sequential pattern plots. Geometry of erosion and deposition are different, since no three-dimensional data are available to calibrate the trend-to-thickness function. The occurrence of paired negative-positive surfaces for the channel and lobes (with and without erosion) are observed in the simulated result, as expected.

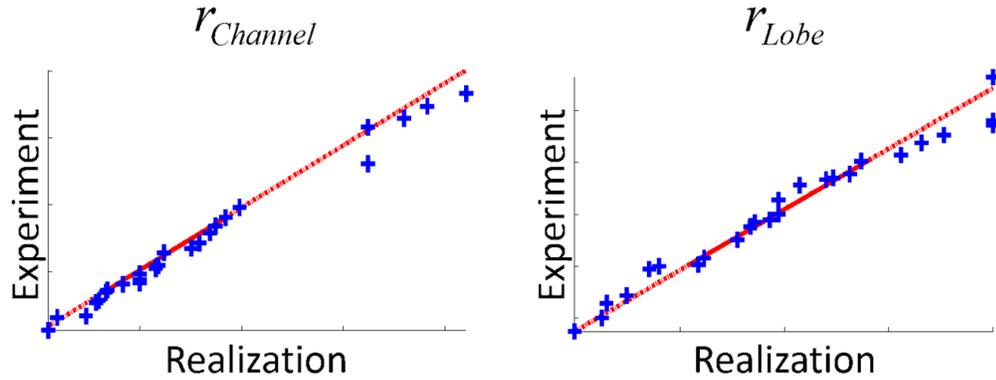


Figure 2.27: The QQ plots prove the reproduction of cumulative distribution functions of the dimensionless ratios in the model realization.

## 2.4 Chapter Summary

A workflow integrating intermediate topographic data measured from geomorphic experiments was developed based on the surface-based modeling techniques. The workflow includes a solution for extracting sequential pattern information on erosion and deposition from intermediate topography records. The geometric information from the experimental data was characterized by dimensionless ratios, such that the information from small experiments is applicable to the real-scale depositional system. A surface-based model of a distributary channel-lobe system was constructed to demonstrate the method. It was demonstrated that using the surface-based model could reproduce input statistics and generate complex and realistic stratigraphy. The experiment used in this chapter only included two-dimensional intermediate topographies; however, the workflow is also applicable to experiments with three-dimensional intermediate topography records. The current surface-based model only accounts for geometric information on deposition and erosion, and further work may involve improvements on depositional rules with experimental data to consider the spatial-temporal clusters of sediment events.

# Chapter 3

## **THE STATISTICAL EXTERNAL SIMILARITY BETWEEN LOBATE ENVIRONMENTS: LINKING EXPERIMENTS TO REAL-SCALE SYSTEMS**

### **3.1 Introduction**

For the integration of experimental data into reservoir models, the first challenge is to build the link between unscaled experiments to real scale depositional systems. In terms of experimental stratigraphy, the similarity between a small version of a system and a large system is defined as the external similarity (Paola et al. 2009). The external similarity is defined distinct from internal similarity, which describes the phenomenon of a small part similar to the whole system. Internal and external similarities are both manifestations of scale dependence. Normally, experimental stratigraphers consider internal similarity as the implication of external similarity. External similarity is the

fundament for linking an experiment to a real depositional system. At present, although we have much to learn about the physics behind the external and internal similarity, they are observed in experiments of various environments. Hence the detection of the external similarity is the objective of this chapter. Unfortunately, neither quantitative solutions to estimate the external similarity, nor methods to specify a geomorphic experiment to a specific real system have been published. Developing a solution linking an experiment to a real depositional system with hydrodynamic laws or conventional stratigraphic method is complicated. From the perspective of hydrodynamics, specifying an experiment to a real system includes setting physical quantities such as flow discharge and sediment grains to be analogous to the unknown paleoflow of the real system, which clearly does not work. Conventional stratigraphy compares depositional systems through the geological history of the systems, which is interpreted based on the spatial-temporal distribution of petrophysical properties (Deptuck et al. 2008; Droz et al. 2003; Gervais et al. 2006; Gervais et al. 2006; Jegou et al. 2008; Pr dat et al. 2010; A. H. Saller et al. 2004; A. Saller et al. 2008). However, geomorphic experiments are extremely simplified systems, in which the sediment composition normally contains only two or three representative types of grains, such as fine and coarse. Thus, stratigraphers rarely consider deposits in experiments to be analogous to a real system.

A statistical method is proposed in this chapter that defines a quantitative and statistical measure on the external similarity between an experimental system and a real-scale depositional system, such that an experiment that is the most similar to the real-scale system can be identified and used as a reference. Then, the external similar experiment is treated as a reference from which the characterized information is applicable in understanding and simulating the real system. The basic assumption of this chapter is that external similarities exist between lobate systems. The solution was devised in a favorable manner for engineering applications, such that the results can be ready to use in surface-based methods for static reservoir modeling. The devised solution consists of two stages of analysis along with a data preprocess step (Section

3.2). Since the intermediate states of the depositional process are recorded by overhead photos, the preprocessing step is required to extract lobate polygons from photos. In the first stage of this method, a treatment to parameterize lobe stacking patterns based on pairwise lobe proximity and an automatic algorithm to quantify experimental lobe hierarchies are presented. The algorithm is required because lobe stacking patterns at various scales in the experiment must be estimated for external similarity, for which manual interpretation is infeasible. In the second stage, lobe stacking patterns are first obtained from the quantified lobe hierarchy. Then, the lobe stacking pattern is characterized as a spatial point process. Two point processes from respective lobe stacking patterns are compared based on the nearest neighbor statistics. A two-sample bootstrap hypothesis test based on L1-norm differences between cumulative distribution functions was designed to estimate the similarity between nearest neighbor statistics with a probability value. Each stage is demonstrated by examples followed by an application with the whole workflow for testing the external similarity between two experimental and two published real-scale lobe stacking patterns (Section 3.3). Finally, results of the method are discussed on a technical level.

## **3.2 Methodology**

The method aims to quantify the similarity between two lobe stacking patterns. If stacking patterns from an experiment are similar to that from a real-scale system, the experiment can be used as a database of sedimentological knowledge for the real system. Therefore, depositional rules can be characterized and calibrated with comprehensive experiment data not available in the real system, and the determined rules are then applicable to model a real-scale system using surface-based methods. Since the intermediate conditions of the experiment were recorded by overhead photos, the preprocessing step for extracting lobes from photos is introduced in Section 3.2.1. Stage One of the method (Section 3.2.2) included parameterization and hierarchy quantification. A lobe stacking pattern was parameterized by the pairwise proximity of every pair of lobes within the stacking pattern. A hierarchical clustering algorithm was

designed to automatically quantify the hierarchy of lobes in experiments based on the parameterization. The lobe hierarchy is a quantitative description of the function between scales of interpretation and lobe stacking patterns, such that lobe stacking patterns could be extracted from the quantified lobe hierarchy at different scales. Once a stacking pattern was obtained from the hierarchy, it was compared to the stacking pattern interpreted from the real-scale depositional environment. Section 3.2.3 describes a special hypothesis test based on bootstrap and L1-norm measures that was designed to estimate a p-value for measuring similarities between the two stacking patterns.

### **3.2.1 Experimental Data and Preprocessing**

Other than DB-03 in Chapter 2, two alternative delta fan experiments are described in this chapter. The settings of the new experiments were similar to DB-03. Both experiments were performed in the same tank as the DB-03, and the sediment infeed points were set on a corner of the tank. The frequency of intermediate topographic records in these experiments were not sufficient to estimate the lobe at fine scales, so only overhead photos obtained at the interval of 30 seconds became available to measure the intermediate status of the lobes in the process. The overhead photos were recorded by a camera over the tank. Since photos were geometrically corrected to eliminate distortions caused by the lens and location of the camera (Fig. 3.1), the shape and length measurements in the horizontal plane of photos are accurate.

The preprocessing on overhead photos involved three tasks:

- Identification of a successive fraction in the overhead photo sequence, such that behaviors of a single active channel and its connected lobe were recorded. Experiments used in this chapter were performed for more than 100 hours, and complex features, such as multiple channels and lobes, were observed. Since the surface-based model in this work only considered one lobe at each step, a very small fraction of the photo sequence was selected, such that only one active channel with its lobe appears in each photo.

- Interpretation and clipping of the lobe region from a photo. Manual interpretation was required to distinguish the lobate geometry from the complex background of an overhead photo. Since photos were recorded at very fine intervals, a new lobe was identified only when a significant shift was observed on the end of the channel between two successive photos. Otherwise, only one lobe is identified from multiple photos (Fig. 3.2).
- Extraction of geometric features for lobes, including the lobe MPP, lobe orientation, and the lobe boundary, from clipped photos. An automatic script based on the image processing toolbox in Matlab was developed for this purpose. The geometric features of a lobe are defined as follows.
  - A lobe boundary is defined by the outmost cells of the cropped lobe in an image.
  - The MPP of a lobe is defined as a lobe boundary point that is the closest to the sediment infeed point.
  - Lobe orientation is defined by the unit vector from the lobe MPP to the geometric centroid.

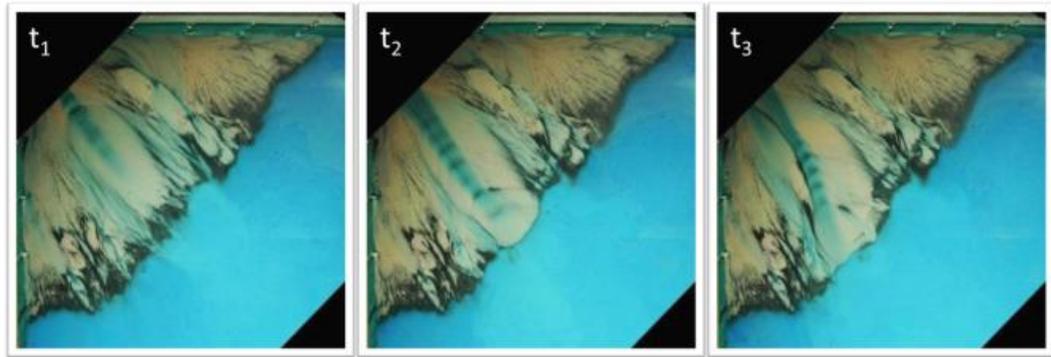


Figure 3.1: Corrected overhead photos of an experiment taken at regular intervals, in which only one active channel was recorded within the successive photos. Courtesy by Prof. Chris Paola, San Anthony Falls National Laboratory.

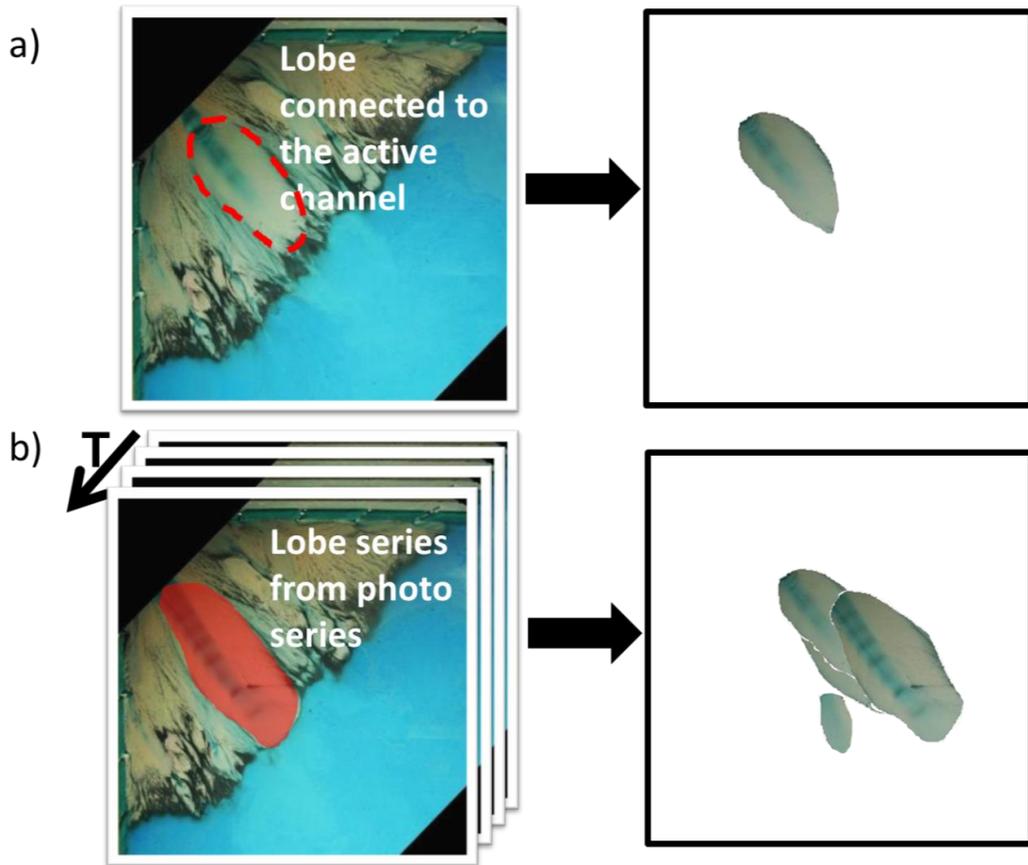


Figure 3.2: The lobe connected to the latest channel was interpreted on every photo, and a series of lobes are obtained.

### 3.2.2 Stage One – Automatic Lobe Hierarchy Quantification

As the prerequisite for a quantitative study, a parameterization strategy for lobe stacking patterns was proposed based on the pairwise proximity between the lobes. Since lobes are polygons with geological interpretations, the pairwise lobe proximity is measured by a geological distance function designed to account for both geometric and geological proximity. Based on the lobe proximity, the hierarchy of parameterized lobes can be quantified with agglomerative hierarchical clustering. A temporal constraint was designed for the clustering algorithm, such that the clustering algorithm followed the sequence of lobes. The method is demonstrated with a sample sequence of lobes.

#### Parameterization of Lobe Stacking Patterns

Stacking patterns are differentiated by the intensity of regional lobe clustering. In terms of spatial statistical analysis, intensity of regional lobe clustering is reflected by the statistics of pairwise proximity between lobate polygons. The pairwise geological proximity of lobes were measured by four parameters, namely, the lobe MPP proximity, the lobe orientation proximity, the polygonal proximity, and the shape proximity.

- **MPP proximity  $d_{MPP}$ :** The MPP of a lobe is defined as the boundary point of a lobe that is the closest to the infeed point of an experiment tank (Fig. 3.3 a). In the conceptual model of flow process in lobate system, the MPP is interpreted as the starting point where the energy of the sediment flow starts weakening. Thus, the spatial proximity between MPPs implies genetic differences between two lobes.
- **Orientation proximity  $d_{\theta}$ :** Lobe orientation is defined by the unit vector from the lobe MPP to the geometric centroid of the polygon (Fig. 3.3 b). The lobe orientation records the orientation at the end of the life of a sediment flow. Similar to MPP, the lobe orientation proximity also reflects the genetic dissimilarity between two lobes.

- **Polygonal proximity  $d_p$ :** Although lobes are usually marked by MPPs, they are geometrically polygons. Hence the polygonal proximity must also be studied to measure the areal density. In this work, the polygonal proximity is measured by the Hausdorff distance (Velkamp 2001) (Fig. 3.3 c) because of its simplicity. The Hausdorff distance is the maximum distance from a point set to the nearest point in another set, defined by Eq (3.1) – (3.2). As indicated, two lobes are represented by the boundary points while calculating the polygonal proximity.

$$h(A, B) = \max_{a \in A} \{ \min_{b \in B} [d(a, b)] \} \quad (3.1)$$

$$H(A, B) = \max \{ h(A, B), h(B, A) \} \quad (3.2)$$

- **Shape proximity  $d_s$ :** Small lobate geobodies within a larger one normally share similar shapes. The shape proximity is measured by applying procrustes analysis on the boundaries of two lobes (Fig. 3.3 d). In the procrustes analysis, two lobes are first translated and rotated such that the MPP and lobe orientation are overlapped. Then, coordinates of a lobe boundary are normalized by the Root Mean Square Distance (RMSD) (Eq. (3.3) – (3.4)), which is a statistical measure of the scale of the lobes. Finally, differences between lobes are calculated along two normalized boundaries.

$$s = \sqrt{\frac{\sum_{i=1}^k (x_i - x_{MPP})^2 + (y_i - y_{MPP})^2}{k}} \quad (3.3)$$

$$\bar{x}_l = \frac{x_i - x_{MPP}}{s}, \bar{y}_l = \frac{y_i - y_{MPP}}{s} \quad (3.4)$$

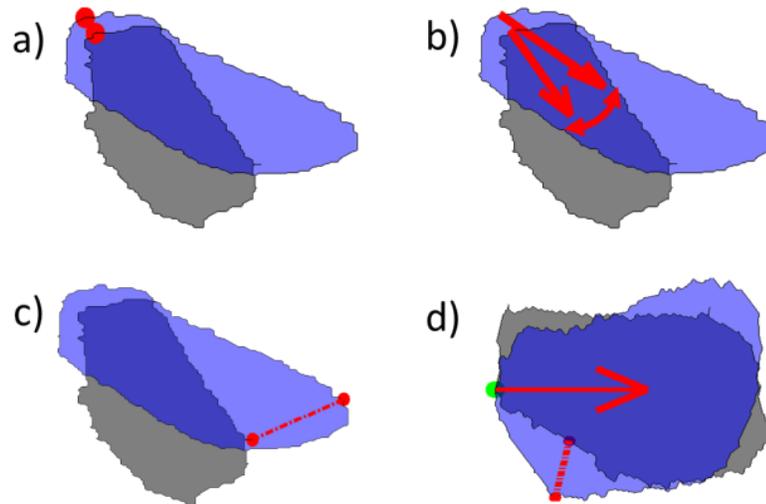


Figure 3.3: The parameterization of lobe stacking patterns: a) the source is defined as the MPP on the lobe boundary to the basin source; b) lobe orientation is defined by the unit vector pointing from the source point to the geometric centroid of a lobe; c) polygonal proximity is measured by the Hausdorff distance between two lobes; d) shape proximity is measured through procrustes analysis.

### Hierarchy of Lobes

Studies on real lobate depositional systems reveal a hierarchical organization of lobate geobodies (Préat et al. 2010). The hierarchy includes event beds, lobes, lobe complexes, and fans, varying from thinnest and smallest to thickest and largest. In real-scale depositional systems, determination of such a hierarchy is based on sedimentologic and stratigraphic studies of the spatial-temporal distribution of petrophysical properties (e. g. grain sizes, stratigraphic formations, paleoflow conditions, and deposit thickness). Within a hierarchy of lobate geobodies, the level of lobes and the lobe elements are of the most interest for petroleum exploration and are the scale to which static reservoir modeling is applied. Hierarchies of lobate geobodies exist in geomorphic experiments as well. However, in experiments, the stratigraphic method is not able to identify a level within the hierarchy that is equivalent to lobes within a real-scale system, since the experimental processes are extremely simplified and upscaled concerning flow conditions, grain sizes, etc. Moreover, data recorded in

geomorphic experiments are extremely substantial (e.g., tens of thousands of overhead photos recorded every 30 seconds) compared to data obtained from real systems through geological field studies and seismic surveys (e.g., several or less than 20 lobes). Therefore, conventional stratigraphic methods based on expert analysis is essentially infeasible.

Based on the assumption that a statistical and external similarity exists between small depositional systems (experimental) and large depositional systems (real scale) (Paola et al. 2009) and the capability of characterizing a lobate system by its hierarchy of lobes, the algorithm presented in this section aims for automatic and quantitative characterization of lobe hierarchies in experiments based on the proximity measure between lobes. A hierarchy of lobes is a description of the function between scales of interpretation and lobe stacking patterns. The scale of interpretation is the scale on which geobodies of petrofacies are interpreted. If the geological dissimilarity between a set of lobes are below a chosen scale of interpretation, they are interpreted as an integrated lobe at the chosen scale. The proximity measure should integrate dissimilarities between two lobes in a spatial, temporal, and geological sense based on the four parameters indicated earlier. For example, MPPs of small lobes forming a higher scale lobe are located within the intermediate vicinity. Moreover, these small lobes within the same higher scale lobe must be successive in time, since temporal gaps lead to the interpretation of multiple lobes. Finally, small lobes of a higher scale lobe usually share similar shapes and orientations. Once the pairwise proximity of lobes are quantified by the four parameters, the overall hierarchy of a dataset is quantified by the hierarchical clustering algorithm. The challenge is to constrain the general statistical algorithm, such that the results are geologically reasonable.

## Agglomerative Hierarchical Clustering with Geological Adjustments

The most extensively applied data-mining algorithm for summarizing data hierarchies is agglomerative hierarchical clustering. Given a dataset  $G$  of  $N$ , the basic process of agglomerative hierarchical clustering is as follows (Johnson 1967).

Step 1: Compute the similarity between every pair of observations in  $G$ , and assign each single observation into its own cluster, defined as  $N$  clusters.

Step 2: Identify the most similar pair of clusters and merge them into a single cluster. Now the number of clusters is  $N - 1$ .

Step 3: Compute the pairwise similarity between the new cluster and the old clusters.

Step 4: Repeat Step 2 – Step 3 until  $N = 1$

The similarity is measured using distance, and Euclidean distance was chosen for this work. Different strategies can be performed in Step 3 to compare the similarity between clusters, such as single-linkage, complete-linkage, etc. In this chapter, the single-linkage, the distance between the centroids of two clusters, is used. Hierarchical clustering has been extensively applied as a powerful tool for data analysis. However, adjustments reflecting stratigraphic interpretation rules must be applied on the normal hierarchical clustering procedure to generate geologically reasonable groups of lobes.

Since the geological proximity between two lobes is measured through the four parameters defined earlier, the first adjustment is the definition of a distance function integrating all four parameters. Given a set of lobes  $G$ , a single lobe  $g$  is measured by a four-dimensional vector  $g = (MPP, \theta, Polygon, Shape)$ . As an alternative to the normal treatment of clustering in the four-dimensional space, the proximity between two lobes is defined by the weighted mean of the proximity in the four respective

components (Eq. (3.5)) for the purpose of accounting for expert judgment on the importance of each parameter in the interpretation.

$$d_f = w_{MPP}d_{MPP} + w_{\theta}d_{\theta} + w_p d_p + w_s d_s, \quad \sum w_k = 1, k = MPP, \theta, p, s \quad (3.5)$$

where empirical weights reflect the importance of each component. In the determination of grouping two small lobes into a higher scale lobe during the interpretation, the component with the higher weight is considered more important. For example, one set of weight values used in this work that give a reasonable clustering of lobes as  $w_{MPP} = 0.4$ ;  $w_{\theta} = 0.2$ ;  $w_p = 0.3$ ;  $w_s = 0.1$ , implying that the interpreter considers the dissimilarity between MPPs as the most important criteria in the interpretation. In this manner, proximities of four parameters are fused into a single value. In the following sections, the fused proximity is named as the geological distance, and the one-dimensional space of geological distance is defined as the interpretation space. Any scale comparison in the following sections is in the interpretation space.

The second adjustment is the temporal constraint, which is a typical stratigraphic consideration that has not been studied in the algorithmic literature on clustering. During the interpretation procedure for a lobate depositional system, two lobes may be close from the perspective of the four parameters but are separated by a relatively long gap from the perspective of time. In the lobate reservoir, especially in deepwater systems, the time gap between two spatially overlapped lobes is usually indicated by thin impermeable shale layers. These thin layers are difficult to measure in experiments. Hence, the purpose of a temporal constraint is to distinguish two clusters of lobes that are spatially and geometrically close but temporally separated. Distinct from assigning empirical weights to four proximity components, which is a ‘soft’ adjustment on the hierarchical clustering, the temporal adjacency constraint is a ‘hard’ constraint, forcing the agglomeration of clusters to occur only in two geobodies that are successive in time. For example, given a lobe sequence  $g_i, i = t_1, t_2, t_3$ , the acceptable agglomerations are  $\{t_1, t_2\}$  or  $\{t_2, t_3\}$ , while  $\{t_1, t_3\}$  is not allowed (Fig. 3.4 a). Instead of directly modifying the standard hierarchical clustering algorithm, which is efficient and

available in most statistical toolboxes, a sequential cumulated distance (Eq. (3.6)) is applied to ensure that the distance between two points closer in time is always closer in the interpretation space.

$$d_{(i,j)} = \sum_i^{j-1} d_{f(k,k+1)} \quad (3.6)$$

where  $i, j$  are temporal indices of a pair of points in the sequence;  $d_{f(m,n)}$  is defined in Eq. (3.5). The essence of the sequential, cumulated distance between the  $i$ th point and  $j$ th point in time is the sum of segments linking each successive pair of points between  $i$  and  $j$ , which ensures that  $d_{(i,j)}$  is always greater than  $d_{(i,k)}, i < k < j$ , such that agglomeration only occurs in temporally successive pairs of clusters (Fig. 3.4 b).

The results of hierarchical clustering can be visualized by a dendrogram (Fig. 3.5), which is a weighted binary tree structure. Each leaf of the tree represents an observation (a lobe) from the clustered dataset, and each branch of the tree structure represents a cluster, or a higher-scale lobe. The height of the junction of two branches represents the distance between two clusters. Thus, a dendrogram describes the hierarchy of lobes as a function between the scale of interpretation and the configuration of clusters.

The lobe hierarchy described by a dendrogram is straightforward; however, the treatment works well only with small datasets. In cases with hundreds of observations and multiple levels in the hierarchy, visualizing lower levels of the tree structure becomes infeasible (Fig. 3.6). Thus, the reachability plot (Fig. 3.7), an alternative form of dendrogram, is introduced for large datasets.

The reachability plot (Ankerst et al. 1999; Breunig et al. 1999; Sander et al. 2003) is a bar chart representing clustering results of density-based clustering algorithms. Studies have proved that a reachability plot is equivalent to a dendrogram in certain cases and can be directly converted from dendrograms resulting from hierarchical clustering. Each observation in the clustered dataset is represented by a bar in the reachability plot, and bars are aligned along the horizontal axis, maintaining the original

order of leaves of the dendrogram. The height of a bar is the distance from an observation to all observations on its left, which is the height of the common junction between a leaf and all leaves on its left in the dendrogram. Tall bars in a reachability plot represent large intercluster gaps and short bars represent short gaps. Thus, valleys in a reachability plot represent a dense set of lobes. The advantage of a reachability plot is that the method is able to properly visualize large datasets with hundreds of observations.

The interesting property of a reachability plot, or a dendrogram, is that the y-axis represents scales of interpretation, and valleys in bars represent a configuration of clustering. In other words, the dendrogram defines a function between scales of interpretation and the stacking pattern of lobes. A scale of interpretation is determined by thresholding the y-axis. Bars greater than the chosen scale intersect with the threshold, and a valley between two intersections is a cluster. Each cluster is then a lobe at the chosen scale of interpretation (Fig. 3.8). If the y-axis of a reachability plot is thresholded from the finest to the coarsest spatial and temporal scales recorded in an experiment, the respective stacking pattern of lobes at the scale quantified by the threshold is obtained. Therefore, a scale-dependent study on the lobe hierarchy becomes feasible (as in Section 3.3). The significance of this method is that the interpretation process in the experiment can be quantified with an automatic algorithm for the first time.

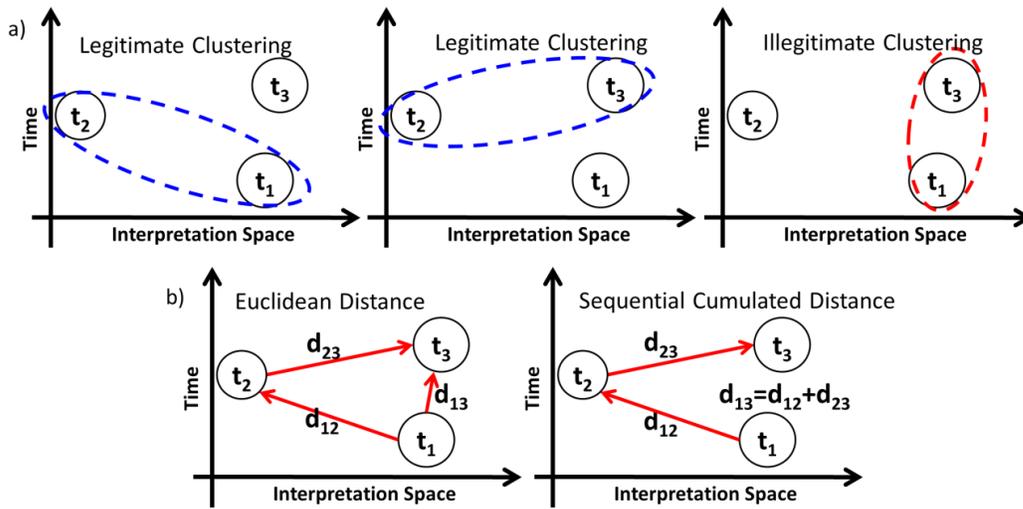


Figure 3.4: A synthetic demonstration of the temporal constraint: a) based on rules in the interpretation, only agglomerations on temporarily successive lobes are legitimate in clustering; b) the sequential cumulated distance from Lobe  $t_1$  to Lobe  $t_3$  is the distance following the sequential route from  $t_1$  to  $t_2$  to  $t_3$ , promising that  $d_{13}$  is always greater than  $d_{12}$  and  $d_{23}$ .

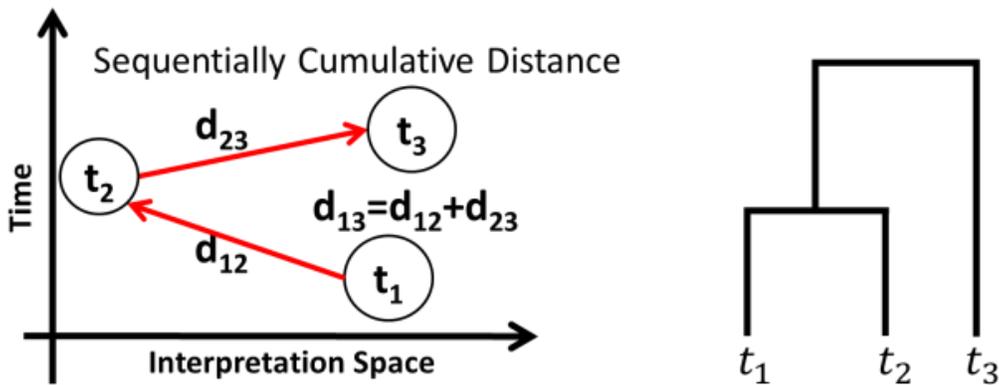


Figure 3.5: The hierarchical clustering result is visualized by a dendrogram. In the synthetic demonstration, distance  $d_{12} < d_{23} < d_{13}$ , the dendrogram visualizes a lower scale group  $(t_1, t_2)$  and a higher scale group  $((t_1, t_2), t_3)$ .

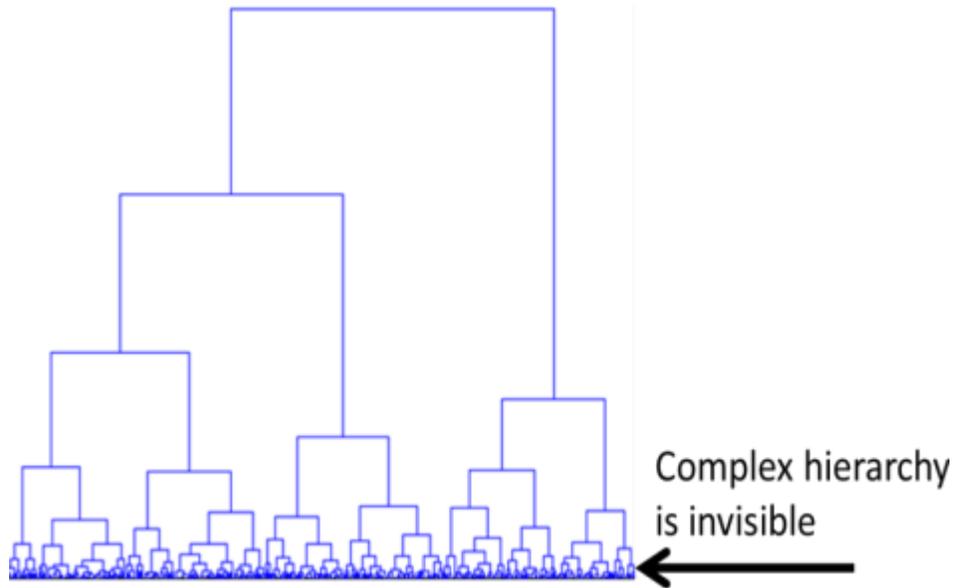


Figure 3.6: The dendrogram of a complex dataset. Visualization of large datasets with dendrograms is inconvenient.

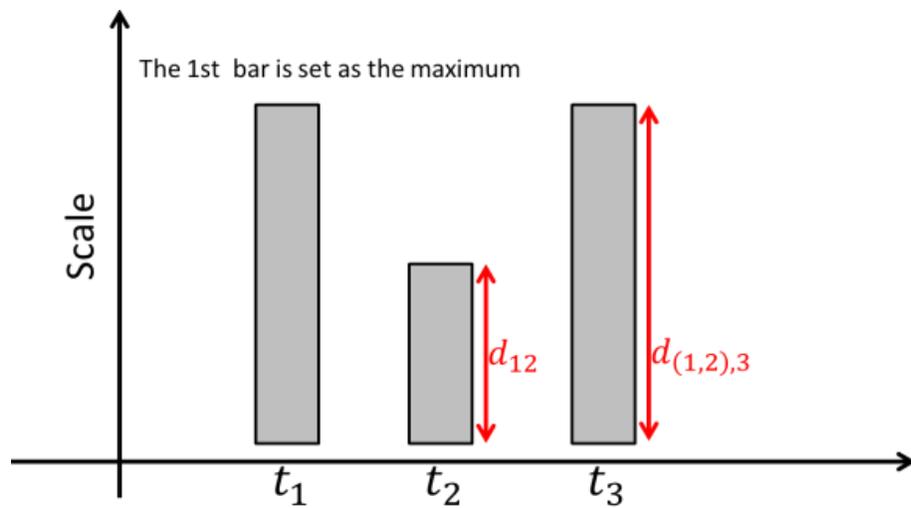


Figure 3.7: The reachability plot as a dendrogram alternative to visualize a large dataset. The height of the bars represent the distance from one observation to all the left neighbors in the plot. The first bar is set to be the maximum height within this clustering.

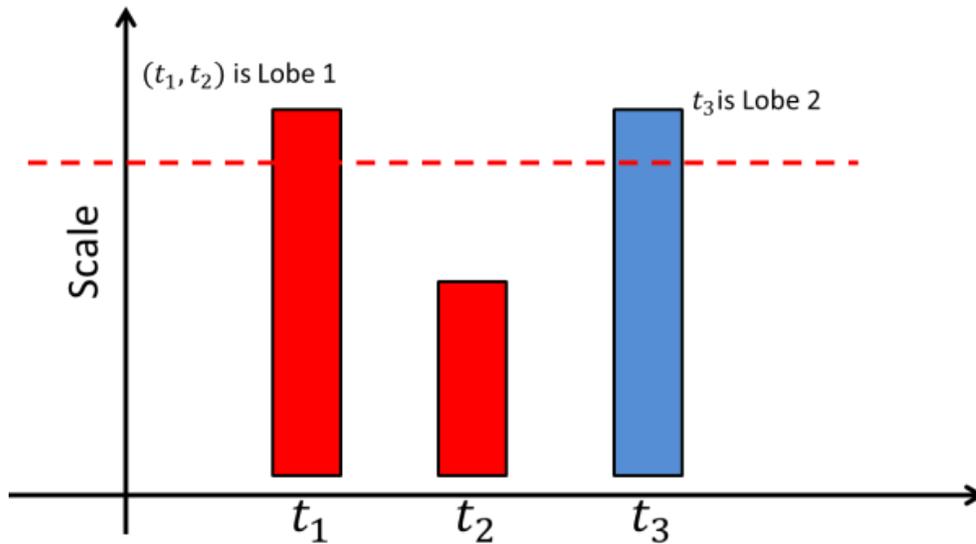


Figure 3.8: The threshold intersects with bars greater than the chosen scale of interpretation within a reachability plot. Lobes between two intersections form a lobe at the chosen scale.

### Example

An exemplary sample of 36 successive lobes (Fig. 3.9) is selected from a geomorphic experiment for the demonstration of the algorithm in Stage One. The dendrogram (Fig. 3.10 a) and reachability plot (Fig. 3.10 b) of the sample are generated with the weighted mean distance. Within the dendrogram, the 36 lobes are organized by the order of clustering in the hierarchical clustering and listed along the horizontal axis, which is the same as the order in the reachability plot. The interesting observation is that the dendrogram and reachability plot reveal a relatively dense group to the left of Lobe 5, indicated by a set of lower branch junctions in the dendrogram and a valley in the reachability plot. The lobe density gradually decreases from Lobe 5 to Lobe 21, indicating several groups with a fewer number of lobes. Finally, Lobes 1, 2, 3, 4 are separated from other lobes with a much longer distance in the interpretation space. With a selected threshold, the sample lobe set is clustered into six groups (Fig. 3.11). The dense group from Lobe 35 to Lobe 22 are clustered into Group 1, the gradually sparse lobes are clustered into Group 2 to 4 with a fewer number of lobes, and the sparsest

lobes are clustered into Group 5, with a single lobe, and Group 6, with three lobes. Higher scale lobes formed by each group are also demonstrated in Fig. 3.11. Lobes in the same group share common MPP and orientation, as well as polygonal overlap and shape, proving the robustness of the algorithm.

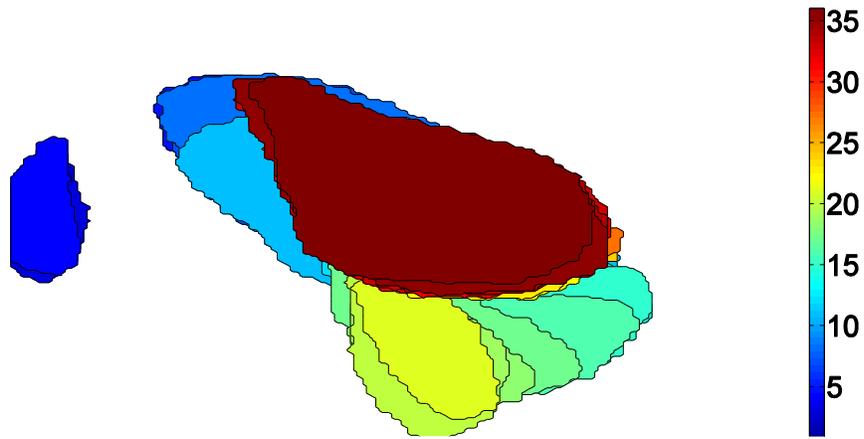


Figure 3.9: A sample lobe sequence of 36 lobes. Color represents lobe location within the sequence, ranging from 1 to 36.

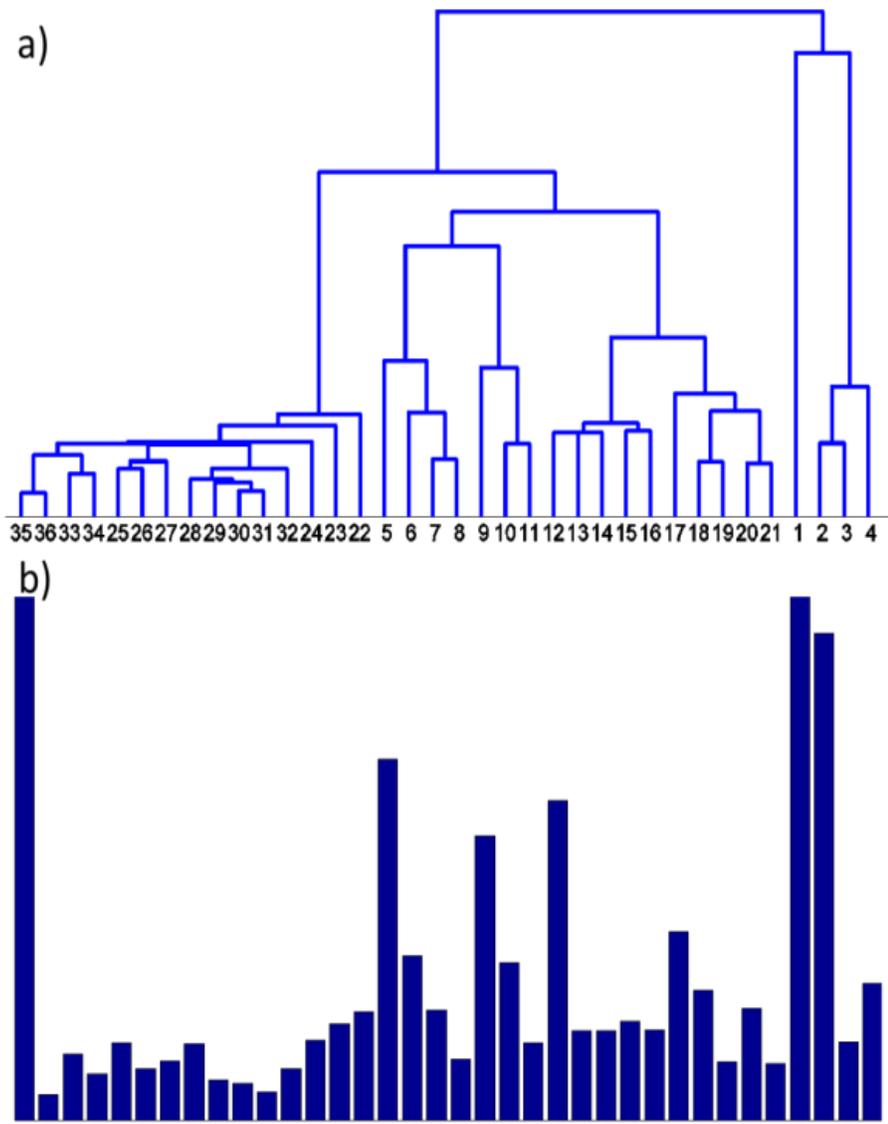


Figure 3.10: a) The dendrogram of the sample lobes, the 36 sample lobes are organized by the order of being clustered in the clustering algorithm, and listed along the horizontal axis; b) the reachability plot of the sample lobe; lobes are listed with the same order as in the dendrogram.

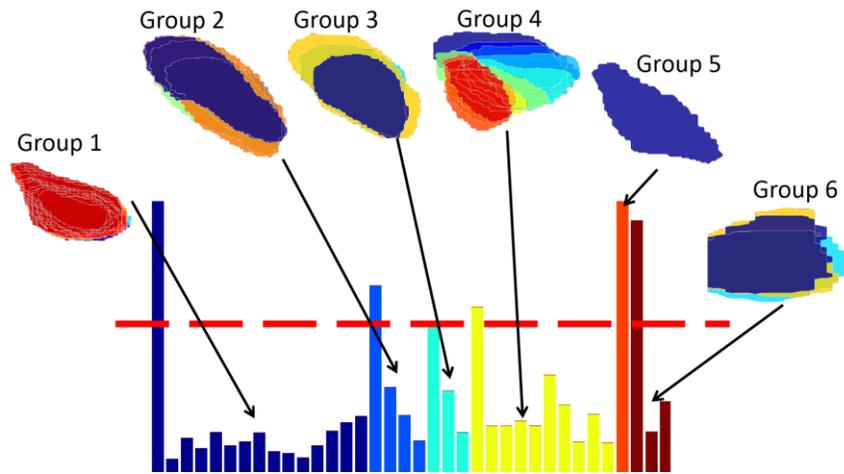


Figure 3.11: Clusters on the exemplary sample of lobes with a selected threshold. Lobes in each group form a lobe at the scale of interpretation.

### 3.2.3 Stage Two – Testing Pattern External Similarity

In this section, a statistical manner of comparing two lobate stacking patterns is introduced. Lobes in a stacking pattern are characterized by a point pattern of MPPs, which is further characterized by its probability distribution function of the nearest neighbor's proximity for the purpose of reflecting the intensity of the local clustering of lobes. Probability distribution functions of the nearest neighbor's proximity are compared with a hypothesis test. Since normal hypothesis testing techniques are not robust in cases of small numbers of lobes, a bootstrap two-sample hypothesis test based on L1-norm measures was devised. The output of the hypothesis test was a p-value representing the similarity of the patterns. A higher p-value implied the experiment should be chosen. This method is demonstrated by an example as well.

### Characterization of Stacking Patterns

An interesting feature of the parameterization for stacking patterns is that each parameter has its own distinctive ability to distinguish lobe stacking patterns from various depositional systems. The most important parameter determining a lobe location

is the MPP, because this parameter implies the genetic origin of a lobe. In this sense, a lobe stacking pattern is converted to a spatial point process of MPPs. The other parameters, including lobe orientation  $\theta$ , polygonal distance  $p$ , and lobe shape  $s$  are treated as properties of the MPPs. If a lobe  $g$  in Stage One is represented as  $g(MPP, \theta, p, s)$ , in which the four parameters are equivalent features, lobe  $g$  is represented as  $MPP_g(\theta, p, s)$  in Stage Two, implying orientation, polygonal distance, and shape are attributes belonging to MMP. Fig. 3.12 a demonstrates a sample set of 60 lobes marked by the MPPs, and Fig. 3.12 b is the spatial point process.

The feature of a lobate stacking pattern that is important to reservoir performance is the intensity of the local clustering of lobes. The intensity of clustering is detected by studying the nearest neighboring lobe dissimilarity. Several techniques are available to characterize the intensity of clustering in a spatial point pattern, such as the empty space distance function  $F$ , the nearest neighbor distance function  $G$ , and pairwise distance function  $K$  (Diggle 2003). The empty space distance is calculated from empty locations within the domain of the point process to their nearest point. The  $F$  function is the empirical cumulative distribution function (ECDF) of empty space distances. The disadvantage of the  $F$  function is that the treatment is sensitive to the domain boundary of the point process. However, the domain boundary for a lobe stacking pattern is the depositional basin boundary, which is an interpretation. A slight difference in the interpretation of the basin boundary will affect the resulting  $F$  function. The  $K$  function calculates pairwise distances of points and plots the average number of point versus the distance of measurement. However, the  $K$  function is not a cumulative distribution function, which is inconvenient as an input for the hypothesis test that will be presented later in this section. The  $G$  function is the ECDF of the nearest neighbor's distance in a point process, which does not depend on boundary measurements and is straightforward as an input for a hypothesis test. In this work, the concept of the  $G$  function is applied to characterize a point process converted from a lobe stacking pattern. For a point process of a lobe stacking pattern, four  $G$  functions,  $G_{MPP}, G_{\theta}, G_p, G_s$ , are calculated

corresponding to the four parameters denoted earlier. Since other parameters are treated as attributes of the MPP, the functions are calculated as follows: First, a nearest neighbor for each lobe is identified in the distance of the MPP, so that  $G_{MPP}$  is obtained as the empirical distribution function of the nearest neighbor's MPP distance. Between each lobe and its nearest MPP neighbor, proximities for the remaining three parameters are calculated to obtain the ECDFs,  $G_\theta$ ,  $G_p$ ,  $G_s$ . In this manner, a stacking pattern is characterized by four ECDFs, defined as the statistical characterization of the lobe stacking pattern, and the similarity between the statistical characterizations of two stacking patterns is defined as the statistical external similarity. A demonstration of statistical characterization of the sample dataset (Fig. 3.12 a) is shown in Figure 3.13. All the stacking patterns must be normalized with procrustes analysis (Section 3.2.2) for the whole stacking pattern, such that patterns measured at different scales are comparable.

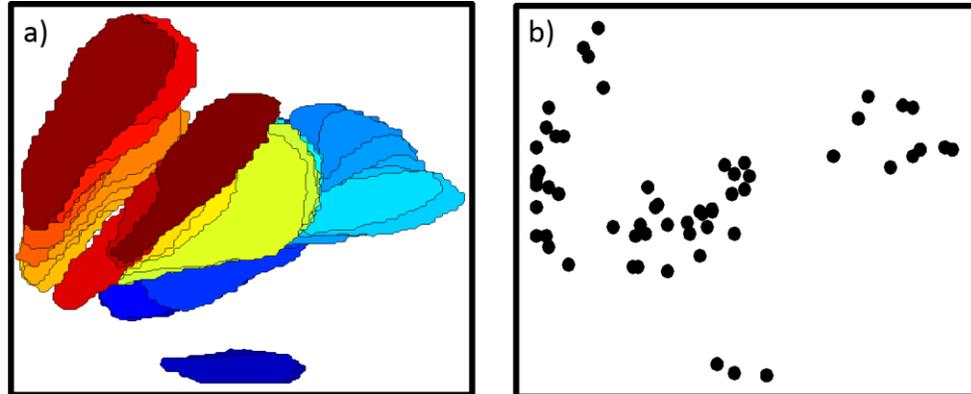


Figure 3.12: a) An exemplary set of 60 lobes, which is a super set of the sample lobe set in Fig. 3.9. b) The point process of MPP of lobes in Fig. 3.12 a.

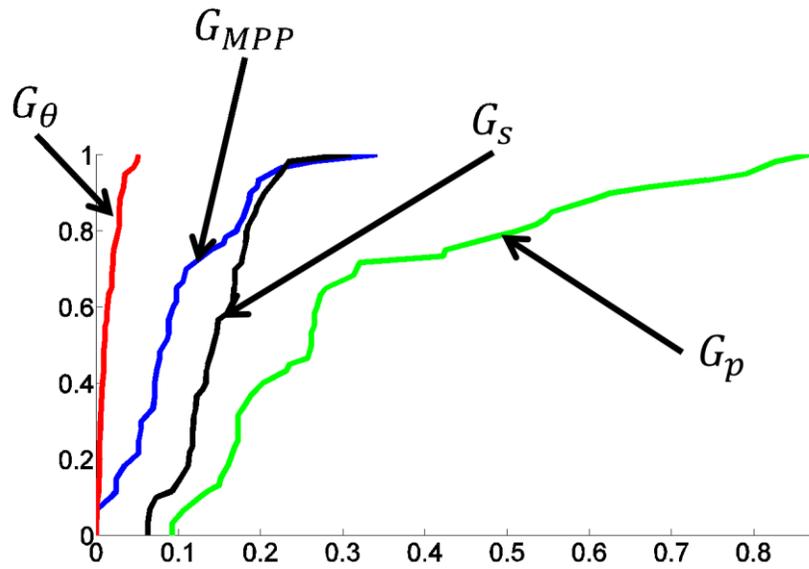


Figure 3.13:  $G$  functions (normalized using procrustes analysis) of the spatial pattern in Fig. 3.12.

### Bootstrap Hypothesis Test for Two Samples

Since each stacking pattern is characterized by four ECDFs, the problem of comparing the statistical external similarity between two stacking patterns is now converted into a two-sample hypothesis test. This test estimates the probability of the ECDFs from two stacking patterns sharing the same underlying population distribution. This hypothesis test is performed for each respective parameter. In cases of lobate depositional environments, the population distributions are considered unknown due to the complexity of the environment and the lack of data. Hence, the nonparametric hypothesis testing techniques are appropriate for this work. Several techniques of nonparametric hypothesis testing are available in the literature, including the Kolmogorov-Smirnov (K-S) test, the Cramer-Von Mises (Cramer) test, and the Pearson's Chi-Square (Pearson's) test (Corder 2009). The K-S and Cramer tests work for medium and large datasets, and Pearson's test is designed for categorical parameters. In the case of lobate environments, a stacking pattern may be a small-size sample of

lobes. Thus, a bootstrap (Efron and Tibshirani 1994) two-sample hypothesis test is devised based on L1-norm measures between ECDFs, such that the test does not rely on parametric distributions and is robust for small samples.

The test involves a sample  $A$  of size  $M$  and a sample  $B$  of size  $N$ . The null hypothesis and the alternative hypothesis are

$H_0$ : Sample  $A$  and  $B$  are drawn from the same cumulative distribution

$H_a$ : Sample  $A$  and  $B$  are drawn from different cumulative distributions

The testing procedures are as follows

Step 1: Merge the two samples into a pool of size  $(M + N)$ .

Step 2: Calculate the L1 norm difference between Sample  $A$  and  $B$ ,  $d_{A,B} = f_{L1}(\widehat{F}_A, \widehat{F}_B)$

Step 3: Draw a bootstrap sample of size  $(M + N)$  observations with replacement from the merged pool.

Step 4: Estimate the ECDF  $\widehat{F}_M$  of the first  $M$  observations and the ECDF  $\widehat{F}_N$  of the remaining  $N$  observations, and calculate the L1 norm difference  $d_{M,N} = f_{L1}(\widehat{F}_M, \widehat{F}_N)$ .

Step 5: Repeat Step 3 and Step 4 for  $K$  (e.g., 3000) times and obtain  $K$  L1-norm difference values.

Step 6: The p-value of the similarity of accepting the null hypothesis is

$$p_s = 1 - \frac{\text{number of } (d_{M,N} < d_{A,B})}{K} \quad (3.7)$$

Distinguished from conventional logic of hypothesis testing, this method also does not require a significance level rejecting the null hypothesis. In this method, the output p-value is interpreted as the probability that one stacking pattern can be used as a correct reference to the other. For example, if a high p-value is obtained from the

hypothesis test with one sample from a geomorphic experiment and one sample from a real scale depositional system, it is safer to apply inferences from the experiment to model the real system than using inferences from an alternative experiment with lower p-values. Since more comprehensive data of initial conditions, intermediate states, and the final stratigraphy are available in geomorphic experiments, making more precise quantitative depositional rules or even quantitative geological dynamic models becomes feasible. To the extent of this study, p-values are used as relative values. A stacking pattern with higher p-values is interpreted as better than a stacking pattern with lower p-values. The work does not attempt to interpret the physical meaning of p-values (e.g., the physical meaning of 0.4, 0.8 etc.).

### **Example**

Three sets of lobe samples were selected to demonstrate our hypothesis-testing algorithm (Fig. 3.14). The sizes of Sample 1, 2, and 3 were 13, 16 and 13, respectively. The sample sizes were designed to be small so that performances of the algorithm on small samples could be demonstrated. Sample 1 and Sample 2 were from the same geomorphic experiment and Sample 3 was extracted from a different experiment. Lobes in Sample 1 and 2 exhibited a high intensity of spatial-temporal clustering, where MPPs were located in a small region and the lobe orientation varied within a relatively narrow range. The shape of lobes was generally similar as well. Since Sample 3 was from an alternative experiment, the stacking pattern was significantly distinguishable from other samples in the sense that the MPPs were more randomly distributed, the orientation varied in a wider range, and the shapes demonstrated obvious distinctions from one to the other. Differences in stacking patterns led to different statistical characterizations (Fig. 3.15). A visual comparison reveals that all  $G$  functions of Sample 1 and 2 are more similar than those of Sample 3. The statistical similarity between Sample 1 and Sample 2, as well as between Sample 1 and Sample 3, were compared using the two-sample bootstrap hypothesis. Resulting p-values are listed in Table 3.1, in which p-values

between Sample 1 and 2 are completely higher than between Sample 1 and Sample 3, proving that the method is capable of identifying similar stacking patterns.

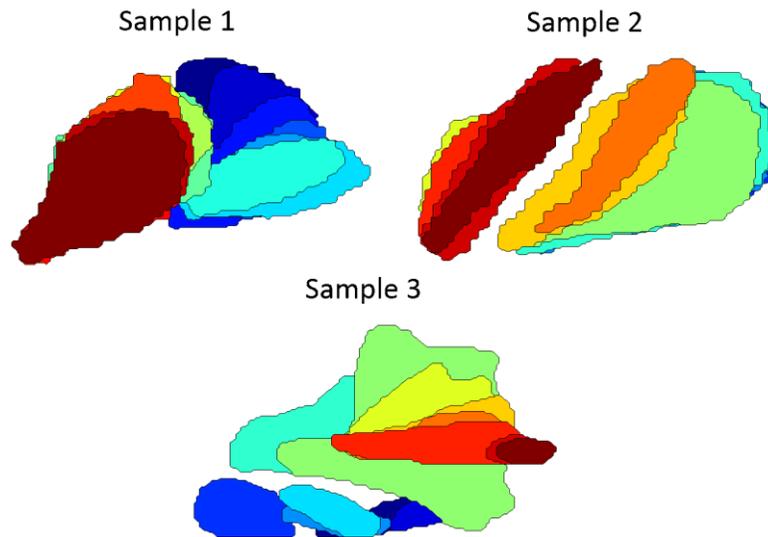


Figure 3.14: a) Lobe Sample 1 of 13 lobes; b) Sample 2 of 16 lobes; c) Sample 3 of 13 lobes. Sample 1 and Sample 2 are from the same depositional system.

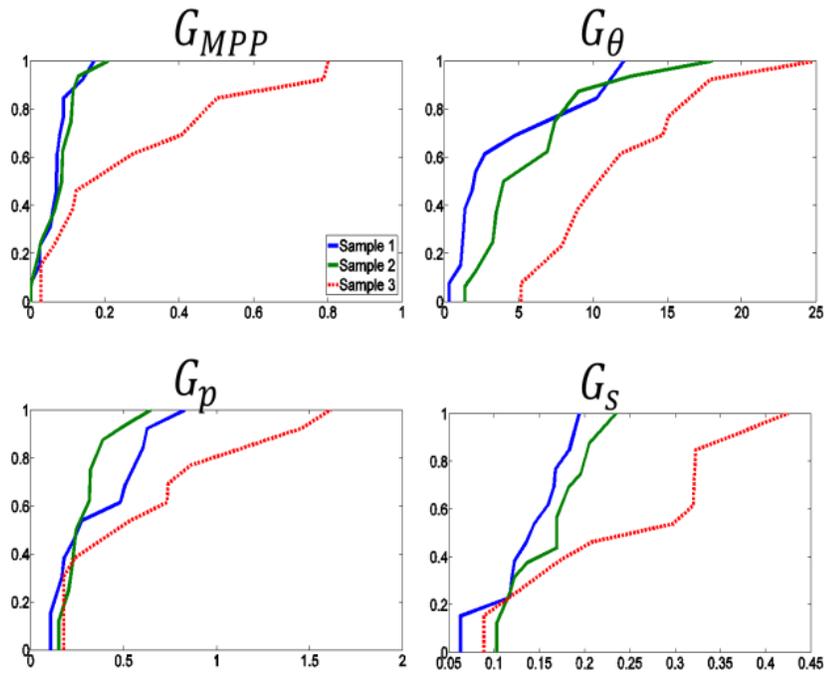


Figure 3.15:  $G$  functions for all three samples, the samples are normalized using procrustes analysis so length measurements from different systems are comparable.

P Values	Sample 1 vs. Sample 2	Sample 1 vs. Sample 3
$P_{MPP}$	0.58	0
$P_{\theta}$	0.076	0
$P_p$	0.026	0.0075
$P_s$	0.042	0

Table 3.1: P-values of hypothesis testing between samples. Sample 1 and 2 are detected to be similar, which Sample 3 is different from 1 and 2.

### 3.3 Application

The proposed algorithms were demonstrated to be capable of quantifying a hierarchy of stacking patterns and estimating the statistical external similarity between stacking patterns. For the application of a scale-dependent analysis of statistical external similarity between an experiment and a real depositional system, the general workflow of this method is demonstrated in Fig. 3.16. Inputs to the workflow included a lobate stacking pattern of a real-scale depositional system and a stacking pattern interpreted from overhead photos from a geomorphic experiment. In the first step, hierarchical agglomerative clustering was applied on the experimental stacking pattern. The hierarchy of experimental lobes was quantified by a dendrogram and reachability plot. Then, the reachability plot was thresholded regularly on the y-axis (scales). At each threshold, the lobes were clustered into several groups. Each group was a higher-scale lobe of the scale of interpretation determined by the threshold. A lobe stacking pattern was also obtained at the scale. Each experimental stacking pattern was compared to the real scale pattern with the bootstrap hypothesis test based on L1-norm distance. At every threshold, a p-value was estimated to measure the similarity. Since the thresholds varied from very fine to large scales, the p-values were estimated from fine scales to large scales as well. Finally, the relationship between p-values and the scales of interpretation was described by a plot. The expected result was that peaks of p-values might be observed at certain scales of interpretation. Such plots were generated for all four parameters. If peaks of all plots were at a common scale of interpretation, the implication was that experimental lobes interpreted at the scale of the peaks are equivalent to lobes in the real system.

Two stacking patterns from real-scale lobate systems were identified from sedimentology literatures (Fig. 3.17 a, b). One stacking pattern was from a basin-floor fan in Kutai Basin, East offshore Kalimantan, Indonesia (A. Saller et al. 2008). The basin floor fan included 18 lobes, whose length ranged from 3 to 12 km and width ranged from 1 to 7 km. The other stacking pattern was extracted from the Amazon fan

(Jegou et al. 2008), including 14 lobes with lengths ranging from 21 to 83 km and widths ranging from 7 and 25 km. The two patterns exhibited different characteristics in terms of stacking patterns. In general, lobes in the Kutai basin exhibited more intense overlaps, while the lobe orientations switched in a relatively large range from one lobe to the other. Lobe shapes varied from long and narrow to oval and wide. While in the Amazon fan, very few lobes overlapped the others. Lobe orientations and shapes were both within a very narrow range. Two hierarchies of lobe stacking patterns were interpreted from two geomorphic experiments, which were provided by the Sedimentology team, San Anthony Falls National Laboratory (Fig. 3.17 c,d), Minnesota. Since the intermediate elevation data was not measured at fine frequencies, the lobe patterns were interpreted from a fraction of the intermediate photos, in which clear lobate geometries were identifiable at intervals of 30 seconds. The first experiment (referred to Exp.A in the following sections) included a sequence of 424 lobes. The sediment cohesion in Exp. A was designed to be unrealistically low, thus normally channels were short and lobes randomly aggradated on previous deposits. The second experiment (referred to Exp. B) included 184 lobes. Because the sediments were designed to be more cohesive, longer channels were formed during the interval of two avulsions. Thus new lobes tended to appear near the channel mouth, such that the pattern exhibited more intense stacking. The reachability plots (Fig. 3.18) quantify the hierarchies of Exp.A and Exp.B. More small scale valleys are observed in the chart of Exp. A, implying that the sizes of spatial-temporal clusters are smaller than in Exp. B.

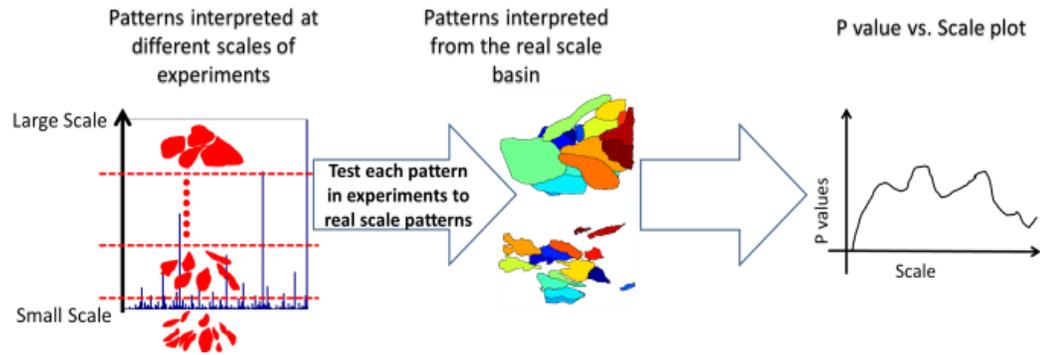


Figure 3.16: The workflow of scale-dependent similarity analysis. The result is a p-value versus a scale plot for each parameter.

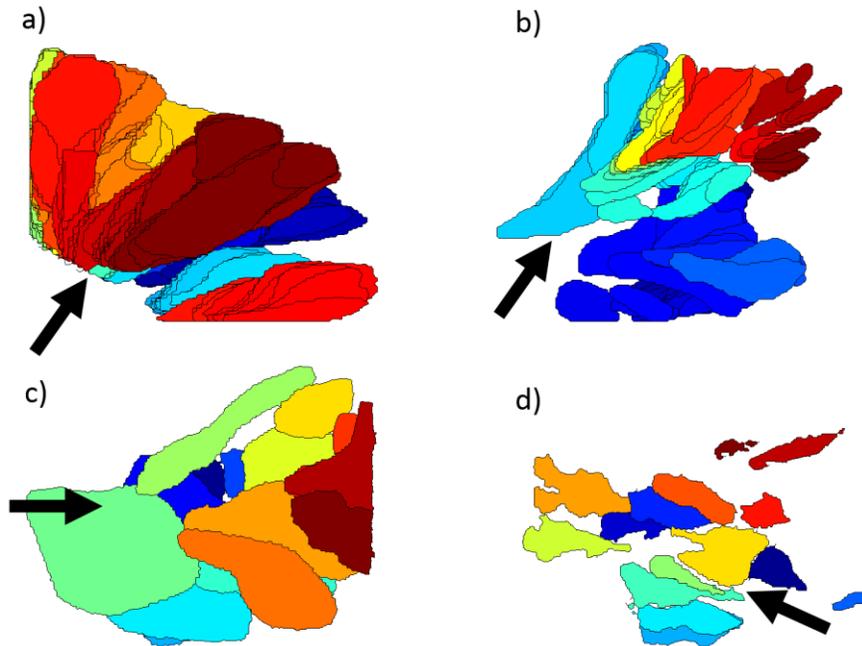


Figure 3.17: Lobe patterns used in the application: a) stacking pattern of Exp. A; b) stacking pattern of Exp. B; c) stacking pattern of Kutai basin; d) stacking pattern of Amazon fan.

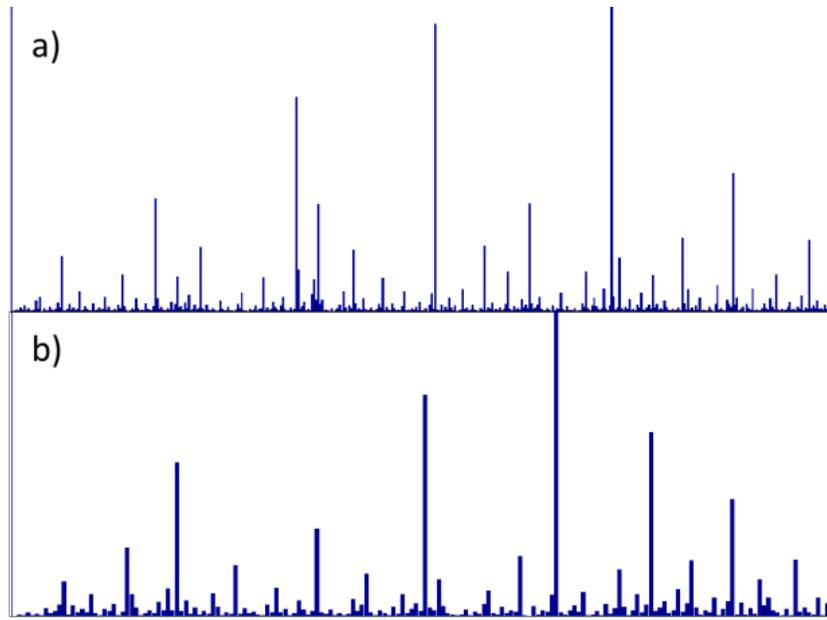


Figure 3.18: Hierarchies of experiments: a) hierarchy of Exp. A; b) hierarchy of Exp. B.

All stacking patterns (experimental and real) were normalized with procrustes analysis so that patterns at different scales became comparable. Then, 60 thresholds at regular intervals from the finest lobes were taken on the y-axis of the reachability plots of Exp. A and Exp. B. In the third step, stacking patterns at every scale of interpretation were obtained by taking the union of lower-scale lobes in each group. Finally, stacking patterns at each scale were compared to the respective real-scale stacking patterns. In this manner, p-values were generated as a function of the scales of interpretation. For the purpose of interpretation, the length unit was converted with respect to the real-scale system. Assuming a stacking pattern of  $N$  lobes, the rescaling factor is the RMSD of the reference real-scale system. The resulting p-values vs. the scale of interpretation plots comparing both experiments to the Kutai basin are plotted in Fig. 3.19. Fig. 3.20 contains the Amazon fan plots.

The experimental stacking patterns exhibited significant distinctions in p-values at various scales compared to both real cases. In both cases, low p-values for the MPP

and  $\theta$  were observed spanning most of the plotted range in Exp. A (Fig. 3.19 a and b, Fig. 3.20 a and b). While higher p-values were observed for the polygonal distance and the shape in both cases (Fig. 3.19 c and d, Fig. 3.20 c and d), the peaks spanned a wide range of scales. Since no common scale of high p-values was obtained for all parameters, Exp. A was interpreted as having no value for providing applicable information to the Kutai basin or the Amazon fan. On the contrary, plots of Exp. B versus the real cases were more informative. Peaks of p-values of all plots located within a narrow range centered at the scale of 15 km (Fig. 3.19). The range of scales for the four peaks was from 14 km to 17 km (the shaded region in Fig. 3.19), implying that lobes interpreted within this range in Exp. B are equivalent to lobes in the Kutai basin. Although peaks were observed in comparing Exp. B to the Amazon fan as well, scales of the peaks ranged from 14 km to 62 km, covering most portions of available scales of interpretation in the plots. Stacking patterns for four peaks in Exp. B compared to the Kutai basin are shown in Fig. 3.21. The stacking patterns for peak scales for the Kutai basin were not only visually similar but also contained a similar number of lobes (from 30 to 38 lobes). Stacking patterns of Exp. B compared to the Amazon fan are shown in Fig. 3.22. The numbers of lobes at the peak scales ranged from 22 lobes to 65, which is significantly different. Hence, it is safer to use Exp. B to provide information to model the Kutai basin. Lobate geobodies interpreted in the range of 14 km to 62 km in Exp. B were equivalent to the scale of lobes in the Kutai basin.

In summary, Exp. A, as physically designed, is unrealistic, thus no information from this experiment are referable to the two real scale systems. Exp. B is more realistic to all realistic systems than Exp. A in the physical design, thus it has higher similarity than Exp. A to both real scale cases. However, Exp. B cannot provide referable information for the Amazon fan either. Because the petrofacies information of neither Kutai basin nor Amazon fan were discussion in the publications, we can only interpret this in the perspective of lobe spatial distribution. Generally, channels in the Amazon fan are obviously longer than those in the Kutai basin and in Exp. B and lobe distribution

are sparse relative to lobe distributions in the Kutai basin and Exp. B, hence, the high p-values are inconsistent in all four parameters.

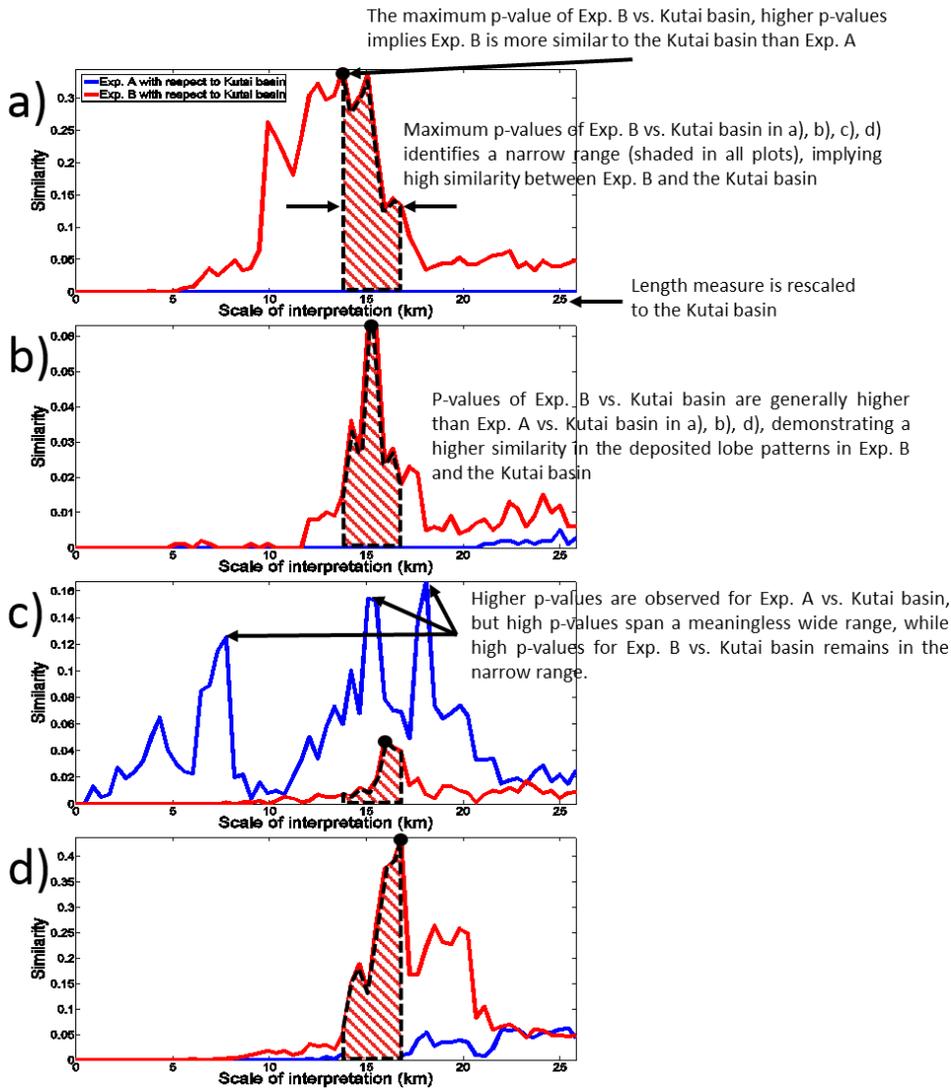


Figure 3.19: P-values versus scales of interpretation plots with respect to the Kutai basin. The distance measurements are rescaled with RMSD of the Kutai basin. Peaks of p-values are marked out by the black dots and a narrow range for scales of interpretation with the highest p-values for all four parameters (shaded regions) is identified. a) MPP; b) Orientation; c) Polygonal distance; d) Shape.

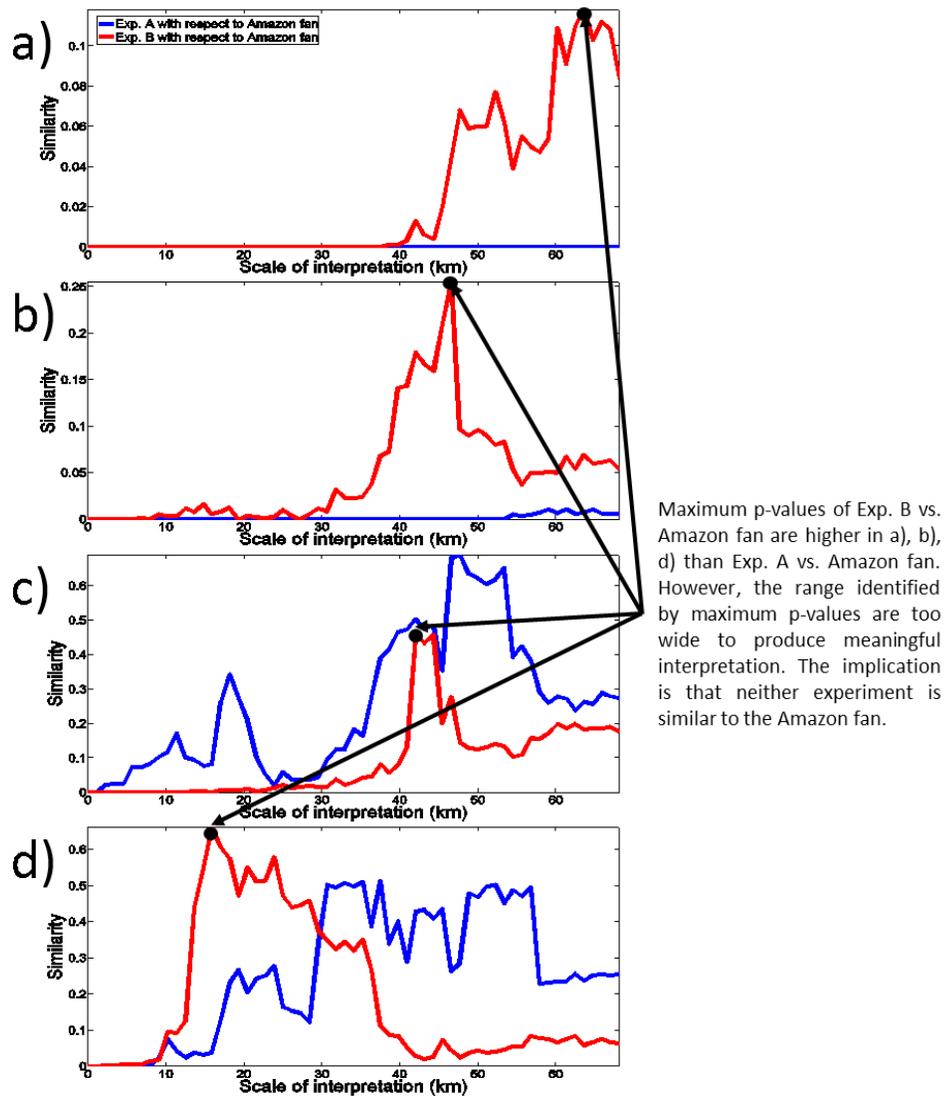


Figure 3.20: P-values versus scales of interpretation plots with respect to the Amazon fan. The distance measurements are rescaled with RMSD of the Amazon fan. Peaks of p-values are marked out by the black dots, but the peaks vary in a wide range of scales. a) MPP; b) Orientation; c) Polygonal distance; d) Shape.

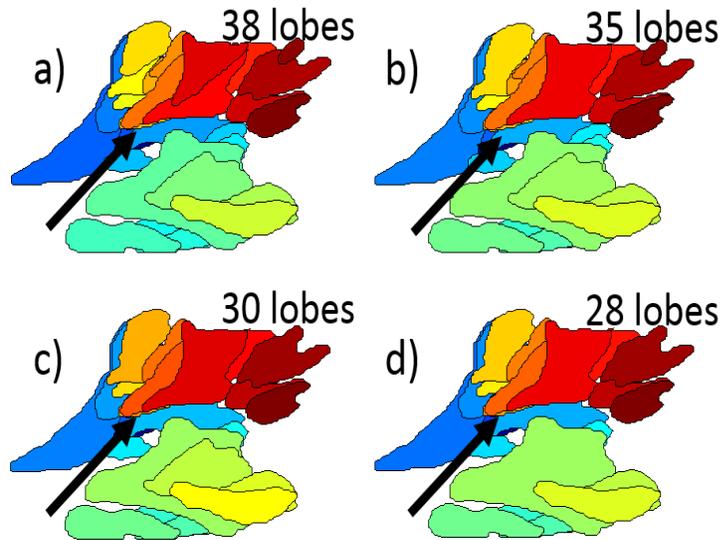


Figure 3.21: Stacking patterns from Exp. B at peaks of p-values with respect to the Kutai basin: a) patterns at the scale of interpretation with the highest p-value for the MPP (38 lobes) at the scale of 13.8km; b) patterns at the scale of interpretation with the highest p-value for  $\theta$  (35 lobes) at the scale of 15.1km for; c) patterns at the scale of interpretation with the highest p-value for polygonal distance (30 lobes) at the scale of 15.8km; d) patterns at the scale of interpretation with the highest p-value for shape (28 lobes) at the scale of 16.8km.

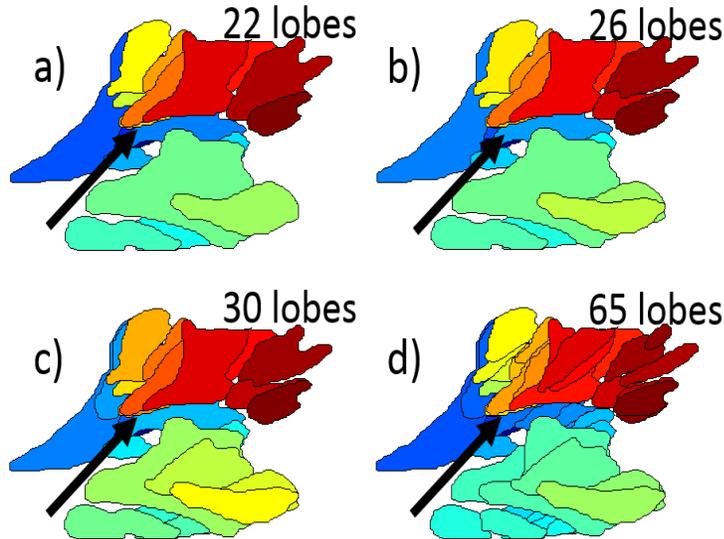


Figure 3.22: Stacking patterns from Exp. B at peaks of p-values with respect to the Amazon basin: a) patterns at the scale of interpretation with the highest p-value for the MPP (22 lobes) at the scale of 62km; b) patterns at the scale of interpretation with the highest p-value for  $\theta$  (26 lobes) at the scale of 47km for; c) patterns at the scale of interpretation with the highest p-value for polygonal distance (30 lobes) at the scale of 42km; d) patterns at the scale of interpretation with the highest p-value for shape (65 lobes) at the scale of 14km.

### 3.4 Chapter Summary

This chapter presented a workflow of identifying a proper geomorphic experiment and the proper scale to interpret the experiment with respect to a stacking pattern interpreted from a real-scale depositional system. The workflow is based on data mining and statistical analysis, and is proved to be capable of generating reasonable results by applications on two experiments and two real systems. In addition, several techniques to quantify geological information, such as the stacking pattern and hierarchy of lobes, were introduced in the workflow, which are useful tools for quantitative studies on lobate systems.

The proposed workflow also leads to future work. First, the stacking patterns were quantified by four parameters, but not all of them could effectively distinguish one stacking pattern from another. One potential improvement could be made on the measurement of polygonal proximity. In this study, the p-values versus the scale of interpretation plots of polygonal proximity, measured using the Hausdorff distance (Fig. 3.19 c and Fig. 3.20 c), were noisy and not as informative as the other three parameters. A possible reason is that the areal proximity measurement should be capable of catching the degree of overlap of two small lobes, which is not considered by Hausdorff distance. Second, the hierarchy of lobe systems was quantified by a dendrogram or reachability plot, which is considered as a quantitative summary of the system. It would be worth developing a technique for measuring the differences between two dendrograms, which would reflect the overall statistical similarity between two lobate systems. Third, the current L1-norm measurement based bootstrap hypothesis test was designed for samples of a small size. The method outperformed a K-S test with small samples; however, p-values estimated with this bootstrap test might be underestimated compared to those estimated with the K-S test. Moreover, the robustness of this hypothesis testing method should be further estimated through applying it to stacking patterns from a large number of experiments, such that the situations in which the p-values are reliable can be determined. Finally, the model identified a scale of interpretation within an experiment

equivalent to lobes at the seismic scale of a real system. Since the interest of static reservoir modeling is to estimate the subseismic scale lobe stacking patterns, available in the quantified experimental hierarchy, the next step would be to develop a simulation algorithm that honors the hierarchical information.

# Chapter 4

## **HIERARCHICAL SIMILARITY OF LOBES: SIMULATION AND COMPARISON**

### **4.1 Introduction**

The results of Chapter 3 are a set of lobes from one experiment that is the most similar to a specific real depositional system within multiple optional experiments. A reasonable assumption is that the geological dynamic process that produces this similarity is itself also similar to the process that generates the real lobe stacking pattern. Based on this assumption, this chapter is dedicated to an alternative implementation of lobe migration dynamic processes, designed to honor information provided by the chosen experimental lobe stacking pattern. The implementation was designed to involve the least number of empirical coefficients so that the model is favorable for uncertainty estimations. As indicated earlier, the frequency of three-dimensional intermediate topography measurement is not sufficient in this experiment, thus the method developed

in this chapter only considers lobe stacking patterns extracted in the two-dimensional horizontal plain.

The most valuable information from an experimental lobe stacking pattern is the sequential migration behavior of lobes. Current lobe migration rules (Fig. 2.16) in surface-based models determine placement of lobes based on conceptual understandings about depositional systems, normally interpreted from field and seismic surveys and the general physical rules. The lack of intermediate topography and the limited amount of lobes maintained in field-scale depositional systems limit more precise understanding of the depositional process and thus more precise quantitative depositional rules. Current surface-based models with conceptual depositional rules are sufficient for applications in the verification or demonstration of hypothesized conceptual rules. However, in applications for engineering purposes, e.g. conditioning, uncertainty estimation, history matching, disadvantages of conceptual depositional rules are obvious. First, the computational implementation of conceptual rules involves a modeler's subjective decisions regarding conversion functions from qualitative concepts, which are normally sensitive to model realizations. Therefore the choice of conversion functions directly increases the uncertainty estimations. On one hand, the ranges of values of empirical coefficients of the conversion functions cannot be narrowed down because there is no sufficient data to quantitatively calibrate the algorithms of conceptual rules. To test uncertainty of a wide ranges of values is an incredible workload. For example, the probability component reflecting impacts of the basin source (Fig. 2.16) is converted from a distance-to-basin-source map, in which the value in every cell of the modeling grid is its distance to the basin source point. The conversion function from the distance to the probability is the modeler's decision, such as logarithmic, gaussian, or exponential functions of the inversed distance etc. Each conversion function has its own coefficients directly affecting the simulated spatial distribution of lobes, e.g. the mean and variance for the gaussian function. Considering there are usually multiple components in the conceptual depositional model, a handful of scenarios and tens of empirical coefficients must be considered in the uncertainty

estimation. On the other hand, the interactions between parameters are extremely unpredictable in a computational implementation of conceptual models with too many empirical parameters. Thus, the response of a surface-based model is probably dominated by noise, leading to difficulty in any inversion problem with the model, such as conditioning. Finally, conceptual depositional rules in current surface-based model implementations rarely attempt to account for any specific input patterns, such as the stepwise characteristics or overall hierarchy of a lobe migrating sequence from an experiment. The objective of this chapter is to present an improved computational implementation for lobe migration in order to minimize the subjective choices of quantification functions and empirical coefficients with a simple algorithm that is capable of using prior information from an input lobe sequence.

The precise lobe migration sequence provides more specific information on the depositional process, which is valuable in the design and calibration of a lobe migration algorithm, such that the algorithm design can be improved, e.g. to reduce the number of parameters, to test and fix proper values of empirical coefficients. The essential meaningful information provided by a lobe migration sequence is the stochastic relationship between intermediate topography and lobe migration. Since the migration sequential information is limited in the horizontal plain, the first task is to define a characterization of the two-dimensional lobe migration sequence such that the surface-based model can account for geologically meaningful information from the input sequence and the intermediate topography (surfaces) in every step of the simulation. In this chapter, the lobe migration is described with a correlated random walk (CRW) (Section 4.2), in which the stepwise behaviors and overall hierarchies of the input stacking patterns are statistically reproduced. In the CRW-based strategy, empirical coefficients are used as infrequently as possible. The new strategy, demonstrated here, is an implicit control on the hierarchy of the lobe MPP. A dissimilarity measure is proposed to validate the reproduction of the MPP hierarchy between two lobe stacking patterns, with which it is proved that MPP hierarchies of model realizations are effectively constrained (Section 4.3). The significance of the CRW-based lobe

migration strategy is that the prior depositional dynamic process model can be updated effectively by replacing input migration sequences analogous to updating training images (conceptual models of finalized petrofacies) in multiple-point statistics, which is meaningful in the rapid updating strategy for history matching (Caers 2013).

## **4.2 A Lobe Migration Model with Input Lobe Sequence**

This section illustrates the implementation of a CRW-based lobe migration mechanism. As aforementioned, the model is expected to honor an input lobe stacking pattern. Since the input experimental stacking pattern is normally at a different spatial scale to the simulated depositional basin, a normalization and rescaling factor is required to convert any spatial measurements from input experiments to the real system to be modeled. The Root Mean Square Distance (RMSD), a statistical measure of the scale of a spatial pattern, is proposed to be the normalization factor (Section 4.2.1). Instead of using the pairwise interlobe proximity of the finalized stacking patterns (Chapter 3), statistics of stepwise pairs in the input lobe migration is considered in this section (Fig. 4.1).

Temporal and spatial information in the horizontal two-dimensional domain of the sequence was precisely recorded in the inputs; however, the intermediate topographic information was available at every step of the simulation besides two-dimensional spatial and temporal information. This led to the requirement of accounting for two types of information: the conceptual compensational stacking rules in response to the intermediate topography and the precise statistics of migration in the two-dimensional horizontal domain. In the method, a bifactor modeling scheme was proposed to combine both types of information (Section 4.2.2). While characterizing and simulating the lobe migration pattern, the focus was on the MPP of the lobes and the lobe orientation (Section 4.2.3 – 4.2.4), which were the most important parameters of the spatial distribution of the lobes. Simulation of the polygonal distances and shape similarities were not considered in this work. Finally, an exemplary run of the model

was performed to demonstrate the new surface-based model with new lobe migration mechanism (Section 4.2.5).

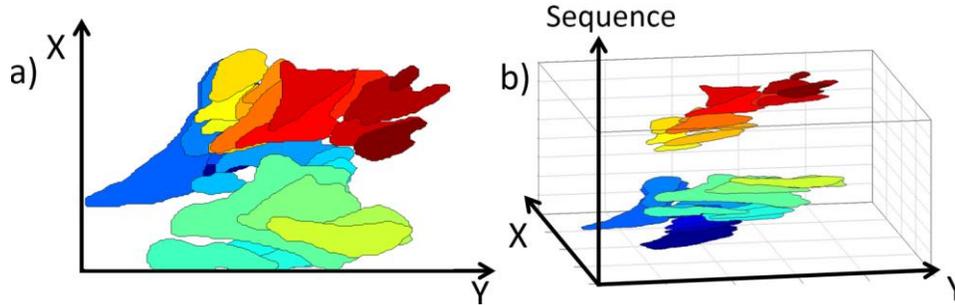


Figure 4.1: a) The plane view of an experimental stacking pattern with the high p-values in Chapter 3, including 38 lobes. b) The sequence view of the experimental stacking pattern, which is explicitly input to the algorithm in this chapter.

#### 4.2.1 Root Mean Square Distances: The Scaling Factor for Stacking Patterns

The input stacking patterns were obtained from geomorphic experiments, in which the length scale was much smaller than that of a real-scale depositional system. As an extension of the strategy of measuring statistics of dimensionless ratios presented in Chapter 1, using a uniform scaling factor that normalized length measurements in both the input stacking pattern and the simulation realizations for a dimensionless space was proposed. The lobe migration sequence was modeled and simulated in the dimensionless space and was rescaled to the unit of the real depositional system after the simulation. The proposed uniform scaling factor was the RMSD. RMSD is a popular statistical measure of the scale of a set of data and is adopted in procrustes analysis for the analysis of shapes of different sizes (Dryden and Mardia 1998; Kendall 1989). For this application, the Mean Square Distance (MSD) was calculated between any point  $p_i(x_i, y_i)$  on the interpreted basin boundary to  $p_c(\bar{x}, \bar{y})$ , the geometric centroid of the basin. The RMSD was simply the square root of the MSD, such that the value was converted from a real unit to a length unit (Eq. (4.1)). Any distance measurement,  $d_{exp}$ , in the experiment was converted to the dimensionless space by scaling it with the RMSD

of the input stacking pattern  $RMSD_{exp}$ , the simulated realizations in the dimensionless space were converted to the real system scale with the RMSD of the real basin  $RMSD_{real}$  (Eq. (4.2) – (4.3)).

$$MSD = \sqrt{\frac{\sum_i [(x_i - \bar{x})^2 + (y_i - \bar{y})^2]}{k}} \quad (4.1)$$

$$\hat{d} = \frac{d_{exp}}{RMSD_{exp}} \quad (4.2)$$

$$d_{real} = \frac{\hat{d}}{RMSD_{real}} \quad (4.3)$$

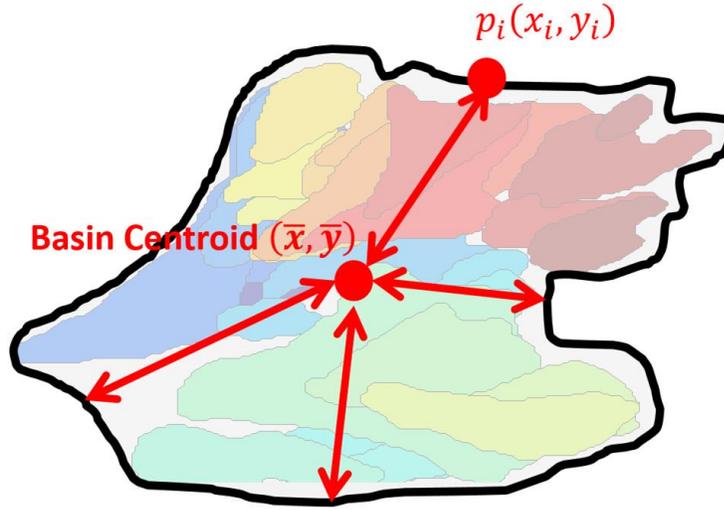


Figure 4.2: Root Mean Square Distance (RMSD) as the scaling factor for stacking patterns. RMSD is determined by the averaged squared distance from the geometric centroid to the interpreted basin boundary. It is a measure of the scale of a stacking pattern and is used as the scaling vector to normalize stacking patterns. All length measurements and coordinates from the experiment are normalized by the RMSD of the experimental pattern, and the simulation is performed in the dimensionless space. The RMSD of the field basin can be used to scale the realizations to the basin scale. Refer to Eq. (4.1) – (4.3).

#### 4.2.2 The Bifactor Modeling Scheme

The first task of modeling a lobe migration sequence was the strategy of accounting for two types of information, including precise statistics from the input horizontal spatial lobe migration sequence and the conceptual compensational stacking

rule in response to the intermediate topography at every step of simulation. From the perspective of the depositional process, the spatial distribution of sediments was governed by the compensational stacking rule, which was a stochastic process endeavoring to transport sediments into lower regions such that the sediment basin was turned into a horizontal plain. On the contrary, accounting for an input lobe sequence implied modeling the quantitative relationship between intermediate topographies and the spatial location of a new lobe. Since the intermediate topography information is not available in the experiment data used for this chapter, building such a stochastic relationship is impossible; hence, the lobe migration mechanism in this chapter only describes the movement of the input lobe sequence without considering the real relationship between intermediate topographies and lobe locations. This scheme is in analogy to the kinematics in classical mechanics, in which only the motions, but not the cause of motions, are considered. The effect of the intermediate topography was considered by combining the conceptual compensational stacking rule at every step of simulation. In detail, the component of conceptual compensational stacking rule always tend to place a new lobe toward the steepest direction away from the previous one, while the input lobe sequence component determine the next MPP location based on statistical behavioral characteristics of the input. The combined modeling scheme of the two factors is formulated in Eq. (4.4). The combination weight was the first empirical coefficient introduced in this modeling strategy to reflect the modeler's judgement on the importance of the two factors.

$$g = b * g_{pattern} \oplus (1 - b) * g_{comp} \quad (4.4)$$

where  $g_{pattern}$  is the prior information from the input lobe sequence;  $g_{comp}$  is the prior information from the conceptual compensational stacking rules;  $\oplus$  represents the combination of two types of priors;  $b$  is a combination weight,  $b = 1$  implies total control by the lobe sequence,  $b = 0$  implies total control by the compensational stacking rules.

### **4.2.3 The Input Stacking Pattern Factor**

#### **Parameterization of Lobe Sequences**

The basic requirement of the characterization strategy was the simplicity, implying both the simplicity of parameterization of the input movements and the resulted simplicity in the mechanism implementation. A geological model has to be as simple as possible to reduce the workload in uncertainty estimations and history matching. Another requirement is that the characterization of a lobe migration sequence should produce geologically reasonable features, such that the simulation is not over constrained to the input sequence. In lobate environments, the depositional process is featured by the stepwise relative movements of lobes, such as progradation, retrogradation, and lateral shift. Progradation and retrogradation are the successive basinward or landward trends of lobe migration movements, in which a new lobe location is defined relative to the previous one. The basic assumption behind studying relative lobe movements is the directional persistent behavior. A conceptual explanation for the directional persistence assumption, also for the previous lobe location component in Chapter 2 (Fig. 2.16), is that a sediment flow tends to follow the channel through which the previous flow went. When the new flow passes through the channel-to-lobe transition point of its previous lobe, the local slope, flow condition, and strength of the channel levee determine either the formation of a new lobe or the continuation of channelization toward the downstream direction. Generally speaking, the change of slope around the channel mouth is drastic, and the channel levee is not well formed. Hence, a new lobe normally occurs in the intermediate vicinity of the previous lobe. Some of the features, such as the migration direction (progradation or retrogradation) and lateral shifting direction, are related to the intermediate regional topographic and fluid conditions. However, other features, such as the statistics of the stepwise migration length and lateral shifting angle of the lobe orientation, can distinguish one lobate system from another, although the physical explanation remains unclear. For example, some lobate systems are characterized by sustained progradation or retrogradation with frequent short migration lengths and a small shifting angle, while others are

characterized by an almost completely random lobe distribution with frequently large migration steps and shifting angles. A strategy to characterize lobe migration behavior with the direction persistence assumption and orientation changes of every single lobe is presented here. The strategy also assumes that lobe migration behavior is stationary and homogeneous within an input lobe sequence. Consistent to the workflow in Chapter 3, MPP is used as the primary parameter indicating the location of a lobe. Thus, the migration distance  $\Delta r$  from the current MPP to the next is measured as the stepwise lobe migration distance (Fig. 4.3 a). A MPP migration shifting angle,  $\Delta\theta_{MPP}$ , is required to determine the coordinate of the new MPP.  $\Delta\theta_{MPP}$  for the step  $t + 1$  is the angle between the previous MPP migration direction, defined as the vector from  $MPP_{t-1}$  to  $MPP_t$ , and the new MPP migration direction, defined as the vector from  $MPP_t$  to  $MPP_{t+1}$  (Fig. 4.3 b).  $\Delta\theta_{MPP}$  is measured from 0 to  $2 * \pi$  for the purpose of recording the precise movement pattern. Besides MPP migration, the lobe orientation is also an important parameter determining the exact location of a lobate geobody. Lobe orientation is defined as the unit vector from MPP to the geometric centroid of a lobe in the horizontal two-dimensional plain. Conceptually, the relative movement between lobe  $t$  and lobe  $t + 1$  is affected by the basinwise topography and overall flow orientation, but in the meantime it is stochastic in response to the regional topography. Instead of measuring it relative to the previous lobe orientation, the lobe orientation shifting angle is measured relative to the interpreted overall basin flow orientation (Fig. 4.3 c), defined by the unit vector from the basin source point to the two-dimensional geometric centroid of the basin. The lobe orientation shifting angle,  $\Delta\theta_L$ , is measured from 0 to  $\pi$ . For each input lobe migration sequence,  $\Delta r$ ,  $\Delta\theta_{MPP}$ , and  $\Delta\theta_L$  are measured through the sequence, resulting in three ECDFs.

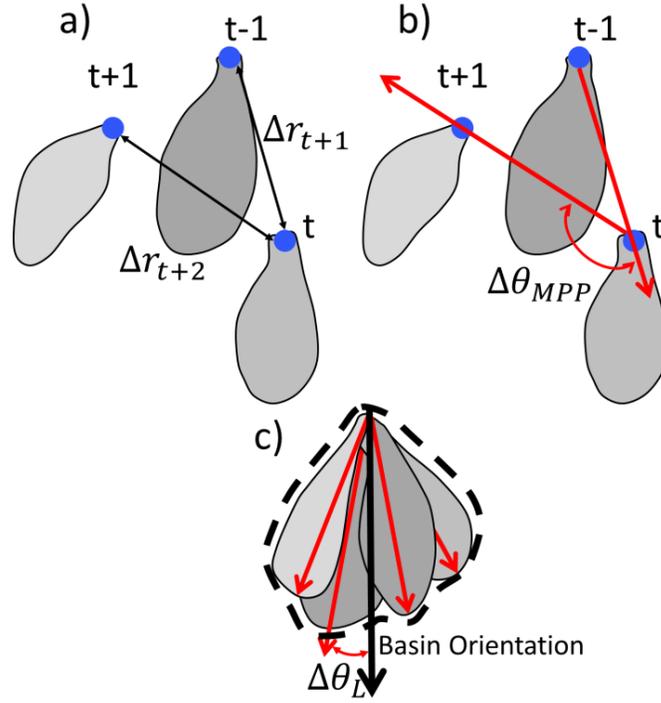


Figure 4.3: MPP migration and lobe relative (to the basin flow orientation) shifting angle are modeled in this chapter. Scale versus stacking pattern plots of the area distance and the shape are noisy, thus the use of them as inputs requires further studies. a)  $\Delta r$ , the stepwise migrating distance of MPP; b)  $\Delta \theta_{MPP}$ , the stepwise MPP migrating orientation, measured from 0 to  $2\pi$ ; c)  $\Delta \theta_L$ , the stepwise lobe relative (to the basin flow orientation) orientation, measured from 0 to  $\pi$ .

### Correlated Random Walk (CRW) for MPP Migration

The lobe sequence characterization with CDFs of the MPP migrating distance  $\Delta r$ , and migration orientation shifting angle  $\Delta \theta_{MPP}$ , naturally leads to the application of the CRW as the modeling technique. CRW broadly refers to the homogeneous and stationary sequential random point movement, in which the location of a new point is determined by the stepwise migrating distance and the turning angles from a previous location and moving direction (Fig. 4.4). In the most common form of CRW, the stepwise migration distance follows a non-negative right-skewed parameterized probability distribution, i.e. weibull distribution (Eq. (4.5)), and the turning angle from

a previous moving direction follows a symmetric distribution, in which values are around zero, i.e., wrapped cauchy distribution (Eq. (4.6)).

$$f(r) = a(r)^{a-1}e^{-(r)^a} \quad (4.5)$$

where the shape parameter  $a$  controls the shape of the distribution;

$$f(\theta) = \frac{1}{2\pi} \left( \frac{1-k^2}{1+k^2-2k\cos[\theta-E(\theta)]} \right) \quad (4.6)$$

where  $k \in [0,1]$  controls the range of the turning angle,  $E(\theta)$  is the mean direction;

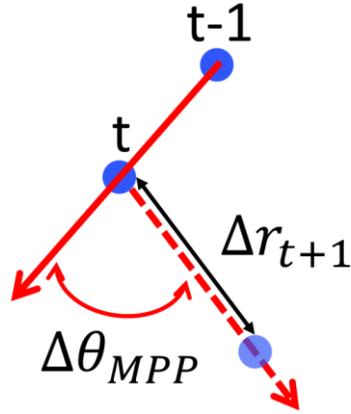


Figure 4.4: The idea of Correlated Random Walk (CRW). CRW simulates spatial point sequences. The point at time  $t + 1$  is determined by deviations from the point at  $t$  by a moving distance  $\Delta r$  and an angle  $\Delta\theta_{MPP}$  deviating from the previous MPP moving orientation.  $\Delta r$  and  $\Delta\theta_{MPP}$  are modeled as stochastic variables with cumulative distribution functions (CDFs). The choice of CDFs depends on the problem. CRW is applied to model MPP migration in this work and controlled by empirical CDFs from the input lobe sequence.

An increment of  $a$  in the weibull distribution, and  $k$  in the wrapped cauchy distribution, increases the probability of getting larger  $\Delta r$  and  $\Delta\theta_{MPP}$  and thus introduces more randomness in the simulated points (Fig. 4.5). Since CRW is simple and has been widely studied, the technique has been applied in various scientific problems. The first CRW model was proposed to analyze and model the migration behavior of butterflies (Kareiva and Shigesada 1983). Since then, and along with the

application of new data obtaining techniques, such as remote sensing image, GPS, etc., the CRW model has been widely applied or modified based on the Kareiv-Shigesada model in studies of animal migration behaviors (Bergman, Schaefer, and Lutlich 2000; Firlie et al. 1998; Forester et al. 2007; Fortin, Morales, and Boyce 2005; Fortin et al. 2005; Fortin et al. 2005; Marell 2002). In a relatively sophisticated example (Fortin, Morales, and Boyce 2005), a CRW with bias to a topography model was applied to study the migration movement of elk. The assumption of their model was that average direction  $E(\theta)$  of the probability distribution of the next turning angle of the elk wandering randomly was a function of two factors: the directional persistence factor and the steepest slope direction. Therefore, they modeled the average moving direction  $E(\theta)$  of the wrapped cauchy distribution for sampling the next turning angle as a weighted mean of the steepest slope direction and the previous moving direction. The distribution functions were fitted to data from 21 elk-migration routes. The idea of CRW with bias to a topography model effectively incorporated the impacts of topography on the migration movement with a simple technical framework, which was adjusted and extended here for the purpose of accounting for topography in the lobe migration mechanism model. The primary modification is that the ECDFs of the MPP migration distance replace the weibull distribution. To determine the MPP migration orientation-shifting angle, the mean migration direction was determined by weighted averaging of a sampled value from the input ECDF of  $\Delta\theta_{MPP}$  and the steepest slope direction. The weighted mean moving direction was used as an average direction in the wrapped cauchy distribution for sampling the final lobe migration orientation. The reason for maintaining the wrapped cauchy distribution in the model was to keep sufficient randomness in the model. The second empirical coefficient  $k$  is introduced in the wrapped cauchy distribution to control the variability of the distribution, or the strength of directional persistence.

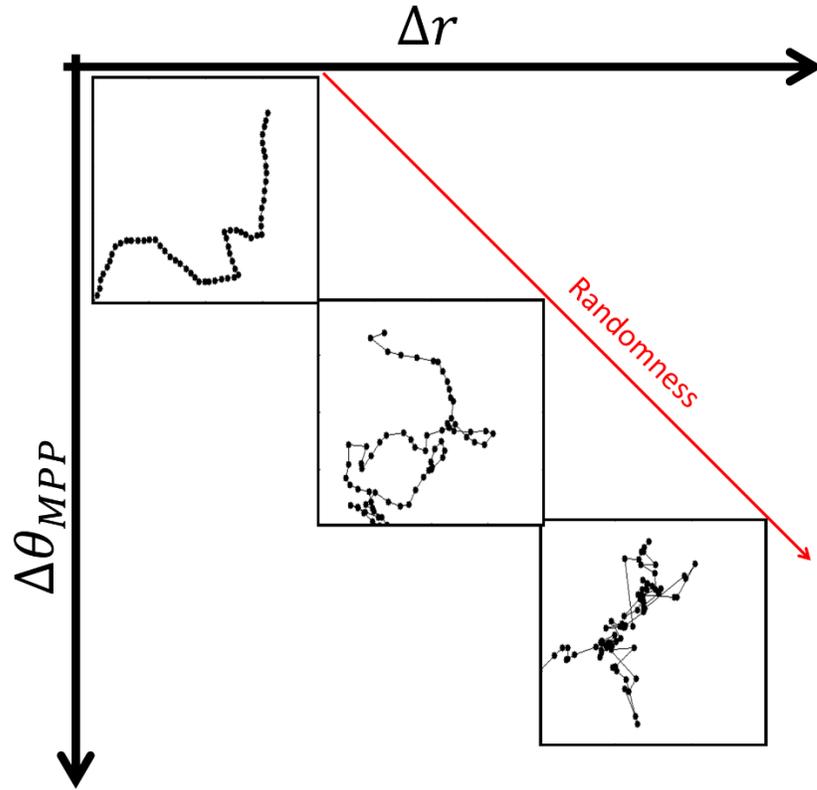


Figure 4.5: The randomness in CRW. The migration distance  $\Delta r$  is usually modeled by a weibull distribution and  $\Delta\theta_{MPP}$  is modeled by a wrapped cauchy distribution. Both distributions provide parameters to control the means and shapes. Distributions with higher probability to large  $\Delta r$  and  $\Delta\theta_{MPP}$  lead to more randomly distributed point sequences, while higher probability for small  $\Delta r$  and  $\Delta\theta_{MPP}$  lead to sequences consistently migrating in the same direction.

#### 4.2.4 The Modeling Algorithms

The bifactor lobe migration mechanism includes two parts: the MPP model (Fig. 4.6 a) and the lobe orientation model (Fig. 4.6 b), both of which are described in this section.

##### Algorithm 4.1 MPP Movement

This algorithm includes two steps. In the first step, the steepest slope direction is calculated at the sampled MPP migration distance from the previous MPP. In the second step, the migrating direction reflecting the input lobe sequence is calculated with CRW.

The two orientations are averaged with a combination weight, which represents the modeler's decision.

Step 1: Sample  $\Delta r$  from the lobe migration distance ECDF of the input lobe migration sequence;

Step 2: Find points  $p_{k,t}$  such that distances  $d_{k,t}$  from the current MPP  $MPP_t$  to all  $p_{k,t}$  satisfying  $d_{k,t} = \Delta r$  ;

Step 3: Find the point  $p_{steep,t}$  within  $p_{k,t}$ , such that the elevation change from the previous MPP to  $p_{steep,t}$  is the steepest;

Step 4: Find the steepest descending direction  $\theta_{steepest,t}$  by the unit vector from  $MPP_t$  to  $p_{steep,t}$ ;

Step 5: Sample the migration orientation shifting angle  $\Delta\theta_{MPP,t}$  from the ECDF;

Step 6: Calculate the new orientation reflecting the input lobe pattern by  $\theta_{MPP,t} = \theta_{MPP,t-1} + \Delta\theta_{MPP,t}$ ;

Step 7: Calculate the new mean orientation for the wrapped cauchy distribution  $E[\theta]_{MPP,t} = b * \theta_{MPP,t} + (1 - b) * \theta_{steepest,t}$ ;

Step 8: Sample the real migration direction with the wrapped cauchy distribution and  $E[\theta]_{MPP,t}$

#### **Algorithm 4.2: Lobe Orientation**

Step 1: Sample the lobe shifting angle  $\Delta\theta_{L,t}$  from the ECDF;

Step 2: Calculate two possible orientations with  $\theta_{L1,t} = \theta_{Basin} + \Delta\theta_{L,t}$  and  $\theta_{L2,t+1} = \theta_{Basin} - \Delta\theta_{L,t+1}$ ;

Step 3: Choose the direction  $\theta_{L,t}$ , such that the elevation changes along the direction are steeper;

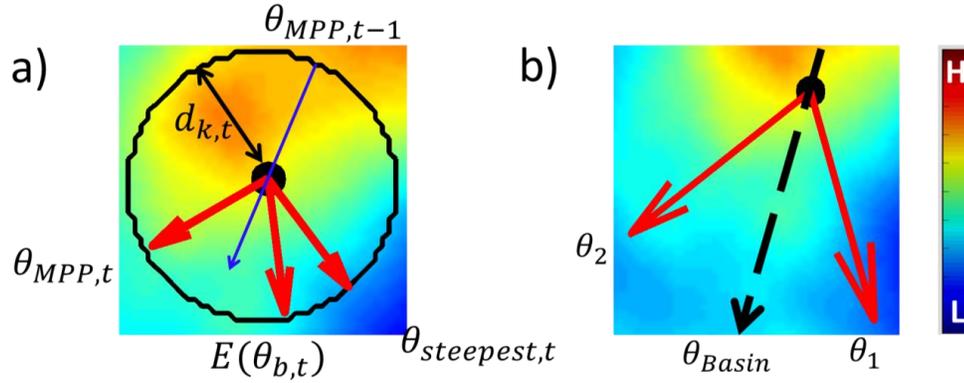


Figure 4.6: The implementation of the combined lobe migration mechanism: a) the bifactor model combining MPP migration information from the input lobe sequence and the conceptual compensational stacking rule in response to the intermediate topography in the simulation, refers to Algorithm 4.1. b) determining the lobe orientation, refers to Algorithm 4.2.

#### 4.2.5 Example

A run was performed on the new lobate model with the CRW-based lobe migration mechanism. Primary parameter values of the simulation are documented in Table 4.1. The purpose of this simulation was to demonstrate the capability of generating a realistic lobe migration sequence and more complex stratigraphy than the model with the conventional conceptual lobe deposition process demonstrated in Chapter 2. For the purpose of simplification, the lobe size was set to be constant. The basin source was stochastically sampled within a small segment located in the middle of an edge, and the overall basin flow orientation was perpendicular to the edge of basin source at the midpoint (Fig. 4.7). The initial topography was a plain surface. The input lobe migration sequence is demonstrated in Fig. 4.1, from which the ECDFs of the MPP migration distance, MPP migration orientation-shifting angle, and lobe orientation-shifting angle were extracted (Fig. 4.8). The simulation was designed to imitate the experiment, so no normalization was performed on the input lobe stacking pattern. The

directional persistence strength control factor for the wrapped cauchy distribution was set to be  $k = 0.6$ , implying that the lobe distribution process was moderated between being completely random and completely directionally persistent. The combination factor  $b$  was set to be 0.5, implying that precise statistics from the input lobe sequence and the conceptual compensational stacking rule accounted for half of the lobe migration mechanisms. 100 lobes were simulated in every realization. The intermediate topographies of five steps from one realization are demonstrated in Fig. 4.9. The depositional process started from the top edge of the model grid, and the first lobe complex was formed at the top-right portion of the grid from Lobe 1 until Lobe 20. The process started to laterally shift from Lobe 20 and filled the up-left region of the grid with a second lobe complex at around Lobe 40. After some random filling around Lobe 50, progradation was observed at Lobe 60 toward the distal empty region of the modeling grid. In the following steps up to Lobe 100, a third lobe complex was in the process of formation. The lobe depositional sequence is visualized at three vertical sections from the proximal to distal in Fig. 4.10. Compared to the lobe sequence section generated by the conceptual depositional rules demonstrated in Fig. 2.26, more natural and realistic spatial-temporal clusters of lobe were observed. Finally, the ECDFs of the statistics from the input lobe migration sequence were compared to that from the realization (Fig. 4.11). Since the MMP migration orientation was a weighted average of two factors, this ECDF is not reproduced in the realization. However, reproductions of the migration distance and lobe-shifting angle were satisfying.

Simulated Object	Lobe only
Modeling Grid	200 cells by 200 cells
Cell Size	1
Lobe Length	$L = 120$ (constant)
Basin Source	$x = 0; 94 < y < 106$
Basin Flow Orientation	Perpendicular to the edge of basin source
Initial Topography	Plain
$k$ (wrapped Cauchy distribution)	$k = 0.6$
Combination Factor	$b = 0.5$
Number of Lobes	$N = 100$

Table 4.1: Values of primary parameters in the simulation.

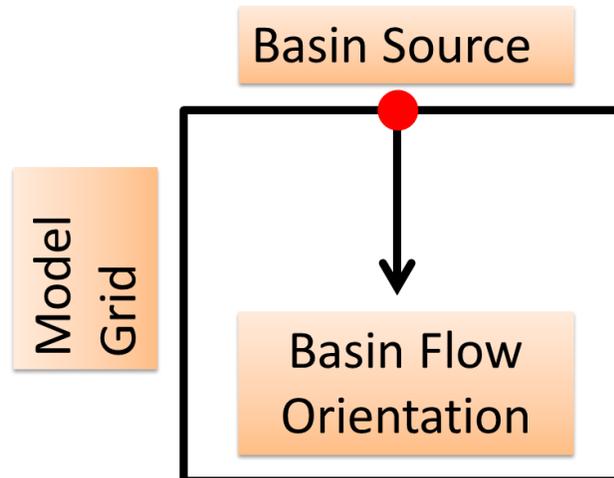


Figure 4.7: Model grid setting for the test simulation. The basin source sample within a small section around the middle point of an edge (refer to Table 4.1) and the basin flow orientation were set to be perpendicular to the edge of the basin source toward the modeling grid.

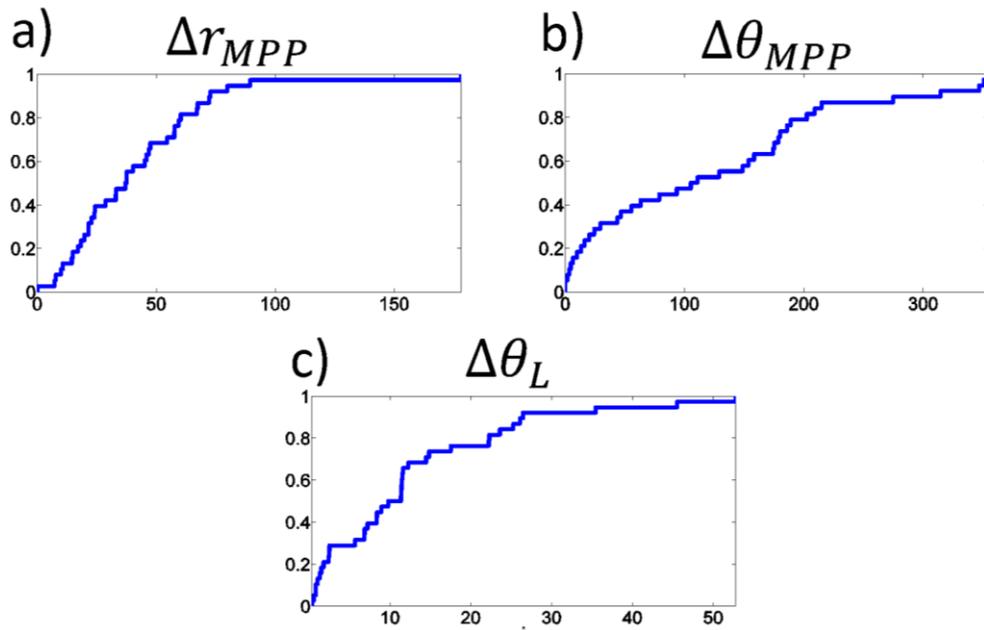


Figure 4.8: ECDFs from the inputs lobe migration sequence.

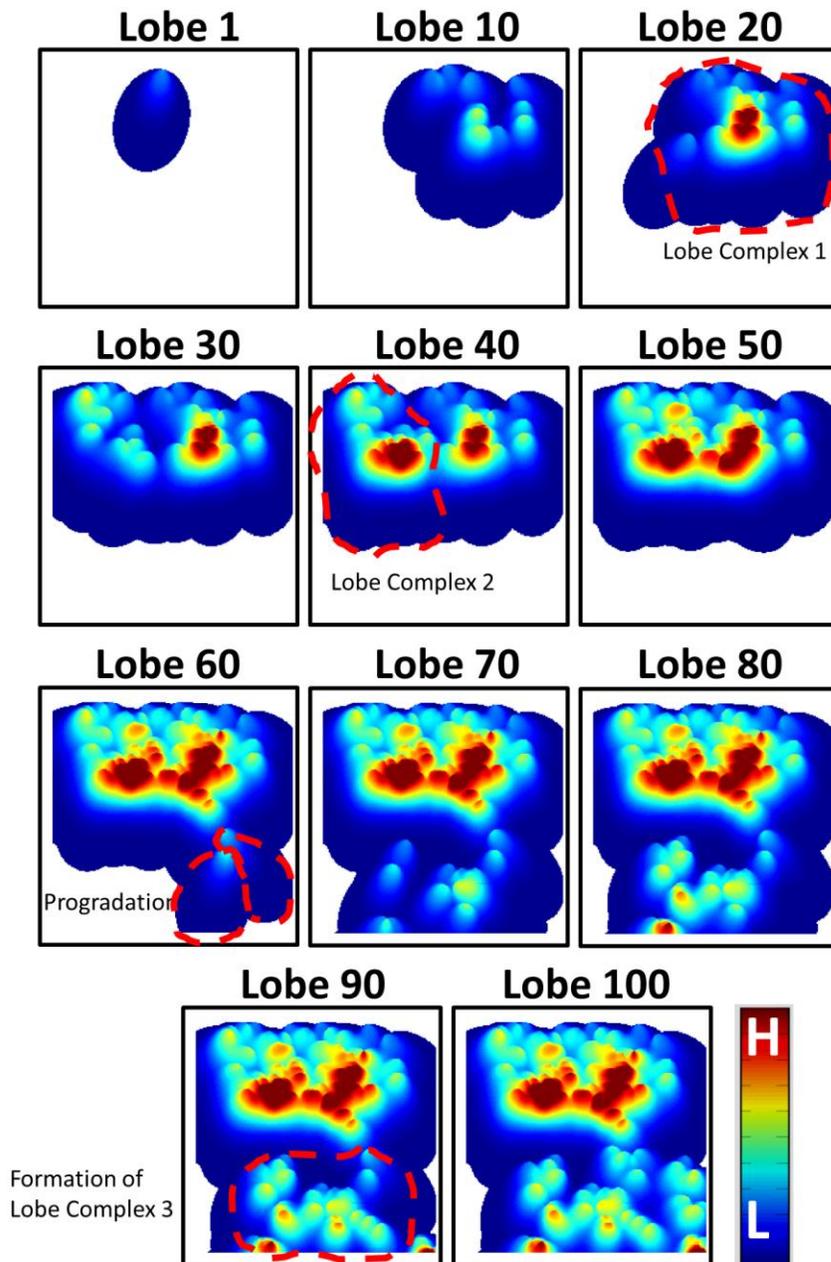


Figure 4.9: Samples of intermediate topographies of a realization with 100 lobes. Several lobe complexes following the compensational stacking rule are observed in the simulation.

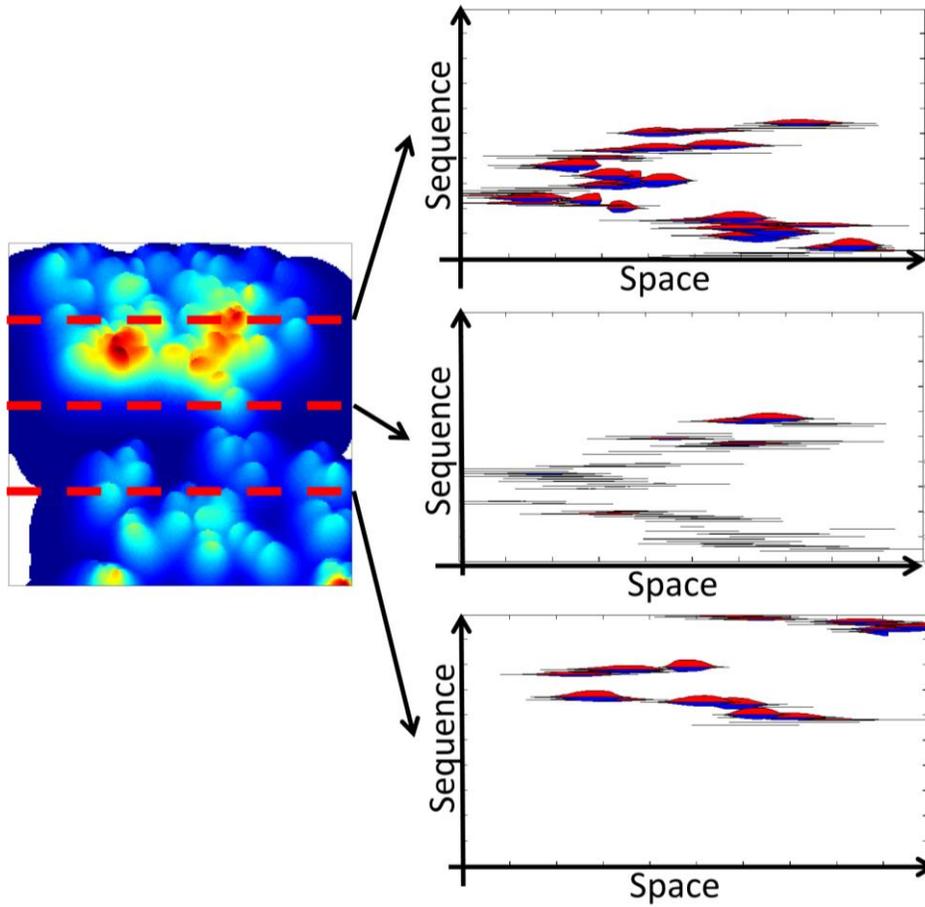


Figure 4.10: One realization of 100 lobes and the proximal, medial, and distal depositional sequence plot. Lateral shifting and spatial-temporal clustering are observed.

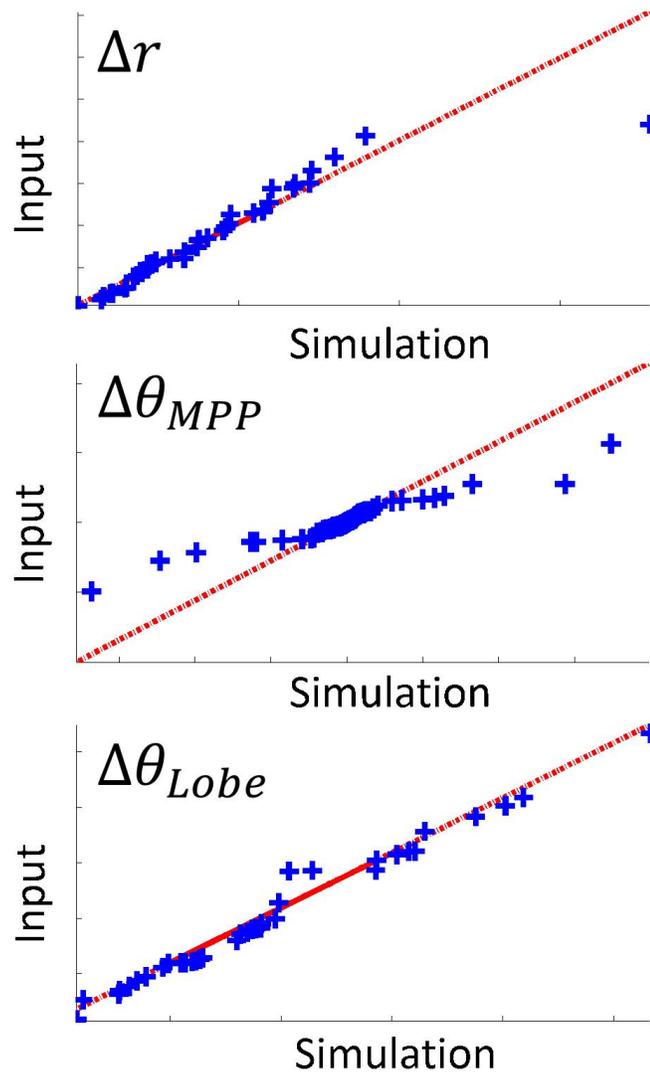


Figure 4.11: The reproduction of input ECDFs of  $\Delta r$ ,  $\Delta\theta_{MPP}$ , and  $\Delta\theta_{Lobe}$ . Relative MPP migrating orientation  $\Delta\theta_{MPP}$  is not reproduced due to the combination with in-situ compensation stacking rules.

### 4.3 Hierarchical Similarity

The ECDFs of the MPP migration distance, MPP migration orientation-shifting angle, and lobe orientation-shifting angle only characterize the sequential pairwise information of the migration process. Nevertheless, an input lobe sequence also

provides overall information on the lobe stacking pattern, such as lobe hierarchies described by the dendrogram or reachability plots. In Chapter 3, the dendrogram of a stacking pattern was calculated with the sequential cumulated weighted mean distance over all four parameters, measuring pairwise lobe proximity. Since the weights are simply chosen by the modeler, reproducing the dendrogram of a stacking pattern is essentially the problem of reproducing dendrograms for each of the parameters: MPP distance, lobe angle, polygon distance, and shape dissimilarity. The optimal lobe migration mechanism should be capable of constraining all four dendrograms, which is complicated. The reason is that characterizing dendrograms requires obtaining all the simulated lobes, while the lobe migration is a forward process in which the final pattern information is unavailable before a simulation is completed. Thus, accounting for the final stacking pattern hierarchies is a computationally intensive and inefficient inversion problem in most situations, which is unfavorable in the geological modeling stage of a reservoir development project. This section demonstrates that the input sequential statistics in the CRW-based mechanism are an implicit control of overall lobe hierarchies in the forward simulation process (Section 4.3.2). Thus, all realizations with the new lobate model are hierarchically similar to the input lobe migration sequence (Section 4.3.3), which is a new feature not available in the model with conceptual depositional rules presented in Chapter 2. The prerequisite of comparing hierarchies is a dissimilarity measure quantifying the proximity with reasonable geological interpretations, introduced in Section 4.3.1.

#### **4.3.1 Quantitative Measure of Hierarchical Similarity**

For the purpose of studying the hierarchical similarity of the lobe MPP, the plane and perspective view of the input lobe migration sequence and of one simulated realization were plotted. The input lobe sequence included 38 lobes (Fig. 4.1), and the realization involved lobes from Step 1 to Step 40 in the example, demonstrated in Section 4.2.5. An observable difference between the two sequences was that the shifting angle of the MPP migration direction in the input sequence (Fig. 4.12) was more

random, although that in the simulated realization (Fig. 4.13) was more directionally persistent, caused by the directional persistence controlling parameter  $k$  and the combination factor  $b$ . The lobe hierarchical similarity based on the sequential migration distance was controlled in the model, given the difference in the MPP migration direction caused by  $k$  and  $b$ . In this dissertation, hierarchies of MPP migration distances are quantified by dendrograms (Fig. 4.14) generated by agglomerative hierarchical clustering with sequential cumulated distances, so the problem of measuring the hierarchical similarity is converted to defining a dissimilarity measure function to quantify the proximity between dendrograms of the patterns.

A dendrogram is formally defined as a rooted, terminally-labeled, and weighted tree in which all terminal nodes are equally distant from the root (Lapointe and Legendre 1995). A dendrogram consists of three components: topology, label, and weights (Fig. 4.15). Topology is the bifurcation structure between leafs only, the label links each leaf to a lobe in addition to the topology, and the weight considers the height of the branches, accounting for the distance between clusters. Comparison of dendrograms includes two parts, a descriptor of the dendrogram and a function measuring the difference of the descriptor. The most popular dendrogram descriptor is the cophenetic distance (Sokal and Rohlf 1962), defined as the interleaf distance within the plot. In a dendrogram, the cophenetic distance between two leaves is the height at which two leaves merge into a branch (Fig. 4.16). The cophenetic distance only considers the label and weights, not topology. The other popular descriptor is the topological distance (Phipps 1971), which considers topological properties only, defined as the number of junctions counted along the route from one leaf to another. Three variations of cophenetic distance and topological distance are proposed (Podani and Dickinson 1984), all of which are considered together as a multivariate descriptor. Functions measuring the difference between descriptors in the literature, unfortunately, are all designed to compare two or more dendrograms generated from the same datasets of  $n$  observations. For example, if two dendrograms are obtained from the same dataset with different hierarchical clustering algorithms, two cophenetic distance values can be calculated for every

observation. Pearson's correlation coefficient between the two sets of cophenetic distances of size  $n$  is estimated to quantify the similarity between the two dendrograms. Known as the cophenetic correlation coefficient (Sokal and Rohlf 1962), the technique is prevalent and is applicable to other descriptors.

Since the compared dendrograms in this study were generated from lobe sequences normally composed by a different number of lobes, the difference function was defined as the ECDF of the cophenetic distance of a dendrogram. There were two advantages of the new descriptor. First, the ECDF of the cophenetic distance characterized the intensity of the clustering effect of the stepwise lobe migration; thus, the comparison was based on the statistical features of geologically interpretable clusters of a migration sequence, e.g. progradation, retrogradation, and aggradation. Second, the ECDFs did not require an equal number of observations in the two dendrograms, such that hierarchies of sequences with different numbers of lobes can be compared. Once the dendrograms were converted to the ECDFs of their cophenetic distances, the L1-norm difference was used to measure the dissimilarity between two distribution functions. The ECDFs of the cophenetic distances of both input lobe sequences and realization lobe sequences are demonstrated in Fig. 4.17. They are visually similar

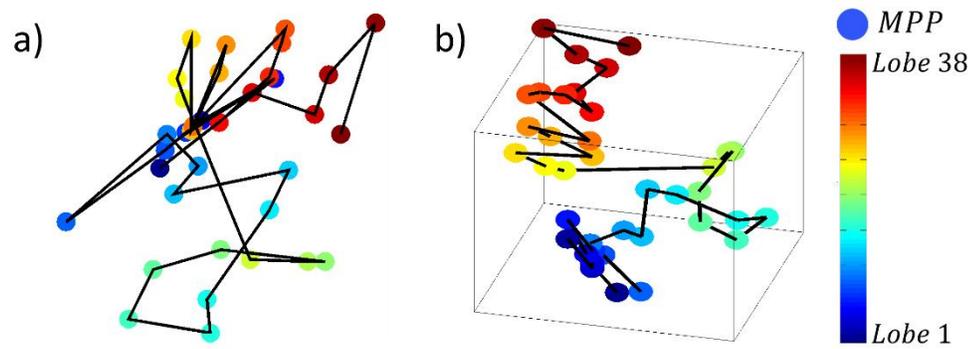


Figure 4.12: MPP migration sequence of the input stacking pattern: a) the plain view; b) the perspective view.

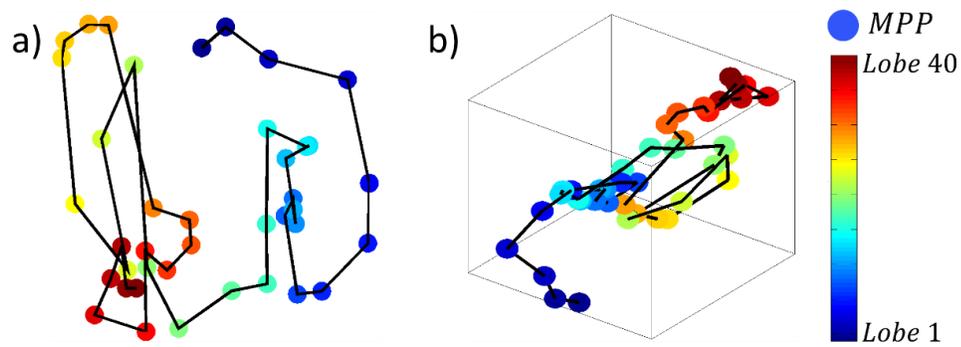


Figure 4.13: MPP migration sequence of the realization: a) the plain view; b) the perspective view.

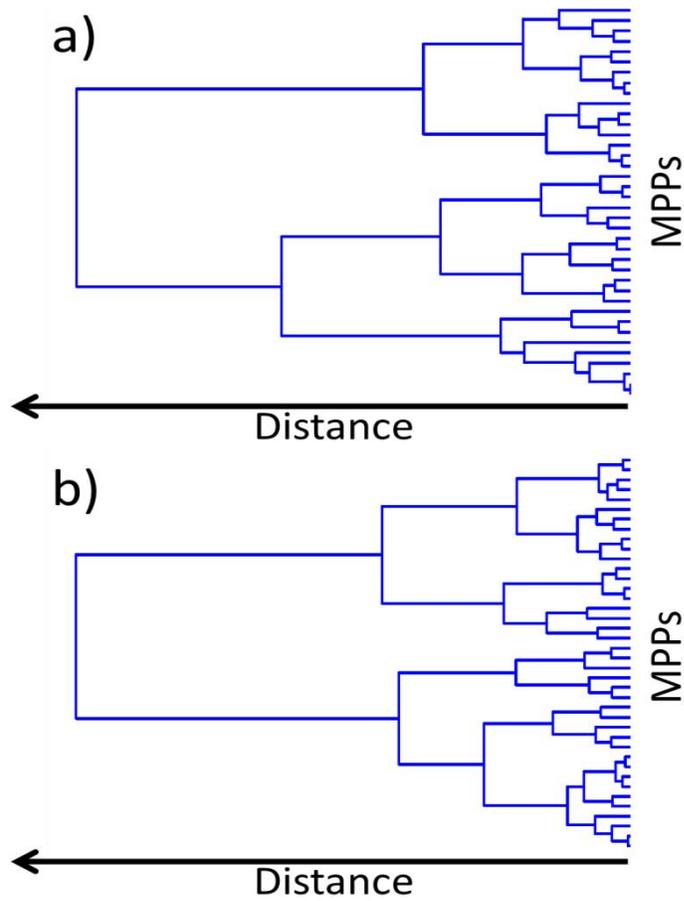


Figure 4.14: The sequential hierarchies of input and realization. The sequential hierarchy is quantified by the dendrogram with the sequential cumulated distance. The hierarchical similarity is defined by the similarity between two dendrograms. a) The input stacking pattern; b) the realization.

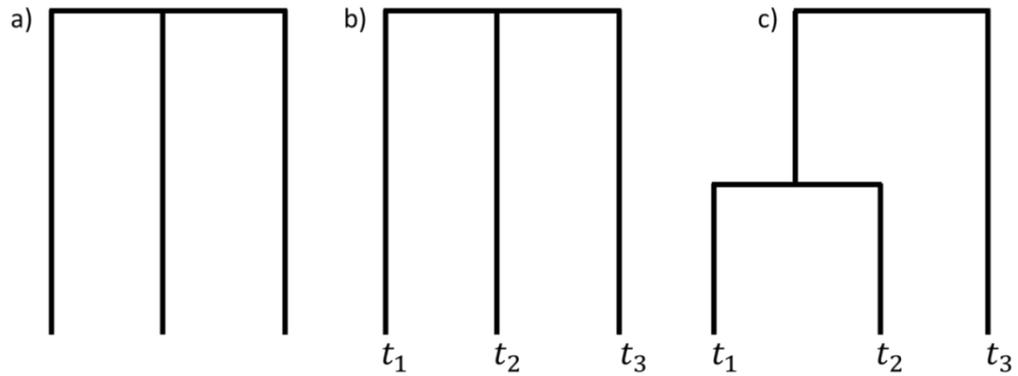


Figure 4.15 A dendrogram is a weighted binary tree with three components: topology, label, and weights. a) topology: the bifurcation structure between leaves of the tree; b) label: links each leaf to a branch; c) weights: the height at which two branches merge, implying the distance between two clusters represented by the two branches.

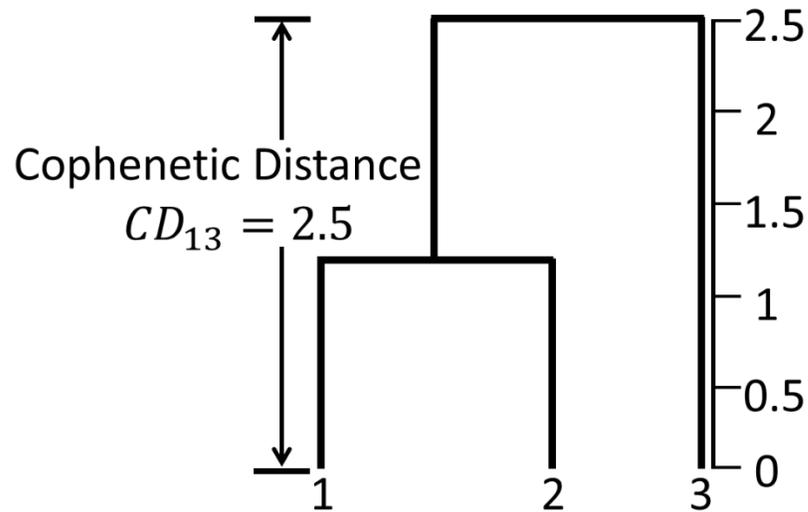


Figure 4.16: Cophenetic distance: distance at which two leaves merge. In the example, the cophenetic distance from Leaf 1 to Leaf 3 is 2.5.

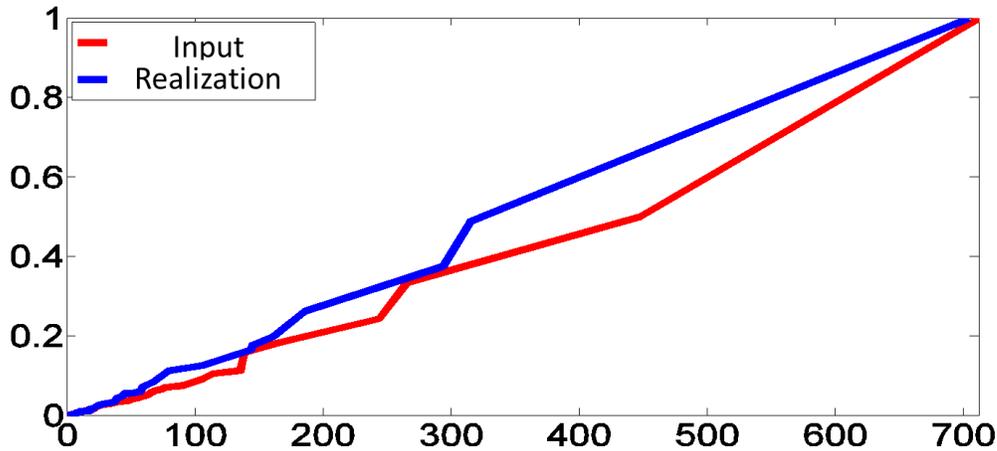


Figure 4.17: ECDFs of the cophenetic distance from the input lobe sequence and from the simulated realization in Fig. 4.14.

### 4.3.2 Implicit Hierarchical Control in the CRW-based Mechanism

The similarity between the ECDFs of the cophenetic distances from the input and from the realization in Fig. 4.17 is not simply a coincidence. In fact, the ECDFs of the cophenetic distance of realizations generated with sequential cumulated distances were implicitly controlled by the ECDF of the stepwise migration distance. During the simulation with the CRW-based mechanism, the ECDF of the stepwise migration was reproduced in model realizations. Hence, the ECDF of the cophenetic distance from the simulated dendrogram generated with the sequential cumulated distance were also constrained. From the perspective of agglomerative hierarchical clustering, the cophenetic distance, or the interbranch distance, within a dendrogram plot is determined by the linkage function chosen to estimate the intercluster distance. Therefore, it is worthwhile to study the relationship between different linkage functions and the resulting ECDF of the cophenetic distance while the sequential cumulated distance is applied in agglomerative hierarchical clustering. The linkage function, as briefly introduced in Chapter 3, is the core of hierarchical clustering algorithms. It is always easier to introduce an algorithm with a synthetic example of a four-point sequence (Fig. 4.18). Hierarchical clustering with a sequential cumulated distance (Eq. (3.6)) starts

from grouping the closest pair of points in the given sequence. In Step 3, two groups,  $\{1,2\}$  and  $\{3,4\}$ , are already formed since the sequential cumulated distance between 1,2 and 3,4 is the shortest. Before Step 4, the distance between the big groups,  $\{1,2\}$  and  $\{3,4\}$ , should be calculated using the linkage function, which is consequently the cophenetic distance between leaf 1 (or 2) and 3 (or 4).

A formal definition of the linkage function is a measure of dissimilarity between two clusters (groups of observations) (Hastie, Tibshirani, and Friedman 2009). Various definitions of linkage functions are proposed in the literature, from which the three most prevalent types, 1) average linkage, 2) complete linkage, and 3) single linkage, are briefly studied here. Let  $G$  and  $H$  be two groups between which the dissimilarity  $d(G,H)$  will be measured. Let  $d_{ii'}$  represents the pairwise dissimilarity between elements of  $G$  and  $H$ , where one member of the pair  $i$  is from Group  $G$  and the other member  $i'$  is from Group  $H$ . The average linkage (Eq. (4.7)) defines the dissimilarity of  $G$  and  $H$  as the average of the pairwise dissimilarity of each member. The complete linkage (Eq. (4.8)) calculates the intercluster dissimilarity to be that of the furthest pair of elements. On the contrary, the single linkage (Eq. (4.9)) uses the dissimilarity of the nearest pair of elements as the intercluster dissimilarity.  $d_{ii'}$  in Eq. (4.7) – Eq. (4.8) is calculated with any non-sequential distance. Figure 4.19 shows dendrograms with average, complete, and single linkages in the experimental and realization sequences in Fig. 4.12 – 4.13. Comparing each column of Fig. 4.19, the observation is that the application of different linkage functions significantly changes the structure of resulting dendrograms. However, no significant hierarchical dissimilarity is observed between dendrograms of the input and realization in each row. In general studies of the linkage function in which the sequential cumulated distance is not used, imprecise relationships (Eq. (4.10) – (4.11)) exist between the original distance (sequential cumulated distance in this dissertation) and the cophenetic distances generated by different linkage functions. A single linkage function generates dendrograms with the shortest intercluster distance, which is even less or equal to the original distance. The complete

linkage function produces the largest cophenetic distance that may be greater than the original distance. The average linkage function ranges between the single and the complete linkage functions. However, no clear relationship between cophenetic distance generated with the average linkage function and the original distance is found in the related algorithmic literature.

$$d_{Average}(G, H) = \frac{1}{N_G N_H} \sum_{i \in G} \sum_{i' \in H} d_{ii'} \quad (4.7)$$

$$d_{Complete}(G, H) = \max_{i \in G, i' \in H} d_{ii'} \quad (4.8)$$

$$d_{Single}(G, H) = \min_{i \in G, i' \in H} d_{ii'} \quad (4.9)$$

$$d_{Single}(G, H) \leq d_{Average}(G, H) \leq d_{Complete}(G, H) \quad (4.10)$$

$$d_{Single}(G, H) \leq d_{Original}(G, H) \leq d_{Complete}(G, H) \quad (4.11)$$

While utilizing the sequential cumulated distance (Eq. (3.6)) in the hierarchical clustering, equations of different linkage functions are simplified. Assuming the temporal indices is  $m_G \leq i \leq n_G$  within Group G,  $m_H \leq i \leq n_H$  within Group H, and  $n_G \leq m_H$ , the relationship between the ECDFs of the cophenetic distances from the results of single and complete linkage functions are defined by Eq. (4.12) – (4.13), where  $d_{m_G, n_H}$  is computed with Eq. (3.6). Since the sequential cumulated distance converges the problem into a one-dimensional geological space, the single linkage function defines cophenetic distance from the last observation in the earlier group to the first observation in the latter group, and the complete linkage function defines cophenetic distance from the first observation in the earlier group to the last observation in the latter group. The relationship in Eq. (4.10) – Eq. (4.11) is maintained in this condition. The distance in complete linkage requires calculating the maximum of the pairwise distance, which is the maximum of the sequential cumulated summations. Thus, the complete linkage produces cophenetic distance measurements much greater than the original distance. The average linkage produces a cophenetic distance that is between the maximum and the minimum. As a demonstration, the L1-norm distance is applied

to the ECDFs of the cophenetic distance to compare dendrograms of each row in Fig. 4.19, in which the distance is 3758.4 for average linkage, 15666 for complete linkage, and 899.3 for single linkage, consistent with the analysis above.

Eq. (4.10) – Eq. (4.11) also define the imprecise relationship between the original stepwise MPP migration distance and the cophenetic distance from the simulated dendrogram. Thus, implicit control on the finalized lobe hierarchies of realizations through the forward simulation process of the CRW-based MPP migration mechanism is proved. In terms of reflecting the relative hierarchical similarity between two lobe sequences, all linkage functions are the same as long as the same function is applied to an ensemble to be compared. However, the single linkage is recommended since it produces a cophenetic distance that is the closest to the original distance.

$$d_{Single}(G, H) = d_{n_G, m_H} \quad (4.12)$$

$$d_{Complete}(G, H) = d_{m_G, n_H} \quad (4.13)$$

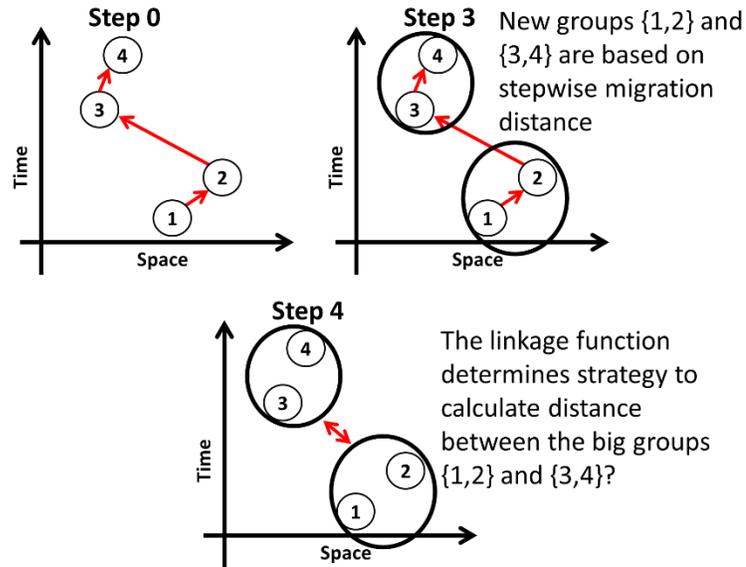


Figure 4.18: The linkage function, a strategy of calculating the intercluster distance in the agglomerative hierarchical clustering, determines the ECDF of the cophenetic distance. An imprecise relationship exists between the ECDF of the stepwise migration distance and the ECDF of the cophenetic distance from the dendrogram.

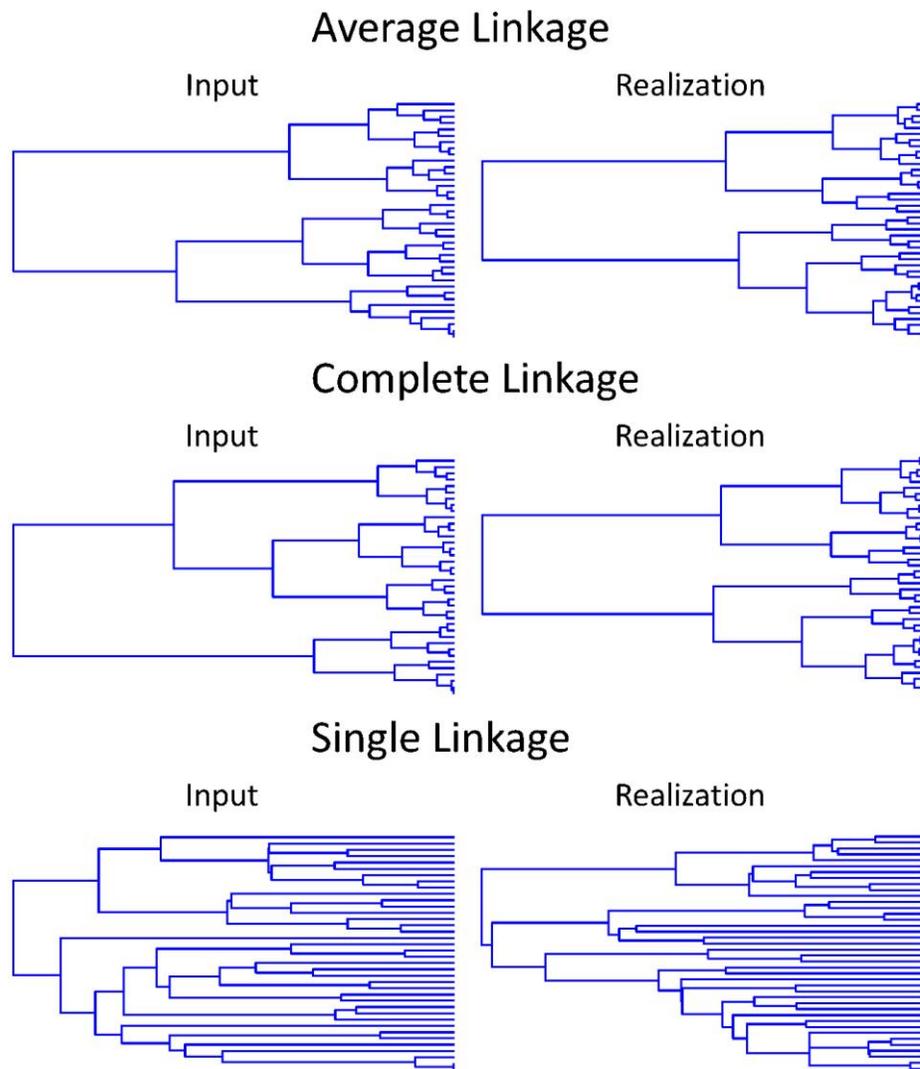


Figure 4.19: The dendrograms generated with different linkage functions using the input lobe sequence in Fig. 4.12 and the realization in Fig. 4.13. The visual comparison can easily identify the difference between the dendrograms with different linkage function on the same set of data. However, no significant difference is observed between dendrograms of the input and realizations with the same linkage function. The L1-norm differences between the ECDFs of the cophenetic distances of the input and realization dendrograms are calculated for each linkage function: 3758.4 for complete linkage, 1566.6 for average linkage, and 899.3 for single linkage, which is consistent with the theoretic analysis.

### 4.3.3 Application of Hierarchical Similarity Control

The hierarchical similarity is a meaningful advantage for extending the application of surface-based models. Updating of conceptual geological models with additional data is always an important step at the appraisal stage of a hydrocarbon reservoir. In a recently proposed strategy for rapidly updating conceptual models (Caers 2013), training images, which is the realization of a process-based model are used in multiple-point statistics. Training images are quantified conceptual models, and the updating is implemented by adding new training images or updating the probabilities of each training image. It is also valuable to test the application of surface-based models in this strategy for cases where training images are unavailable or solving a process-based model is not allowed by the limited time of a project. Analogous to training images in multiple-point statistics, an input lobe migration sequence is a quantification of prior knowledge about the dynamic depositional process. We will further demonstrate that different input sequences can effectively distinguish MPP hierarchies in all model realizations, such that surface-based models can be directly used in the rapid updating workflow as an alternative to multiple point statistics.

An algorithm comparison method based on an analysis of distance (Tan, Tahmasebi, and Caers 2013) was applied to visualize differences between hierarchies of realizations generated by models with conventional conceptual depositional rules and with the CRW-based lobe migration mechanism. First, 210 realizations with different lobe shapes were generated with the model built in Chapter 2, including 70 respective realizations for narrow, medium, and fat lobes. Then, two different input lobe sequences were chosen and 210 realizations of different lobe shapes were generated with each of the sequences. One input sequence was the 38 lobe case from Exp. B, while the other was a 296 lobe case from Exp. A. The lobe size in all realizations was constant, and the model setting is demonstrated in Fig. 4.7. With the hierarchical similarity measure defined in Section 4.3.1 – 4.3.2, MPP hierarchies of all realizations are plotted with the first two components of multiple dimensional scaling (Scheidt and Caers 2009) in Fig.

4.20. Dots in the plot represent all model realizations, and crossings represent two input lobe sequences. Realizations with the conventional conceptual deposition mechanism are very difficult to update because the prior knowledge is quantified by tens of empirical coefficients, while realizations with the new CRW mechanism can be effectively distinguished by simply changing the input lobe sequences. Although the control of MPP hierarchy is imprecise, the method demonstrates a new methodology of improving surface-based models.

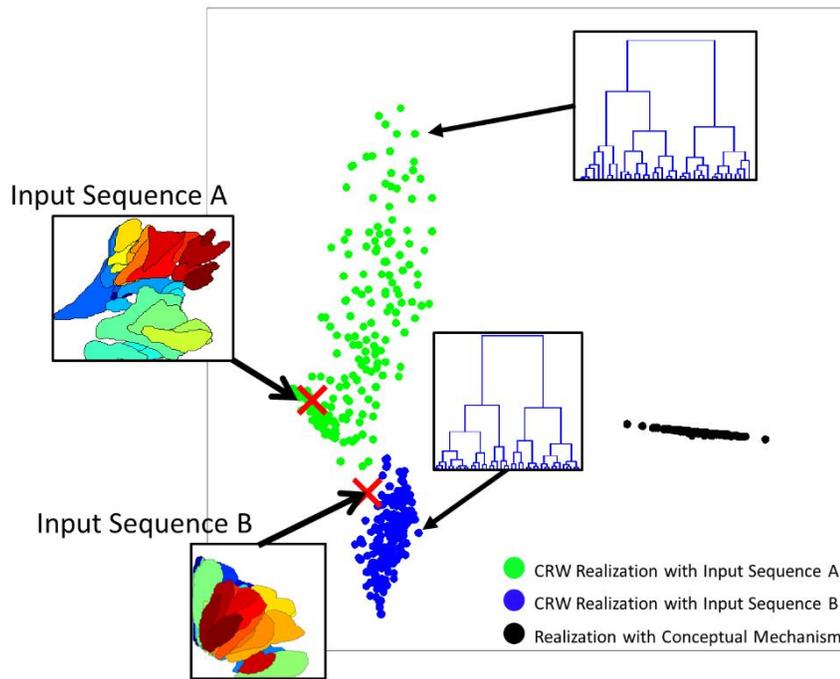


Figure 4.20: First two component multidimensional scaling plot of hierarchies of realizations, demonstrating the effect of an input lobe sequence on constraining the MPP hierarchies of model realizations, which is difficult with the model built in Chapter 2.

## 4.4 Chapter Summary

The high p-value stacking patterns were treated as the quantified prior information to control the spatial distribution of lobes. Instead of characterizing information from the final stacking patterns, a lobe stacking pattern is treated as a lobe migration sequence. Assuming the lobe migration process is homogeneous and stationary at the spatial and temporal scale of the lobe sequence, the migration of the lobe MPP is characterized by the ECDF of the stepwise MPP migration distance and the MPP migration orientation-shifting angle. With the ECDFs as inputs, the migration sequence as CRW was described. The reason for choosing a technique as simple as CRW is that it was desirable to control the number of parameters in the implementation, such that the difficulty of conditioning a surface-based model to wells and thickness data was reduced. Since only two dimensional horizontal movement data were available for the stacking pattern, and both the migration movement of the input lobe sequence and the conceptual rule of compensational stacking should be accounted for, a bifactor scheme was designed such that the final movement of the MPP was determined by the weighted average of the input lobe sequence and the compensational stacking rule. Once the location of a MPP was calculated, the lobe orientation was obtained by a lobe shifting angle from the overall basin flow orientation. The lobe shifting angle also followed its statistics from the input lobe sequence. The first advantage of this implementation is the simplicity. Essentially, aside from the ECDFs from inputs, only two empirical coefficients were employed to produce realistic movement of MPPs, including clustering, progradation, retrogradation, compensational stacking etc., which is significantly reduced compared to the model built in Chapter 2. Second, CRW is an implicit control of the hierarchies of realizations. To prove this, a quantitative dissimilarity measure based on the cophenetic distance was proposed to quantify the hierarchical similarity between the dendrograms of two lobe sequences. There was no attempt to account for other types of spatial pattern information of the input lobe sequence, such as polygonal proximity and shape proximity.

The problem of modeling lobe migration is a spatial-temporal random process, which should be ideally modeled as a function between the intermediate topography and the spatial attributes of the next lobe, including location, orientation, polygonal pairwise proximity and shape, in a three dimensional space. Essentially, the bifactor modeling scheme is but a temporary solution because the lack of intermediate topographic information in lobe sequences. However, the idea of a CRW-based implementation is still a practical method to three-dimensional data.

# Chapter 5

## CONCLUSIONS AND FUTURE WORK

### 5.1 Conclusions

Solutions to three problems were developed in this dissertation. First, the lack of intermediate topographic information is a challenge in the surface-based modeling approach. However, the data is difficult to obtain in real-scale analogues or active real depositional systems. Geomorphic experiments were proposed to be a knowledge database for static reservoir modeling (Chapter 2), since an experiment can produce more complex depositional features than process-based models and can be recorded comprehensively. The solution for this problem includes a method of extracting intermediate geometric information from experimental records and a surface-based model of distributary channel-lobe systems with the intermediate information as inputs. Second, since the most similar experiment provides the most referable information, a practical appraisal project aimed at specifying one experiment (among multiple options) the most similar to a real depositional basin under appraisal. A statistical solution was devised based on the spatial similarity between lobe stacking patterns from the experiment and from the real system (Chapter 3). The solution consists of a set of

algorithms to quantify geological concepts, such as lobe hierarchies, and algorithms to analyze the concepts, which are also useful in comparative studies on lobate environments. The last problem was how to honor the information of an experimental sequence of lobes in the simulation of surface-based models. A new modeling strategy was described in which all model realizations were hierarchically similar to the identified lobe sequence. A quantitative measure of the hierarchical similarity between lobe sequences was introduced (Chapter 4).

The integration of geomorphic experimental data in surface-based models in Chapter 2 was initially proposed to improve erosion on shale drapes in deepwater lobate systems. In terms of conceptual models in deepwater lobate environments, thin subseismic scale shale drapes appear between sandy lobate bodies at different scales of the stratigraphy and behave as flow barriers in a reservoir. Erosion on shale drapes removes shale layers and causes directly contacted large sand bodies, thus impacting the performance of a reservoir. Essentially, erosion caused by subsequent lobes requires understanding the intermediate topography of the depositional process; however, the intermediate topographic information is removed in real-scale outcrops and measuring intermediate topography in active depositional systems is infeasible. Thus, erosion is described with imprecise conceptual rules in current surface-based models. The conceptual rules introduce large numbers of empirical coefficients, thus increasing model complexity and workloads in uncertainty estimations. The geomorphic experiment used in Chapter 2 includes intermediate topographic information measured at fine temporal scales. Erosion and deposition are identified, extracted, and visualized from experimental intermediate topographies, through which statistics of erosion-deposition are interpreted. Since the visualized intermediate information implies correlated erosion-deposition events, the lobate model was designed with negative-positive surfaces. The negative-positive surfaces represent erosion and deposition, respectively. The correlation between erosion and deposition geometries is characterized by a dimensionless ratio between the maximum erosion depth and the

maximum deposition thickness, such that information from small experiments is scalable to large real systems.

To identify an experiment that is the most similar to a specific real depositional system (Chapter 3), physical and conventional stratigraphic methods are not applicable in experiments. The reason is that physical quantities of ancient sediment flows are not measurable, and petrophysical properties used as the basis of interpretation in conventional stratigraphy are simplified in geomorphic experiments. The proposed solution attempts to find a statistical similarity between lobe stacking patterns from different systems. A lobe stacking pattern is characterized by the ECDFs of the lobe proximity parameters, and the similarity between two stacking patterns is estimated by a p-value obtained from a two-sample hypothesis test on whether two empirical distribution functions are from the same underlying population distribution. If two stacking patterns are from the same underlying distribution, a reasonable assumption is that the two processes are mutual references and information from the experiment is applicable in modeling the real system. Lobes can be identified at different scales within a lobate system. Thus, the solution provides a method to identify the proper scale of interpretation in the experiment, at which the lobe pattern is more similar to a specific real system than at other scales. Hierarchical clustering with constraints is applied on parameterized lobe patterns so clustering is forced to follow the order of the lobe sequence. Interpretation of large scale lobes from small scale lobes is automatically performed with the algorithm. The significance of this solution is not limited to linking experiments to real-scale systems, but also to quantitative definitions of lobe hierarchy and methods applicable to the quantitative comparison of different experiments.

Lobe stacking patterns identified with the aforementioned method are quantified scenarios of the depositional dynamic process. The concept of a geological dynamic process is proposed as an extension of depositional rules to define forward processes of distributing interpreted geobodies in response to the intermediate topography. Surface-based models with lobe sequence inputs are the proper way to describe depositional

processes. The integration of input lobe sequences in surface-based models includes characterizing and simulating geologically meaningful statistics of the input sequence with an algorithmically simple mechanism (Chapter 4). The input lobe sequence is characterized by the stepwise relative migration distance and the migration orientation-shifting angle based on the assumption that the geological dynamic process is stationary in an input lobe sequence. The cumulative distribution functions of the lobe migration distance and lobe migration orientation-shifting angle are used to control a CRW describing locations of lobe MPPs. Since only two-dimensional horizontal patterns are provided in the input lobe sequence while the compensational stacking rule has to be accounted for as a conceptual rule in the simulation, a bifactor scheme is devised to combine impacts from both the input lobe sequence and the conceptual compensational stacking rules. With a quantitative definition of the lobe hierarchical similarity, it is demonstrated that the CRW mechanism is an implicit control on the hierarchy of the *MPP*, although the control is imprecise.

## **5.2 Recommendations for Future Work**

Distinguished from other static modeling techniques, the surface-based method is the only means of explicitly modeling a forward geological dynamic process, instead of non-forward statistical methods and forward physical methods. Explicitly considering the forward geological dynamic process produces realistic realizations and empowers a surface-based model to be the only method capable to study uncertainties caused by a geological dynamic process, which is the essential determinant of spatial heterogeneity of petrofacies. However, a large number of empirical coefficients in conceptual depositional models limits applications of estimating uncertainty for geological processes. With sufficient data provided in geomorphic experiments, design and calibration of precise mathematical models for geological dynamic processes becomes possible. This work is dedicated to the fundamental work of the quantification of geological concepts, the characterization of geological dynamic processes, and the

simulation with constraints in geological processes. Suggestions for future work to broaden and improve this work is discussed below.

- **Quantitative Assessment Criteria for Surface-based Algorithms**

A most frequent question in assessing surface-based models is how to determine that one implementation of a surface-based model is better than another. Generally, the assessment of algorithms is based on the purpose of models. As indicated earlier, simple and fast algorithms are better for uncertainty estimations while more complex, but realistic, algorithms are more proper for testing hypotheses about depositional processes. However, quantitative estimations of the comparison results are expected for serious research. In this dissertation, geological concepts of interest were quantitatively defined and then compared with the reproduction of these concepts with inputs, such as the hierarchical similarity and lobe stacking pattern similarity. The comparison criterion of algorithms can definitely be quantified with the degree of reproduction of inputs in the realizations. For future work, alternative quantification of geological concepts may be studied and compared to the dendrogram proposed in this dissertation. For example, an alternative popular scale-dependent measure is Ripley's K function. In spatial-temporal processes, geological concepts of lobe migration are normally too complex to be characterized by a single descriptor. A comparable study on advantages and disadvantages of both techniques are valuable. The objective would be to understand differences and relationships between different measures of geology and to implement a surface-based lobate model constrained in all measures. Or, a new statistical measure that combined advantages of both techniques could be devised.

- **The Relationship between Reservoir Performances and Dynamic Geological Processes**

Reservoir performances are affected by the heterogeneity of petrofacies, which is further determined by the geological dynamic process sequentially distributing interpreted regular shape geobodies of petrofacies. Hence, a reasonable hypothesis is

that the geological dynamic process is sensitive to the heterogeneity of petrofacies. Thus, the direct uncertainty estimation of reservoir performances with parameters and coefficients of a geological dynamic model may be more effective than using the finalized models, such as training images. As indicated in Chapter 1, surface-based and process-based models have been applied to select the most possible training images. However, it is still valuable to study the relationship with the improved models, for example, studying whether similarities will also be observed in the spatial distribution of petrofacies given that hierarchies are statistically reproduced in realizations of the CRW-based mechanism.

- **Identification of Three-Dimensional Geology with Two-Dimensional Measurements**

One assumption of Chapter 3 is that the horizontal spatial-temporal distributions of real lobes are interpretable from field data. However, this assumption is strong because such a spatial-temporal interpretation may not always be available. A weaker, but still valuable, assumption will be that only several vertical sections of interpreted channels and lobes are available from the field. If only several two-dimensional geologies are available, identification of the three-dimensional geology, such as the three-dimensional lobe hierarchy, will be significantly more complicated. More uncertainties are involved in the problem, such as the number and the placement of cross sections, etc. The hierarchically constrained CRW-based mechanism in this study may be a starting point to study this problem. Given that an ensemble of realizations with similar lobe hierarchies are generated, hierarchies of two-dimensional sections of lobes may be available from various numbers of vertical cross sections obtained at different locations. Thus, the statistical relationship between hierarchies within two-dimensional realizations and the input patterns may be estimated. The challenges will still be huge uncertainties associated with two-dimensional sections.

- **Simulation of Subseismic Stratigraphies with Known Hierarchies in Experiments**

One of the original interests of static reservoir modeling is the reconstruction of stratigraphies at subseismic scales. Since the experiment equivalent scale of lobes in the real system was identified with the workflow in Chapter 3 and the experimental lobe hierarchy is available, it is reasonable to utilize hierarchical information in the experiment to model subseismic scale lobate bodies in the real environment. The primary challenge of this problem is how to control the overall hierarchy of lobes in the forward simulation scheme of surface-based models with an efficient implementation.

- **Thorough Utilization of the Experimental Data**

The dissertation presented a simple implementation of a lobe migration mechanism to demonstrate improvements led by the integration of experimental data in the surface-based approach. However, more possible improvements can be expected if the experimental data are more thoroughly utilized.

- i) Three-dimensional intermediate information

The prerequisite of modeling geological dynamic processes is the availability of three-dimensional intermediate topographies measured at a sufficient temporal resolution. Temporary solutions are implemented in this dissertation since the three-dimensional intermediate topography is not available in the experiments. Once the three-dimensional information is available, possible improvements are expected as follows: First, three-dimensional geometries of erosion and deposition can be extracted, so the correlation between the real erosion and deposition surfaces can be modeled more precisely than a general thickness function and a dimensionless ratio. Second, the noisy scale versus p-value plots for the polygonal similarity and shape similarity in Chapter 3 may exist either because of the improper measures or because of the insufficiency of describing the lobe with simply two-dimensional patterns. Thus, more complex

polygonal distance or shape dissimilarity measures are worth studying with three-dimensional geometries. Finally, the bifactor modeling scheme developed in Chapter 4 is no longer needed with three-dimensional intermediate information, since the conceptual compensational stacking rule can be precisely quantified. The CRW-based characterization and implementation can be extended to the three-dimensional space and is thus a good starting point. Lobe orientation, which is determined by the regional intermediate topography, can be quantitatively characterized with respect to intermediate topography instead of a relative angle apart from an interpreted basin flow orientation.

ii) Hierarchical simulation on more parameters

The presented CRW-based implementation of the lobe migration mechanism only constrains the hierarchy of MPPs in model realizations. As indicated earlier, it is optimal to maximize the use of input sequences by constraining hierarchies for all the parameters. However, the challenge is that implementations are expected to control the global hierarchy in a simple and efficient forward simulation scheme with the least number of empirical coefficients.

iii) Characterization and simulation of distributary channels

Distributary channels are also observed in experiments, thus channel sequences are sufficient for developing precise quantitative measures for channel geometry and channel hierarchy. Compared to lobes, channels are more complex in both morphology and hierarchy. Channel morphology is a stochastic function of intermediate topography. The stochastic function varies along the channel centerline toward the downstream direction. Hierarchy is determined by avulsion points, thus the number, location, and migration of avulsion points may be studied for modeling channel hierarchies. Techniques are expected to characterize, compare, and simulate this higher order problem.

## BIBLIOGRAPHY

- Alpak, Faruk O., Mark D. Barton, and David Castineira. 2011. "Retaining Geological Realism in Dynamic Modelling: A Channelized Turbidite Reservoir Example from West Africa." *Petroleum Geoscience* 17 (1): 35–52. doi:10.1144/1354-079309-033.
- Ankerst, Mihael, Markus M. Breunig, Hans-Peter Kriegel, and Jörg Sander. 1999. "OPTICS: Ordering Points to Identify the Clustering Structure." In *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data*, 49–60. SIGMOD'99. New York, NY, USA: ACM. doi:10.1145/304182.304187.
- Arpat, G. Burc, and Jef Caers. 2007. "Conditional Simulation with Patterns." *Mathematical Geology* 39 (2): 177–203. doi:10.1007/s11004-006-9075-3.
- Bergman, C. M., J. A. Schaefer, and S. N. Luttich. 2000. "Caribou Movement as a Correlated Random Walk." *Oecologia* 123(3): 364–74. doi:10.1007/s004420051023.
- Bertoncello, Antoine. 2011. "Conditioning Surface-Based Models to Well and Thickness Data". Phd Dissertation, Stanford University.
- Bertoncello, Antoine, Tao Sun, Hongmei Li, Gregoire Mariethoz, and Jef Caers. 2013. "Conditioning Surface-Based Geological Models to Well and Thickness Data." *Mathematical Geosciences* 45 (7): 873–93. doi:10.1007/s11004-013-9455-4.
- Breunig, Markus M., Hans-Peter Kriegel, Raymond T. Ng, and Jörg Sander. 1999. "OPTICS-OF: Identifying Local Outliers." In *Principles of Data Mining and Knowledge Discovery*, edited by Jan M. Żytkow and Jan Rauch, 262–70. Lecture Notes in Computer Science 1704. Springer Berlin Heidelberg.
- Caers, Jef. 2013. "Rapid Updating of Conceptual Model, Trend and Reservoir Facies Model at the Appraisal Stage" presented at the 2013 SCRF Annual Affiliate Meeting, May.
- Corder, Gregory W, and Foreman. 2009. *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*. Hoboken, N.J.: Wiley.

- Deptuck, Mark E., David J. W. Piper, Bruno Savoye, and Anne Gervais. 2008. "Dimensions and Architecture of Late Pleistocene Submarine Lobes off the Northern Margin of East Corsica." *Sedimentology* 55 (4): 869–98. doi:10.1111/j.1365-3091.2007.00926.x.
- Deutsch, C. V. 2002. *Geostatistical Reservoir Modeling*. Oxford; New York: Oxford University Press.
- Deutsch, C.V., and L. Wang. 1996. "Hierarchical Object-Based Geostatistical Modeling of Fluvial Reservoirs." In Society of Petroleum Engineers. doi:10.2118/36514-MS.
- Diggle, Peter. 2003. *Statistical Analysis of Spatial Point Patterns*. London; New York: Arnold ; Distributed by Oxford University Press.
- Droz, L., T. Marsset, H. Ondras, M. Lopez, B. Savoye, and F.-L. Spy-Anderson. 2003. "Architecture of an Active Mud-Rich Turbidite System: The Zaire Fan (Congo–Angola Margin Southeast Atlantic): Results from ZaAngo 1 and 2 Cruises." *AAPG Bulletin* 87 (7): 1145–68.
- Dryden, I. L., and K. V Mardia. 1998. *Statistical Shape Analysis*. Chichester; New York: John Wiley & Sons.
- Efron, Bradley, and Robert Tibshirani. 1994. *An Introduction to the Bootstrap*. New York: Chapman & Hall.
- Firle, Sascha, Riccardo Bommarco, Barbara Ekblom, and Mario Natiello. 1998. "The Influence of Movement And Resting Behavior On The Range Of Three Carabid Beetles." *Ecology* 79 (6): 2113–22. doi:10.1890/0012-9658(1998)079[2113:TIOMAR]2.0.CO;2.
- Forester, James D., Anthony R. Ives, Monica G. Turner, Dean P. Anderson, Daniel Fortin, Hawthorne L. Beyer, Douglas W. Smith, and Mark S. Boyce. 2007. "State-Space Models Link Elk Movement Patterns To Landscape Characteristics in Yellowstone National Park." *Ecological Monographs* 77 (2): 285–99. doi:10.1890/06-0534.
- Fortin, Daniel, Hawthorne L. Beyer, Mark S. Boyce, Douglas W. Smith, Thierry Duchesne, and Julie S. Mao. 2005. "Wolves Influence Elk Movements: Behavior Shapes A Trophic Cascade in Yellowstone National Park." *Ecology* 86 (5): 1320–30. doi:10.1890/04-0953.
- Fortin, Daniel, Juan M. Morales, and Mark S. Boyce. 2005. "Elk Winter Foraging at Fine Scale in Yellowstone National Park." *Oecologia* 145 (2): 334–42. doi:10.1007/s00442-005-0122-4.
- Garland, C. R., P. Haughton, R. F. King, and T. P. Moulds. 1999. "Capturing Reservoir Heterogeneity in a Sand-Rich Submarine Fan, Miller Field." *Geological Society, London, Petroleum Geology Conference Series* 5: 1199–1208. doi:10.1144/0051199.
- Gervais, A., T. Mulder, B. Savoye, and E. Gonthier. 2006. "Sediment Distribution and Evolution of Sedimentary Processes in a Small Sandy Turbidite System (Golo System, Mediterranean Sea): Implications for Various Geometries Based on Core Framework." *Geo-Marine Letters* 26 (6): 373–95. doi:10.1007/s00367-006-0045-z.

- Goovaerts, Pierre. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press.
- Groenenberg, Remco M., David M. Hodgson, Amandine Prédât, Stefan M. Luthi, and Stephen S. Flint. 2010. “Flow–Deposit Interaction in Submarine Lobes: Insights from Outcrop Observations and Realizations of a Process-Based Numerical Model.” *Journal of Sedimentary Research* 80 (3): 252–67. doi:10.2110/jsr.2010.028.
- Hajek, Elizabeth A., Paul L. Heller, and Benjamin A. Sheets. 2010. “Significance of Channel-Belt Clustering in Alluvial Basins.” *Geology* 38 (6): 535–38. doi:10.1130/G30783.1.
- Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. *The Elements of Statistical Learning, Second Edition: Data Mining, Inference, and Prediction*. New York: Springer.
- Honarkhah, Mehrdad, and Jef Caers. 2012. “Direct Pattern-Based Simulation of Non-Stationary Geostatistical Models.” *Mathematical Geosciences* 44 (6): 651–72. doi:10.1007/s11004-012-9413-6.
- Jegou, I., B. Savoye, C. Pirmez, and L. Droz. 2008. “Channel-Mouth Lobe Complex of the Recent Amazon Fan: The Missing Piece.” *Marine Geology* 252 (1–2): 62–77. doi:10.1016/j.margeo.2008.03.004.
- Johnson, Stephen C. 1967. “Hierarchical Clustering Schemes.” *Psychometrika* 32 (3): 241–54. doi:10.1007/BF02289588.
- Kareiva, P. M., and N. Shigesada. 1983. “Analyzing Insect Movement as a Correlated Random Walk.” *Oecologia* 56 (2-3): 234–38. doi:10.1007/BF00379695.
- Kendall, David G. 1989. “A Survey of the Statistical Theory of Shape.” *Statistical Science* 4 (2): 87–99.
- Lapointe, François-Joseph, and Pierre Legendre. 1995. “Comparison Tests for Dendrograms: A Comparative Evaluation.” *Journal of Classification* 12 (2): 265–82. doi:10.1007/BF03040858.
- Li, Hongmei. 2008. “Hierarchic Modeling And History Matching Of Multi-Scale Flow Barriers In Channelized Reservoirs”. Phd Dissertation, Stanford University.
- Li, Hongmei, and Jef Caers. 2011. “Geological Modelling and History Matching of Multi-Scale Flow Barriers in Channelized Reservoirs: Methodology and Application.” *Petroleum Geoscience* 17 (1): 17–34. doi:10.1144/1354-079309-825.
- Marell, A. 2002. “Foraging and Movement Paths of Female Reindeer : Insights from Fractal Analysis, Correlated Random Walks, and Levy Flights.” *Can J Zool* 80: 854–65. doi:10.1139/z02-061.
- McHargue, T., M.J. Pyrcz, M.D. Sullivan, J.D. Clark, A. Fildani, B.W. Romans, J.A. Covault, M. Levy, H.W. Posamentier, and N.J. Drinkwater. 2011. “Architecture of Turbidite Channel Systems on the Continental Slope: Patterns and Predictions.” *Marine and Petroleum Geology* 28 (3): 728–43. doi:10.1016/j.marpetgeo.2010.07.008.
- Michael, H. A., H. Li, A. Boucher, T. Sun, J. Caers, and S. M. Gorelick. 2010. “Combining Geologic-Process Models and Geostatistics for Conditional

- Simulation of 3-D Subsurface Heterogeneity.” *Water Resources Research* 46 (5). doi:10.1029/2009WR008414.
- Miller, James, Tao Sun, Hongmei Li, Jonathon Stewart, Coralie Genty, Dachang Li, and Colin Lyttle. 2008. “Direct Modeling of Reservoirs Through Forward Process-Based Models: Can We Get There?” In Society of Petroleum Engineers. doi:10.2523/12729-MS.
- Paola, Chris, Kyle Straub, David Mohrig, and Liam Reinhardt. 2009. “The ‘unreasonable Effectiveness’ of Stratigraphic and Geomorphic Experiments.” *Earth-Science Reviews* 97 (1–4): 1–43. doi:10.1016/j.earscirev.2009.05.003.
- Phipps, J. B. 1971. “Dendrogram Topology.” *Systematic Biology* 20 (3): 306–8. doi:10.2307/2412343.
- Podani, J., and T. A. Dickinson. 1984. “Comparison of Dendrograms: A Multivariate Approach.” *Canadian Journal of Botany* 62 (12): 2765–78. doi:10.1139/b84-369.
- Prélat, A., J.A. Covault, D.M. Hodgson, A. Fildani, and S.S. Flint. 2010. “Intrinsic Controls on the Range of Volumes, Morphologies, and Dimensions of Submarine Lobes.” *Sedimentary Geology* 232 (1–2): 66–76. doi:10.1016/j.sedgeo.2010.09.010.
- Prélat, A., D. M. Hodgson, and S. S. Flint. 2009. “Evolution, Architecture and Hierarchy of Distributary Deep-Water Deposits: A High-Resolution Outcrop Investigation from the Permian Karoo Basin, South Africa.” *Sedimentology* 56 (7): 2132–54. doi:10.1111/j.1365-3091.2009.01073.x.
- Pyrzcz, Michael J., Octavian Catuneanu, and Clayton V. Deutsch. 2005. “Stochastic Surface-Based Modeling of Turbidite Lobes.” *AAPG Bulletin* 89 (2): 177–91. doi:10.1306/09220403112.
- Pyrzcz, Michael J., Timothy McHargue, Julian Clark, Morgan Sullivan, and Sebastien Strebelle. 2012. “Event-Based Geostatistical Modeling: Description and Applications.” In *Geostatistics Oslo 2012*, edited by Petter Abrahamsen, Ragnar Hauge, and Odd Kolbjørnsen, 27–38. Quantitative Geology and Geostatistics 17. Springer Netherlands.
- Pyrzcz, Michael J., and Sebastien Strebelle. 2006. “Event-Based Geostatistical Modeling: Application to Deep-Water Systems.” In *Reservoir Characterization: Integrating Technology and Business Practices, Gulf Coast Section, SEPM Twenty-Sixth Annual Research Conference*, 893–922.
- Saller, Arthur H., Jesse T. Noah, Alif Prama Ruzuar, and Rhys Schneider. 2004. “Linked Lowstand Delta to Basin-Floor Fan Deposition, Offshore Indonesia: An Analog for Deep-Water Reservoir Systems.” *AAPG Bulletin* 88 (1): 21–46.
- Saller, Arthur, Ken Werner, Fransiskus Sugiaman, Andre Cebastian, Ron May, David Glenn, and Craig Barker. 2008. “Characteristics of Pleistocene Deep-Water Fan Lobes and Their Application to an Upper Miocene Reservoir Model, Offshore East Kalimantan, Indonesia.” *AAPG Bulletin* 92 (7): 919–49. doi:10.1306/03310807110.
- Sander, Jörg, Xuejie Qin, Zhiyong Lu, Nan Niu, and Alex Kovarsky. 2003. “Automatic Extraction of Clusters from Hierarchical Clustering Representations.” In

- Advances in Knowledge Discovery and Data Mining*, edited by Kyu-Young Whang, Jongwoo Jeon, Kyuseok Shim, and Jaideep Srivastava, 75–87. Lecture Notes in Computer Science 2637. Springer Berlin Heidelberg.
- Scheidt, Céline, and Jef Caers. 2009. “Uncertainty Quantification in Reservoir Performance Using Distances and Kernel Methods—Application to a West Africa Deepwater Turbidite Reservoir.” *SPE Journal* 14 (4). doi:10.2118/118740-PA.
- Sheets, Benjamin A., Chris Paola, and J. Michael Kelberer. 2009. “Creation and Preservation of Channel-Form Sand Bodies in an Experimental Alluvial System.” In *Sedimentary Processes, Environments and Basins*, edited by Gary Nichols, Ed Williams, and Chris Paola, 555–67. Blackwell Publishing Ltd.
- Shmaryan, L.E., and C.V. Deutsch. 1999. “Object-Based Modeling of Fluvial / Deepwater Reservoirs with Fast Data Conditioning: Methodology and Case Studies.” In Society of Petroleum Engineers. doi:10.2118/56821-MS.
- Sokal, Robert R., and F. James Rohlf. 1962. “The Comparison of Dendrograms by Objective Methods.” *Taxon* 11 (2): 33–40.
- Straub, Kyle M., Chris Paola, David Mohrig, Matthew A. Wolinsky, and Terra George. 2009. “Compensational Stacking of Channelized Sedimentary Deposits.” *Journal of Sedimentary Research* 79 (9): 673–88. doi:10.2110/jsr.2009.070.
- Strebelle, Sebastien. 2002. “Conditional Simulation of Complex Geological Structures Using Multiple-Point Statistics.” *Mathematical Geology* 34 (1): 1–21.
- Sullivan, Morgan, Gerrick Jensen, Frank Goulding, David Jennette, Lincoln Foreman, and David Stern. 2000. “Architectural Analysis of Deep-Water Outcrops: Implications for Exploration and Development of the Diana Sub-Basin, Western Gulf of Mexico.” In *Deep-Water Reservoirs of the World: Gulf Coast Section SEPM Foundation 20th Annual Research Conference*, 1010–32.
- Sylvester, Zoltán, Carlos Pirmez, and Alessandro Cantelli. 2011. “A Model of Submarine Channel-Levee Evolution Based on Channel Trajectories: Implications for Stratigraphic Architecture.” *Marine and Petroleum Geology* 28 (3): 716–27. doi:10.1016/j.marpetgeo.2010.05.012.
- Tan, Xiaojin, Pejman Tahmasebi, and Jef Caers. 2013. “Comparing Training-Image Based Algorithms Using an Analysis of Distance.” *Mathematical Geosciences*: 1–21. doi:10.1007/s11004-013-9482-1.
- Van Wagoner, J. C., DCJD Hoyal, N. L. Adair, T. Sun, R. T. Beaubouef, P. A. Deffenbaugh, C. Huh, and D. Li. 2013. “Energy Dissipation and the Fundamental Shape of Siliciclastic Sedimentary Bodies.” Accessed September 30.
- Veltkamp, R.C. 2001. “Shape Matching: Similarity Measures and Algorithms.” In *Shape Modeling and Applications, SMI 2001 International Conference On.*, 188–97. doi:10.1109/SMA.2001.923389.
- Zhang, Tuanfeng, Paul Switzer, and Andre Journal. 2006. “Filter-Based Classification of Training Image Patterns for Spatial Simulation.” *Mathematical Geology* 38 (1): 63–80. doi:10.1007/s11004-005-9004-x.