

Worked examples of Optimal Cycle Analysis

To demonstrate how this method works, two synthetic vertical successions (Fig 1A) have been constructed and analysed to identify what cyclicity they contain. The analyses successfully identify an optimal cyclic arrangement of facies classes in each succession. This is then the basis for further analysis to determine how much evidence of order in the form of cyclicity is present in each succession. Matlab code to carry out the analysis is available at <https://csdms.colorado.edu/wiki/Model:OptimalCycleID>.

How the synthetic strata were generated

Each synthetic succession of strata is 15m thick and composed of fifty lithological units, classified as five distinct facies (Fig. 1A) coded as shown in Table 1. Both successions were generated initially with a perfectly cyclical arrangement of five facies from mudstone (mst), to siltstone (slt), then limestone (lst), fine sandstone (fst) and medium sandstone (msst). Note that the facies code is only the same as the position in the cycle for the limestone lithology facies; this significance of this for the depiction of the cyclicity in a transition probability matrix will become apparent.

Taking as a starting point this 15m thick perfectly cyclical succession with 10 cycles composed of 50 lithological units, the two synthetic vertical successions were constructed by introducing either 10 or 20 randomly assigned facies codes and thicknesses at random points in the succession. For each randomly chosen facies code introduced, a check was conducted to ensure that this introduction did not lead to juxtaposition of the same facies, and if it did another random facies code was chosen randomly until juxtaposition did not occur. In the 10 substitution strata, there are still three complete examples of the original cycles starting at 1m, 9m and 14m (Fig. 1A). In the 20 substitution strata, there are no examples of the original cycles preserved; the longest remnant is an arrangement of four facies at 16.5m that is missing the medium sandstone at its top (Fig. 2A).

These two synthetic examples simulate a situation where strata are being analysed to test a hypothesis that an optimum most cyclic arrangement of facies can be determined. Once this most cyclic arrangement of facies has been identified, the strata can be further analysed to determine if they represent an ordered arrangement of facies and thicknesses unlikely to have occurred by chance (Burgess, 2016).

Constructing the transition probability matrices and calculating m and d values

A transition probability (TP) matrix for a vertical succession of strata records the probability of transition from one facies class to each of the other facies classes. The TP matrix is constructed by recording all older-to-younger facies transitions in the matrix and converted these counts to probabilities by dividing by the total number of transitions from each facies class. An m statistic, indicative of the degree of order present in the vertical succession of strata (Burgess, 2016), can be calculated from a TP matrix such that

$$m = \operatorname{argmax}_{j=1..F-1} \left\{ \frac{\sum \operatorname{diag}(T_j) + \sum \operatorname{diag}(T_{-(F-j)})}{F} \right\} - \operatorname{argmin}_{j=1..F-1} \left\{ \frac{\sum \operatorname{diag}(T_j) + \sum \operatorname{diag}(T_{-(F-j)})}{F} \right\} \quad \text{eq. 1}$$

where F is the number of facies classes, in this case 5, j is the offset value from the main matrix diagonal, diag is a function to find all the elements in a diagonal of matrix T with offset j from the main matrix diagonal, and argmin and argmax are mathematical functions to find the minimum and maximum values in a series composed of each j th offset diagonal in the matrix (Fig. 1B). The algorithm represented by this equation calculates an average probability for each offset matrix diagonal, for $j=1..4$. For each diagonal offset value j the total number of cells being compared is always the same because it includes one or more cells from both side of the main matrix diagonal (Fig. 1B). Once the averages are calculated, m is calculated as the difference between the highest average probability found in an offset diagonal, and the lowest average probability found in a different offset diagonal. In the case of the 10-swap strata for the original facies class coding, the highest average probability in any diagonal is in the $j=2$ and $j=-3$ diagonal (Fig. 1B), and the lowest is in the $j=1$ and $j=-4$ diagonal, and the latter subtracted from the former gives $m=0.1824$ (Fig. 1B).

Burgess (2016) showed that the value of the m statistic is sensitive to the order in which the facies classes are arranged in the TP matrix. For example, for the 10 substitution strata, two different facies class orders in the TP matrix give $m=0.1824$ for one row ordering (Fig. 1B) and $m=0.7124$ for another row ordering (Fig. 1D). This sensitivity can be exploited to find the most cyclic arrangement of facies for the vertical succession. This is done by calculating a TP matrix for the vertical succession with each possible row and column order for the facies classes. Since in this case there are five facies classes in each succession, there are $5!$ or 120 possible combinations of the facies classes, so a TP matrix and a resulting m value is calculated for each of these 120 combinations. For each TP matrix a mean value of the $j=1$ and $j=-4$ diagonal is also calculated so

$$\mu = \frac{\sum \operatorname{diag}(T_{j=1}) + \operatorname{diag}(T_{j=-(F-1)})}{F} \quad \text{Eq. 2}$$

The one-offset diagonal, marked with blue outlined squares in each TP matrix in Figures 1B and 2B, is particularly important because any asymmetric cyclicity in the strata leads to high probability values on this diagonal (Fig. 1D). Note that symmetric cycles are slightly more complicated because the highest p values occur on both the $j=1$ and the $j=-1$ diagonals (Burgess, 2016), but this cyclicity is still amenable to analysis by this method. Finally, for each TP matrix, the product of m and the one-offset diagonal mean is calculated

$$d = m\mu \quad \text{Eq. 3}$$

The value of d will be high in cases where high probabilities occur on the one-offset diagonal, so d can be used to identify TP matrices that have both a high m value, indicating an element of order in the strata, and a high value of μ which may be strong evidence for cyclicity.

Using the m and d values to find optimally cyclical arrangements of facies classes

From the 120 TP matrices, calculated from the 10 substitution strata for all the possible facies class row orders, three row and column orders show m values ranging from the lowest calculated, which

is $m=0.1824$ (Fig 1C), to the highest value calculated, which is $m=0.7124$ (Fig. 1D). For this 10 substitution succession the original facies class coding happens to also give the lowest m value (Fig. 1B). However, for the 20 substitution succession the original facies coding gives $m=0.0772$ (Fig. 2B) and a different facies row order gives the lowest value $m=0.0717$ (Fig. 2C).

For the 10 substitution vertical succession, most calculated TP matrices give m values between 0.1 and 0.4 with only 20 of the total permutations giving values of 0.7124 (Fig. 1E). Similarly most permutations give $d < 1.5$, with only five permutations giving $d=5.0745$ (Fig. 1F). These five permutations represent an ordering of the five facies classes that defines the optimal cycle for the 10 substitution succession of strata (e.g. Fig. 1D). Most importantly, this is the same order of facies classes that specifies the originally defined cycles in the strata prior to the random substitution of 10 facies units (Table 1).

For the 20 substitution vertical succession, the m values range between 0.07 and 0.32, with an equal number of permutations giving values between 0.15 and 0.40 (Fig. 2E). Most permutations give $d < 1.2$ with only five permutations giving the highest value $d=1.3689$ (Fig. 2F). These five permutations represent an ordering of the five facies classes that defines the optimal cycle for the 20 substitution succession of strata (e.g. Fig 2D). Again, most importantly, this is the order of facies classes that specifies the originally defined cycles in the strata prior to the random substitution of 20 facies units (Table 1).

These two examples demonstrate how this method can extract the most cyclic facies class order from a vertical succession of strata. With these two synthetic strata cases we know that there was a strongly ordered cyclical arrangement of facies classes present prior to the random substitutions. Elements of this order are still readily apparent in the 10 substitution succession (Fig. 1A), but less so in the 20 substitution succession where there is not one single case of a cycle as originally defined preserved in the strata (Fig. 2A). Nevertheless, in both cases, analysis of the successions by calculating m and d values from the TP matrices for all permutations of facies class row ordering reveals the originally cyclical succession of facies classes (Figs. 1D and 2D). Success with controlled synthetic cases suggests that the method could be usefully applied for the same purpose to vertical successions measured from outcrop or from subsurface data such as boreholes.

The final step – testing for evidence of order

Burgess (2016) presented a method to use the m value calculated from a TP matrix to determine what evidence strata contain for order, for example in the form of cyclicity. The optimal facies class order identified for a succession can be used in this method.

Application of this analysis reveals strong evidence for order in the 10 substitution succession. The calculated m value is 0.712, related to high probabilities in the one offset TP matrix diagonal (Fig 3B). This m value lies well outside the range of m values produced for randomly shuffled versions of the strata (Fig. 3C), giving $p=0.0000$ (Table 2), showing that the strata are highly unlikely to have arisen by chance, and therefore providing strong evidence of order. This is hardly surprising since we know that the succession was produced from an originally perfectly cyclical section with only 10 random facies substitutions, leaving substantial remnants of cyclicity present (Fig. 3A). However, this is much

more apparent from the analysis using the optimal facies class ordering; the p value from the same method used with the original facies class coding gave $p=0.4046$ which provides no evidence for order in the strata (Burgess, 2016). This indicates the utility of knowing the optimal arrangement of facies classes to yield most evidence for order in the analysis.

Considering the 20 substitution strata, this should contain less evidence for order since the number of random substitutions is higher and therefore less cyclicity should be preserved in the strata. For the 20 substitution strata $m=0.322$ (Table 1), reflecting a less obviously ordered TP matrix (Fig 3E). This m value falls with the range m values produced for randomly shuffled versions of the strata (Fig. 3F), giving $p=0.0342$ (Table 2), providing only weak evidence for order in the strata. However, even this weak evidence is more than provided by the same analysis with the original facies coding, from which $m=0.077$ and the $p=0.9406$.

Conclusions

These worked examples demonstrates that searching all permutations of the TP matrices to find the facies class row numbering that gives the highest m and d value can be used to define an order of facies classes that is optimal for defining and identifying any order that may be present in the strata.

Table 1

Lithology	Position in Cycle	Facies Code
Mudstone	1	2
Siltstone	2	1
Limestone	3	3
Fine sandstone	4	5
Medium sandstone	5	4

Table 2

Name of succession	M statistic	P value	Interpretation
10 swap succession, optimal coding	0.712	0.0000	Strong evidence for order
10 swap succession, original coding	0.182	0.4046	No evidence for order
20 swap succession, optimal coding	0.322	0.0342	Weak evidence for order
20 swap succession, original coding	0.077	0.9406	No evidence for order

Figure 1. A. The 10 swap vertical succession of facies. Note the examples of preserved original cycles starting at 1m, 9m and 14m. B. A TP matrix calculated using the original facies coding, shown on the right, giving $m = 0.1824$ and $d = 1.338$. Note that the values of j indicate numbering of the offset matrix diagonals referred to in equations 1 and 2. The one-offset diagonal cells are indicated with a blue outline and include the single cell in the $j=-4$ diagonal due to the wrap-around effect in the matrix. C. A TP matrix calculated using a low-scoring facies coding, shown on the right, giving $m=0.1824$ and $d=3.1$. D. A TP matrix calculated with the optimal facies coding, shown on the right, giving $m=0.7124$ and $d=7.1$. A value of m this close to 1 is unlikely to occur by chance arrangement of facies. E. A histogram showing frequency of m values from the TP matrices calculated for all 120 facies class permutations. Most m values are between 0.1 and 0.4 with only 20 of the total permutations giving the highest m values of 0.7124. F. A histogram showing the frequency of d values from the 120 TP matrices. Most permutations give $d<1.5$, with only five permutations giving $d= 5.0745$ and defining the optimal cycle shown in D. This is the same order of facies classes for the originally defined cycles in these strata (Table 1).

Figure 2. A. The 20 swap vertical succession of facies. Note there are no examples of preserved original cycles in this succession. B. A TP matrix calculated using the original facies coding, shown on the right, giving $m = 0.0772$ and $d = 2.083$. C. A TP matrix calculated using a low-scoring facies coding, shown on the right, giving $m=0.0717$ and $d=2.2$. D. A TP matrix calculated with the optimal facies coding, shown on the right, giving $m=0.3217$ and $d=4.3$. E. A histogram showing frequency of m values from the TP matrices calculated for all 120 facies class permutations. All m values are between 0.05 and 0.40. F. A histogram showing the frequency of d values from the 120 TP matrices. Most permutations give $d<1.2$ with only five permutations giving the highest value $d=1.3689$ and defines the optimal cycle shown in D. Once again, this is the same order of facies classes for the originally defined cycles in these strata (Table 1) and it has been found despite no occurrences of this order of facies in the vertical succession (A).

Figure 3. A. The 10 swap vertical succession of facies. B. The TP matrix calculated for the facies class row ordering that generated the highest m and d values. C. Results from the Monte Carlo analysis showing that this m value lies well outside the range of m values produced for randomly shuffled versions of the strata, giving a p value of 0.0000 (Table 2), showing that the strata are highly unlikely to have arisen by chance, and therefore providing strong evidence of order. D. The 20 swap vertical succession of facies. E. The TP matrix calculated for the facies class row ordering that generated the highest m and d values. F. Results from the Monte Carlo analysis showing that this m value lies within the tail of the m values produced for randomly shuffled versions of the strata, giving a p value of 0.0342 (Table 2), providing only weak evidence for order in the strata, even when the strata are analysed with this optimal facies class row ordering.

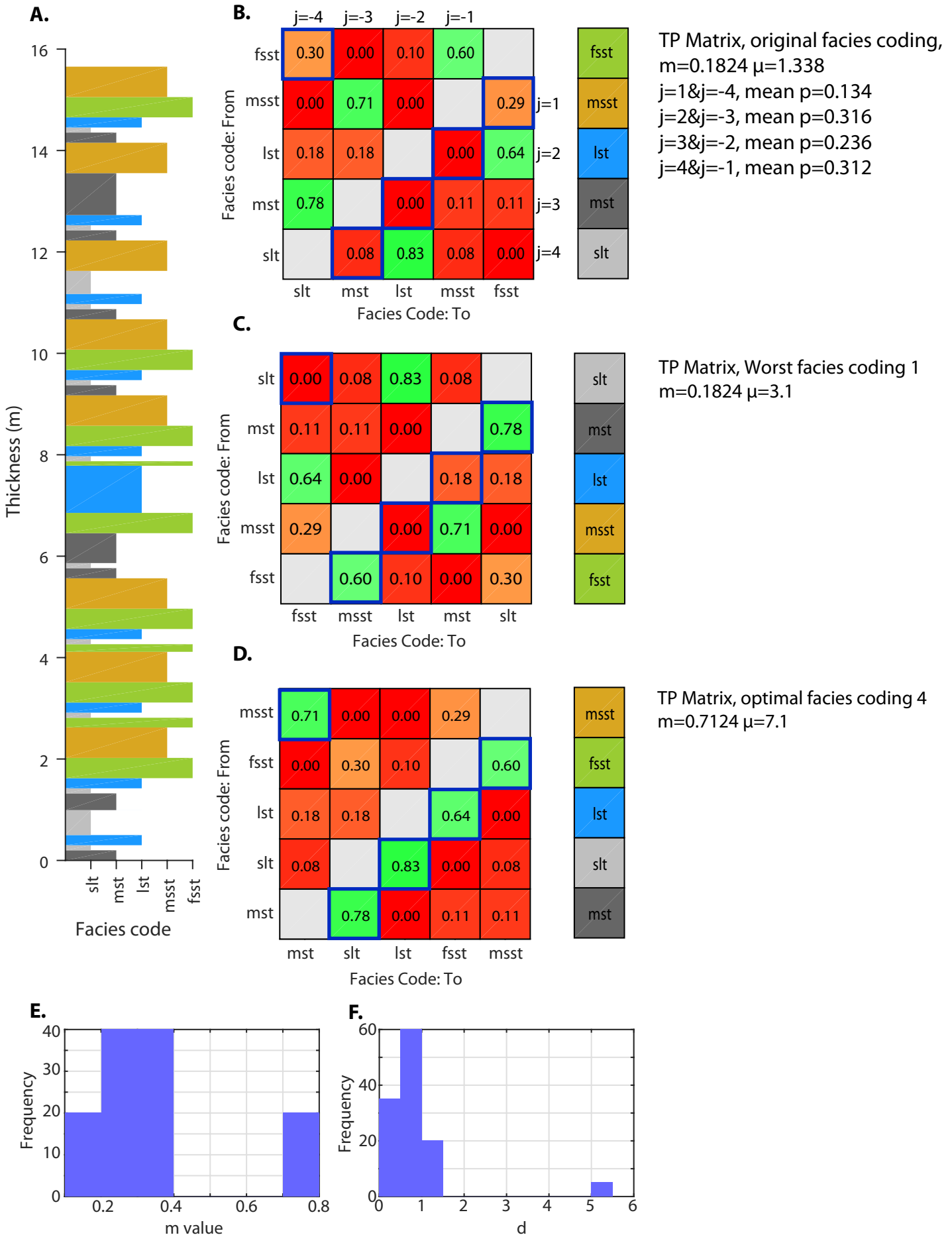


Figure 1

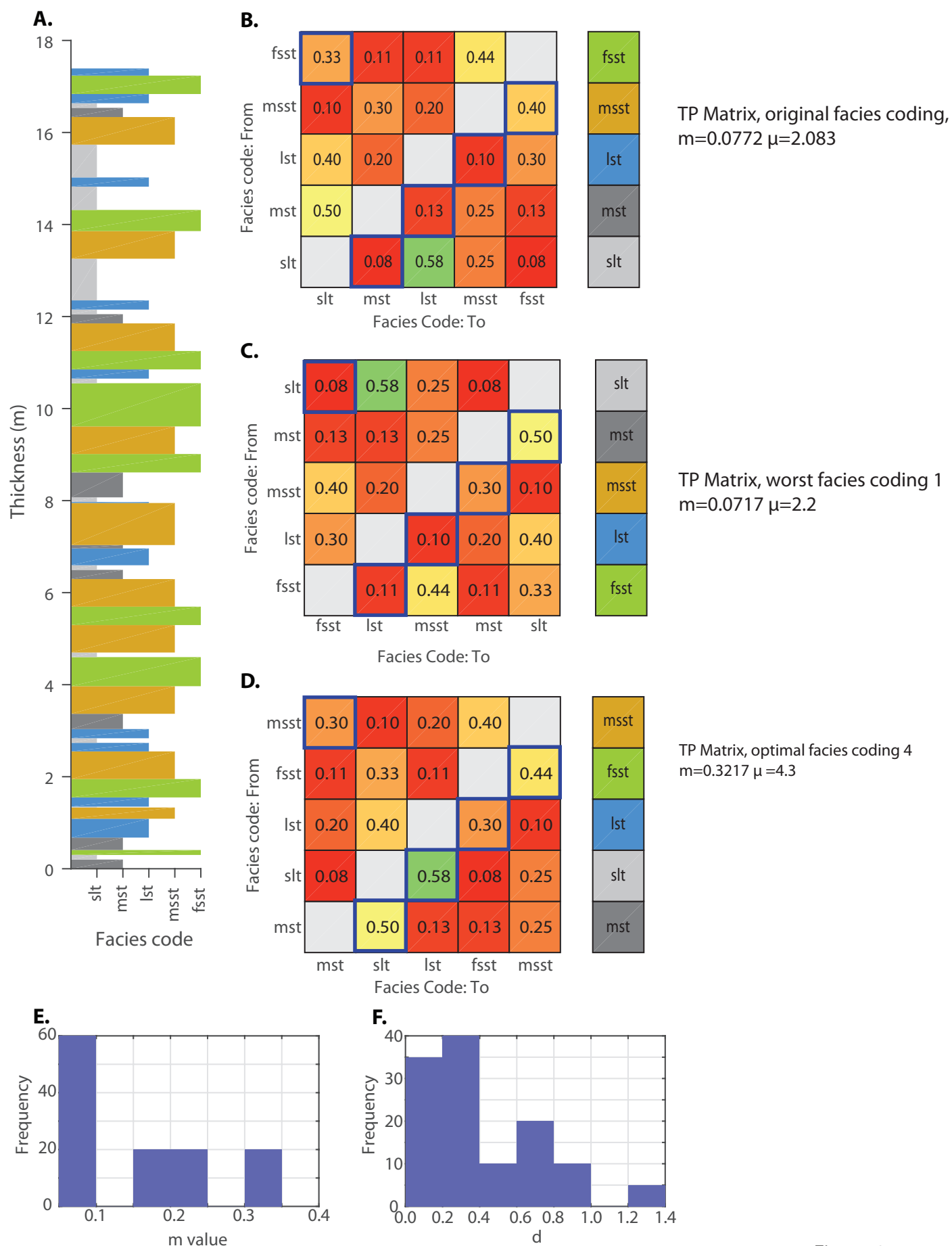
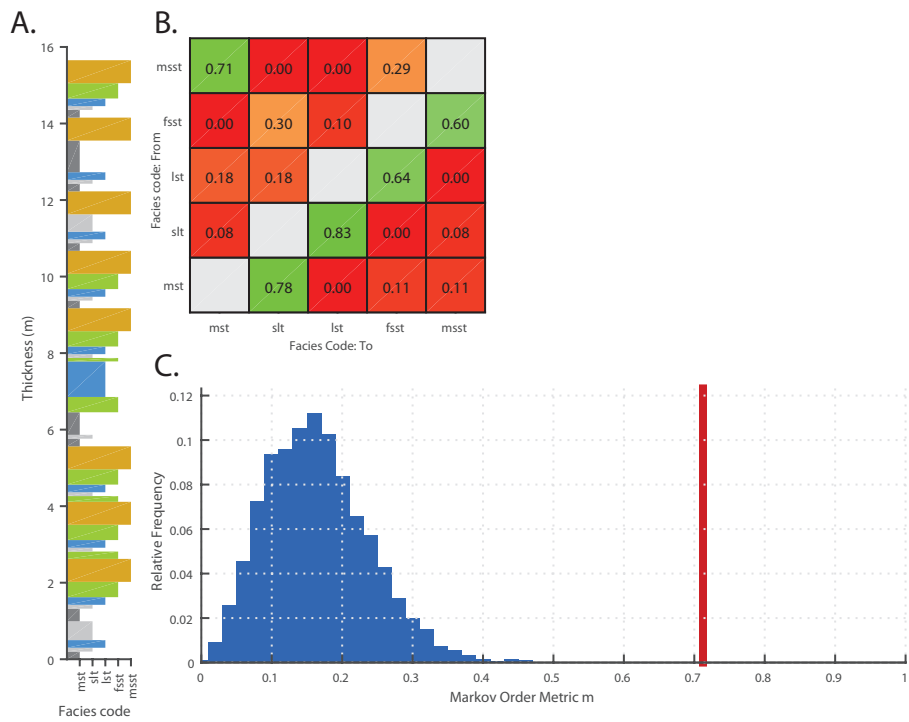


Figure 2



3

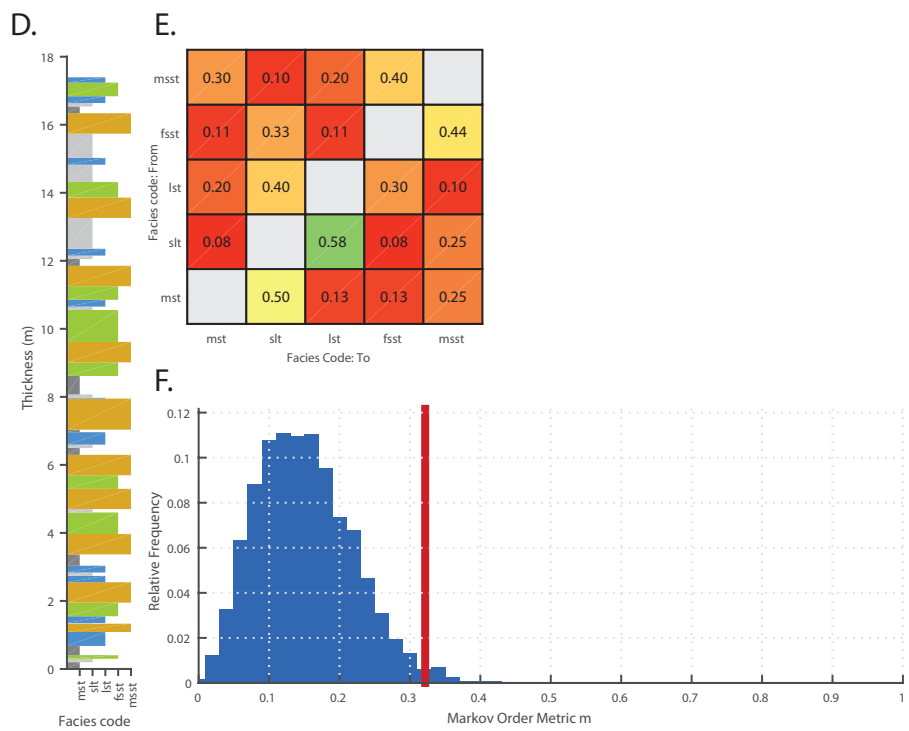


Figure 3