AI seismic domains: case study showing geological utility with reference to a fractured reservoir.

Cathal G Dillon[1]

[1]*Cathal Dillon is a software engineer at Badley Geosciences, North Beck House, North Beck Lane, Hundelby, Spilsby, Lincolnshire, UK, PE23 5NB.*

This study evaluates the utility of AI-derived seismic domains for identifying fracture-prone regions within a Cretaceous chalk reservoir in the Gorm Field, southern North Sea, where fracture density exerts a primary control on reservoir quality and permeability. AI domain models were generated through unsupervised classification of seismic motifs, producing spatially coherent categories that reflect commonality in seismic response rather than explicit geological classes. Fracture-domain relationships were quantified using an enrichment ratio comparing the probability of a domain occurring at fracture locations relative to its background occurrence along the well trajectories. The presence of significant spatial autocorrelation, in both fracture distributions and seismic domains, invalidates conventional independence-based statistical testing. To address this, a suite of stochastic baseline models was generated using segmented fractal random fields with comparable autocorrelation structures to the AI domain models. These baseline models served as null hypotheses against which the AI-derived enrichments were evaluated (AI domain models vs stochastic domain models). Results show that AI seismic domain models consistently produce a single enriched domain associated with fracture occurrence, accompanied by depletion in the remaining domains, whereas stochastic baseline models typically generate diffuse enrichment across multiple categories. This pattern is reproducible across simulations and resolutions. Furthermore, sensitivity analysis of the enrichment scores shows that the stochastic domain models are markedly less stable. Taken together, all the analysis indicated that the AI-derived seismic domains capture a genuine geological signal linked to fracture distribution within the Top Chalk reservoir. The results suggest that, even where fracture signatures lie below direct seismic resolution, AI domain analysis can provide predictive information regarding relative fracture abundance and therefore reservoir quality. This demonstrates the potential of AI seismic domains as a practical reservoir characterization tool in structurally complex fractured systems.

Keywords: *AI/ML, domains, T7, fractures, reservoir.*

Introduction

Seismic domains are effective at mapping subtle changes in seismic signal that may or may not be perceptible. Very often these are related to systematic changes in geological, geophysical and reservoir character. The study looks at how AI domains may help map the presence of more intensely fractured zones even where their explicit signature lies below seismic resolution. A Cretaceous chalk reservoir from the North Sea, where one of the main controls on reservoir quality is fracture density, provides a good case study.

Seismic domains

It is important to note that variability in the output from the AI analysis - namely seismic domains - reflect regions with similar seismic motifs which is strongly dependent on any geological character that may impact contrast in acoustic impedance: pore fluid, mineralogy, structural character, diagenesis etc. And for that reason, the seismic domains should be viewed as a method of finding commonality in seismic signal that reflects, imperfectly, commonality in geology rather than being a function of geology itself. The technique employed here is outlined in detail in Dillon 2026 under the 3D section.

How to interpret AI domain models

The output of an AI domain model can be viewed as a coded map in which regions assigned similar codes tend to share seismic motifs that are more similar to one another. The method is pseudo-stochastic in nature (Dillon 2026) so each new model is a non-unique solution. The specific domain codes themselves have no intrinsic meaning beyond classification. Different models or simulations can converge on very similar solutions, producing maps that appear nearly identical even when the assigned code values have been permuted or shuffled – referred to as category flipping. Consequently, the spatial or geographical pattern is more significant than the explicit numerical values of the domain codes themselves.

As outlined in Dillon (2026), most of the variation in geological systems can likely be characterized using only four or five domains. This appears reasonable given the relative homogeneity of the reservoir chalk examined in this study. Accordingly, five domains were selected for the present analysis.

Geological setting

The current study looks at a Cretaceous chalk reservoir (hereafter Top Chalk) in the Gorm Field, southern North Sea (Fig. 1). Its geological architecture is best described as dome-shaped and faulted owing to a combination of salt diapirism and regional east-west extension (Fig. 1). The chalk is thinner above the diaper than on its flanks. Freeman et al. (2015) attribute this primarily to mechanical emplacement of salt during continued diaperism and significant concurrent fault movement along large NW-SE trending faults.

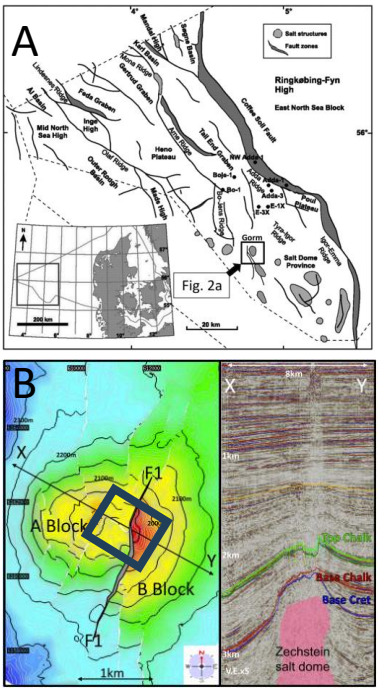


Figure 1. Location of Grom field and top chalk reservoir (after Freeman et al. 2015). A) Location map of Grom field, North Sea. B) Overview of the top chalk surface and diapiric setting. The inset rectangle denotes the location of the current study area.

The result is that the chalk unit has been significantly altered by continued periods of brittle deformation giving rise to areas of intense fracturing on the crest. For that reason, the author decided to focus on this part of the

reservoir since this is where supporting well data is concentrated.

Seismic data

AI analysis is carried out on a single depth converted 3D amplitude volume. This same volume was used in Freeman et al. (2015) for the purposes of advanced geomechanical modelling. They judged that the seismic resolution was limited to faults with offsets more than 20m thus indicating the possibility of sizable sub-seismic faulting.

Well data

A total of 13 wells were available with fracture records derived after interpretation of image logs (Fig. 2). It should be noted that the wells are radially arranged, do not penetrate the crest of the reservoir and are concentrated in the main western graben. It must also be noted that as with all well image log interpretations, fracture logs of this kind often suffer from some degree of orientation bias (fractures striking concordant to the well azimuth are often under-sampled), drilling induced fractures can be difficult to decipher and smaller fractures in particular have an ambiguous origin. Therefore, the current study follows a similar criterion to remove structurally insignificant fractures as set out in

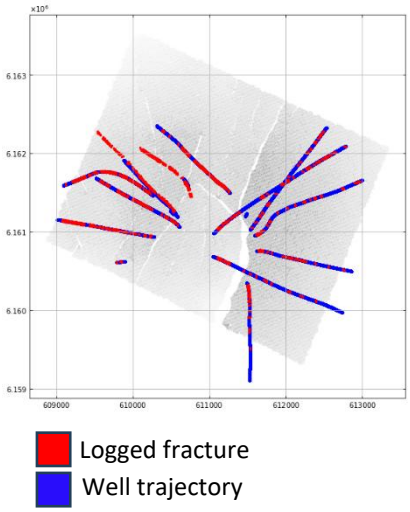


Figure 2. Wells used in this analysis.

Freeman et al. (2015) - those that do not fully transect the borehole are removed as they are considered here to have lesser significance in terms of subtle seismic character and therefore possible seismic response.

Methods

All the domain models are sampled at trajectory points along the well path (fractured and unfractured). Once extracted as well data the model data can be analysed to decipher any systematic relationship between domain code and fracture locations. The method employed here is known as the enrichment ratio (E) and is defined by:

$$E_i = \frac{P(C_i|F)}{P(C_i)} \quad [1]$$

where $P(C_i|F)$ is the probability of finding domain code C_i at a fracture location, while $P(C_i)$ is the probability of finding the same domain code at any point along the well – i denotes a specific domain code within a model. The data is typically presented as an enrichment profile (Fig. 3). Hereafter, we use $P(C|F)$ and $P(C)$ where general rules hold for any single domain within a model:

$$E = \frac{P(C|F)}{P(C)} \quad [2]$$

In short, $P(C)$ acts as our background calibration by-which we can measure whether the presence of C is depleted or enriched with respect to fracture locations.

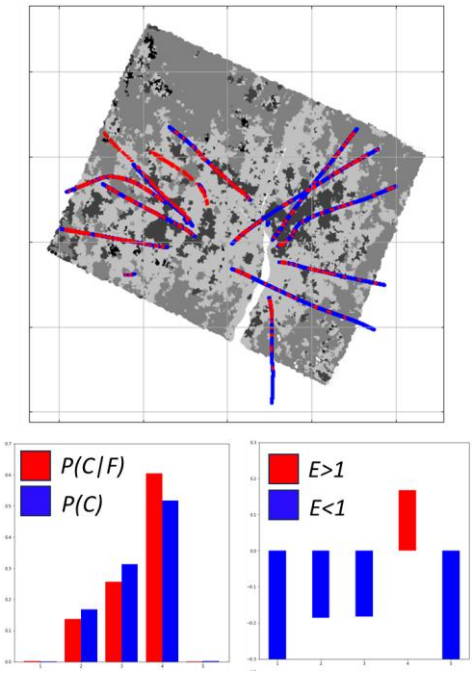


Figure 3. Example of AI domain model with associated probability values for each domain (bottom left) and computed enrichment (bottom right).

It should be noted that around c 61% of the well locations lie within the lateral resolvable seismic distance (c. 18m) of a fracture location (Fig. 2). That means the perfect enrichment for a single domain code (denoted by subscript *i*) will be E_i = 1.6. To elaborate, a domain code that is found only in the fractured zones (*i.e.* $P(C_i|F)=1.0$) will be found at 61% of the well locations ($P(C_i)=0.61$). This does not mean that *E* cannot be greater than 1.6, just that for our given data configuration that is the perfect enrichment. Indeed, *E* can become very large when *P(C)* is very small. This is known as denominator explosion and is typical when dealing with marginal categories. Other configurations that can lead to unduly high enrichments (here $E > 2$) is where a code is sparsely distributed but clustered. By chance these can give high *E* values. That's why it is important to run many simulations to ensure "flukes" don't dominate the signal.

Test Criteria

For this study and data configuration the following rules can be established:

$E > 1$: Enrichment.
$E = 1$: Neutral.
$E < 1$: Depletion.
$E > 2$ (may be superficial, check across many runs)

Finally, the support for seismic domain-fracture relationships would require a number of conditions to be met: 1) enrichment (*E>1*) for a single domain code with corresponding depletion (*E<1*) in the others (see stochastic discussion below) 2) this pattern should be repeated over all or most simulations and 3) the distribution (spatial variation) of domain codes should be similar between simulations.

Baseline model

Because the fractured regions are "continuous" in nature they are also likely highly autocorrelated (i.e. nearby areas tend to have similar character, rather than varying independently from one point to the next). This can be seen in Fig. 2 where fracture locations tend to be clustered along regions of the well trajectories rather than randomly and homogenously distributed. This is also true of the domain models themselves Fig. 3 and verified by variogram in App. 2A. This poses a problem for standard population statistical methods such as the Fisher test as individual data points within a set are not independent.

To emphasise the point, App. 1 shows a series of random indicator fields (random field => segmented => domains) with different degrees of autocorrelation (β) generated for the Top Chalk and analysed in terms of the well data. All these stochastic domain models had uniform domain population distributions. The sampled data (well locations) shows that despite being purely stochastic in nature, and having a uniform domain population distribution, as we increase the autocorrelation the expected uniform distribution of the modelled data is lost (App. 1B). This highlights the redundancy of standard statistical approaches (e.g. Fisher test) and the need for a baseline model. Appendix 1C presents the *E* data for each of these stochastic

domain models. These show that, irrespective of beta, each model typically has two or more domains that show enrichment (*E>1*). Under a limited category system (as here: 5) and the sparse signal principle, a genuinely selective pattern would be expected to manifest as concentrated enrichment within a single category and depletion in the remainder.

As stated, the seismic domain models also have a strong autocorrelated pattern. This has serious implications for assessing the significance of any *E* which also assumes data point independence; this rules out the sole validation of our hypothesis using *E* alone. By adopting a baseline model, we can assess whether any relationship is driven by autocorrelation, or even chance, rather than genuine concordance. In short, the baseline model acts as our null hypothesis. Using an indicator random field (Fig. 4) as our base case we can assess whether any *E>1* (for a domain-fracture relationship) are statistical artefacts (e.g. due to autocorrelation) rather than real. Should these models outperform, or equal the performance of the AI seismic domain models, in terms of "Test Criteria" as outlined above, then the null hypothesis is accepted.

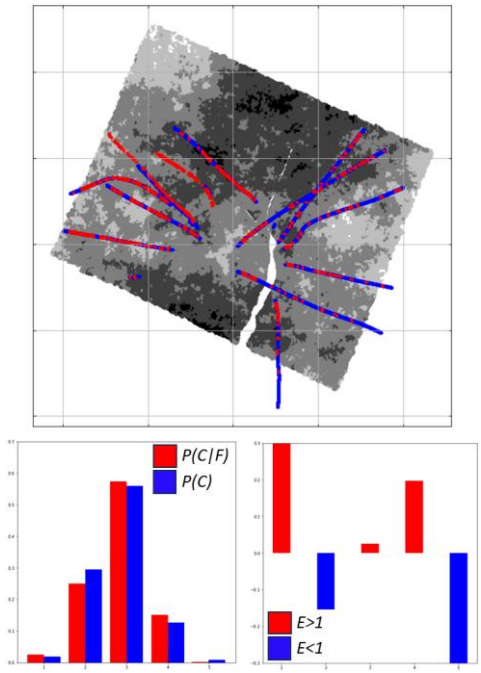


Figure 4. Example of a stochastic domain model with associated probability values for each domain (lower left) and computed enrichment (lower right).

Again, a number of baselines simulations will need to be run corresponding to each AI domain model. The baseline models were generated using a fractal noise generator followed by segmentation into categories thus producing random indicator fields (Deutsch, 2002) – hereafter denoted as stochastic domain models). These all had similar or only slightly different spatial autocorrelative character to the AI domain models (App. 4 and 5). It should be noted at this point, that the baseline cases were segmented using constant bin ranges; this created models with variable population distributions reflecting the actual AI domain models.

Sensitivity analysis

As stated in the Method section, estimates of enrichment can be highly sensitive to peculiarities in the dataset - these can be many and are often difficult to account for. Thankfully there are a number of standard global solutions for identifying, if not solving, problems of bias in our enrichment estimates such as Laplace smoothing (using an additive smoothing parameter: Ewan and Grant 2005).

A similar but different technique is employed here to assess the stability of *E* and thus identify results that are likely unreliable even if encouraging. Here we add a fixed proportion (0-1) to both numerator and denominator:

$$E(e) = \frac{P(C|F) + e}{P(C) + e} \qquad [3]$$

Like a Laplace smooth this helps regularise the result but also acts as a shrinkage factor - for a domain where *E>1*, as *e->1, E(e)* trends to neutrality. By assessing the rate of decay in *E(e)* we can assess the stability of an *E* result for a given domain in a given model. Furthermore, by normalising *E(e)* by *E(0)* we can compare models directly and refer to their curves (Fig. 5).

So, models that produce *E(e)* curves that decay quickly and have a large range (*E(0)-E(1)*) are probably not reliable, especially if the models

have more than one domain with *E>1*. Conversely, those models that have a single domain with *E>1*, stable *E(e)* curve and a small *E(e)* range, are likely capturing structural signal.

Results and discussion

The following section first uses the test criteria outlined above to assess the main character of the results and provides some discussion as to their significance. Afterwards more discussion is given to the significance of the *E* results and how much confidence we place in them.

Test criteria 1: single domain enrichment

The results from the analysis indicate that only the AI generated domain models produce *E* density profiles where there is a single enriched domain (App. 2C and 3C). The stochastic models typically have two or more domains where *E>1* (App. 4C and 5C).

The fact that AI models exhibit enrichment profiles with significantly greater categorical concentration than the stochastic models, which tend to produce diffuse enrichment density profiles, suggests that the AI models are capturing a more coherent and selective relationship between the predictor categories and fracture occurrence. Consequently, the AI-derived enrichments can be interpreted with greater confidence as reflecting structured geological signal rather than enrichment patterns arising from autocorrelated stochastic variability alone.

Test criteria 2: repeated over all simulations

Appendices 2 and 3 illustrates that the pattern described in answer to test criteria 1 above is consistent over all simulations.

Test criteria 3: spatial variation similar between simulations

Although more subjective in nature, the variograms and maps presented in Appendices 2 and 3 suggest some variation in the structure of the AI-derived domain models. However, this variation is substantially lower than that observed in the stochastic models against which they are compared. These differences indicate that the AI domain models converge on a common, geologically driven solution rather than reflecting internally generated stochastic variability.

Interpretation of the E values

It should be clear that *E* values for all of the AI domain models are moderate (1.1 to 1.3) with the lowest values typically found in the low-resolution versions (App. 2C and 3C). Despite being modest the enrichment does only occur in a single domain and consistently so. The difference between the higher and lower resolutions is to be expected, since running at a lower resolution is essentially a smoothing process that leads to blurring between fields that is likely compounded by possible aliasing effects. But the effect here is relatively small.

In contrast, the stochastic domain models can have very high enrichment ratios spread over many domains. These higher enrichment ratios are likely due to the more homogenous spread of *P(C)* and *P(C|F)* over all domains (as explained above). This creates *E* values that are not as stable as those for the AI models as the sensitivity analysis shows (Fig. 5) - it is clear that the enrichment figures for the stochastic domain models produce highly unstable *E(e)* curves (*E(0)-E(1)* are very large and variable).

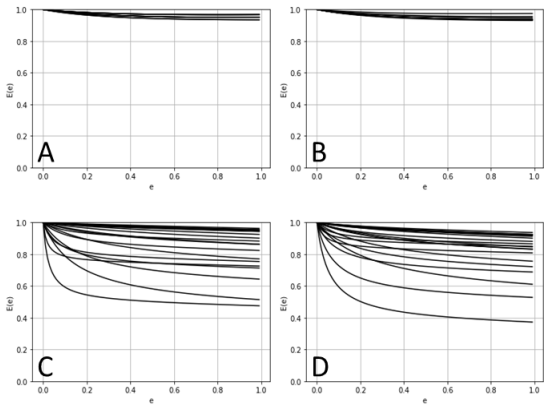


Figure 5. Plots showing *E(e)* against *e*. A) For low resolution AI domain model. B) For high resolution AI domain model. C) For β=2.5 stochastic domain model. D) For β=3.0 stochastic domain model.

This is to be expected, as the *E>1* values are supported by lower probability values, making them inherently more sensitive to small fluctuations and local variations (App. 4 and 5).

Conclusion and summary

In summary then, it seems reasonable to reject the null case and accept that the AI seismic domain method, as implemented in T7, does map to this single geological characteristic in this geological scenario. The modest *E* scores are likely due to other secondary factors that influence lateral changes in the nature of the seismic signature so that the relationship is not perfect – e.g. smaller fractures may play their part. Furthermore, it should also be noted that we are dealing with a closed system, with a known and attainable perfect enrichment score of 1.6. Therefore, *E* values ranging from 1.1 to 1.3 will carry more significance when compared to enrichment scores attained in other subject areas: $E > 1e^4$ in mineral deposits of minor elements.

Finally, the parameter evaluated here conditions the probability of a seismic domain occurring in the presence of fractures (*i.e. P(C|F)* in Eq. 2). For practical application, however, this conditional relationship must be reversed for predictive purposes, such that the probability of fracturing is conditioned on the domain value at a given location: P(F|C). It can be shown (see App.6) that:

$$\frac{P(F|C)}{P(F)} = \frac{P(C|F)}{P(C)} \quad [4]$$

In essence, validating the domain–fracture enrichment relationship can also provide a means of predicting relative enrichment in fracture abundance, in this case by approximately 10–30% for certain domains. While this may appear modest, such an increase can be highly significant in reservoirs where permeability is strongly controlled by fracture density.

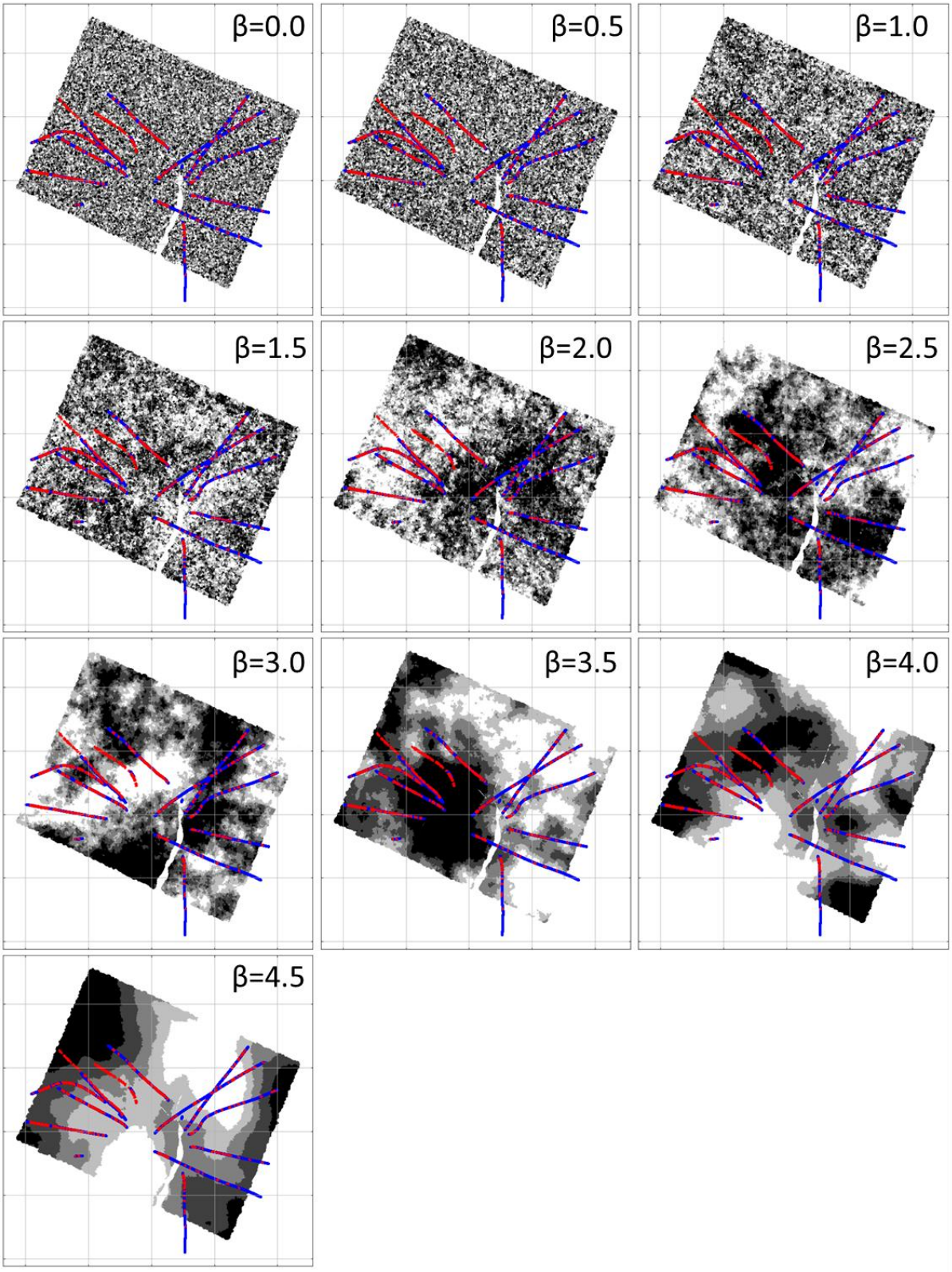
References:

Deutsch, C.V. (2002). Geostatistical reservoir modelling, Oxford University Press, Oxford.

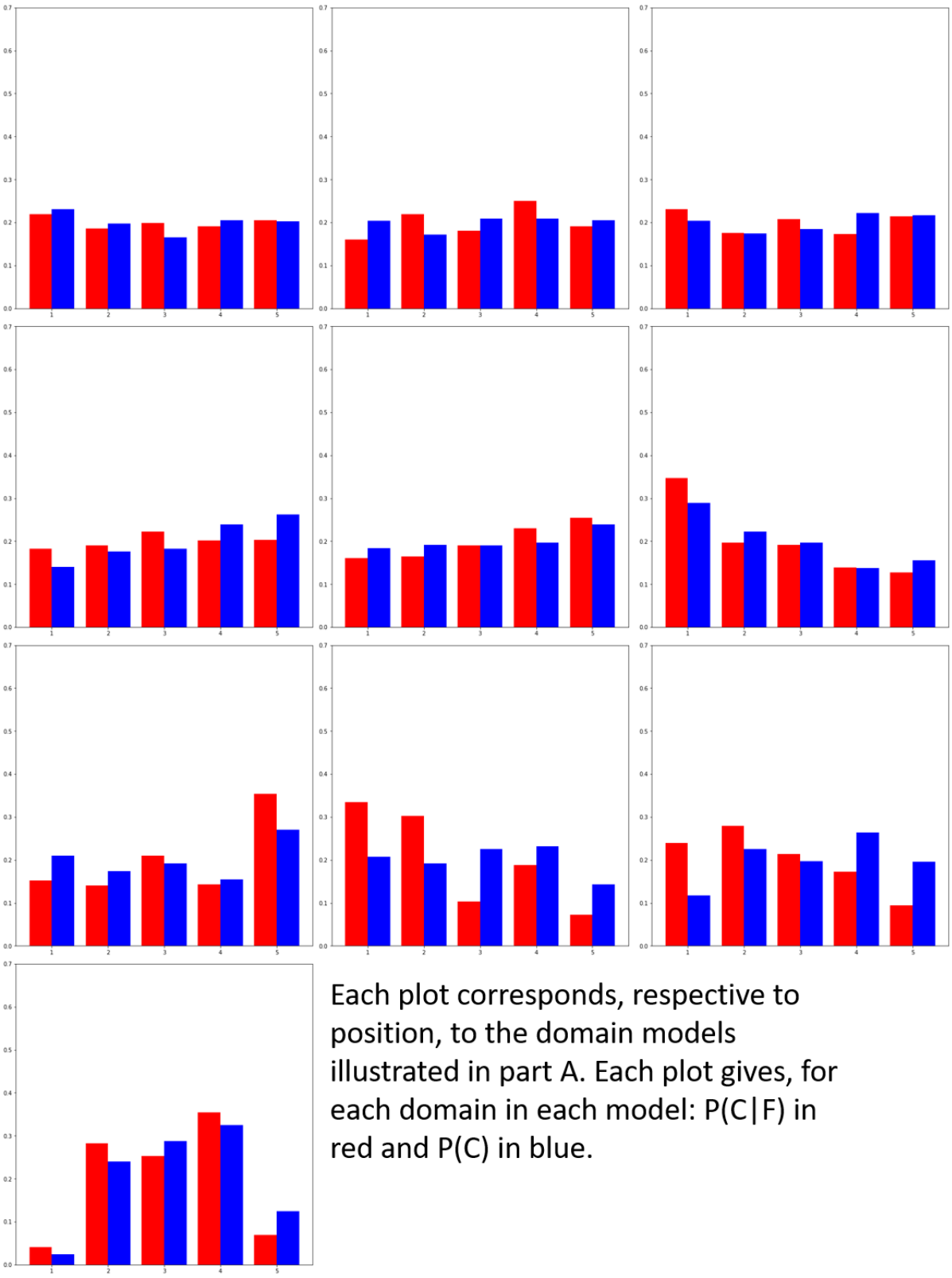
Dillon, C. (2026). AI seismic domains: a new tool to help find hidden signal contrasts within amplitude volumes, BGL Technical Note.

Ewens W.J and Grant G. (2005), Statistical Methods in Bioinformatics: An introduction, New York City, Springer-Verlag New York Inc.

Freeman B., Quinn D.F., Dillon C.G., Arnhild M. and Jaarsma B. (2015), Predicting subseismic fracture density and orientation in the Gorm Field, Danish North Sea *in* Richardson, N. J., Rippington, S. J., Wilson, R. W. and Bond, C. E (*eds*), Industrial Structural Geology: Principles, Techniques and Integration. Geological Society, London, Special Publication.

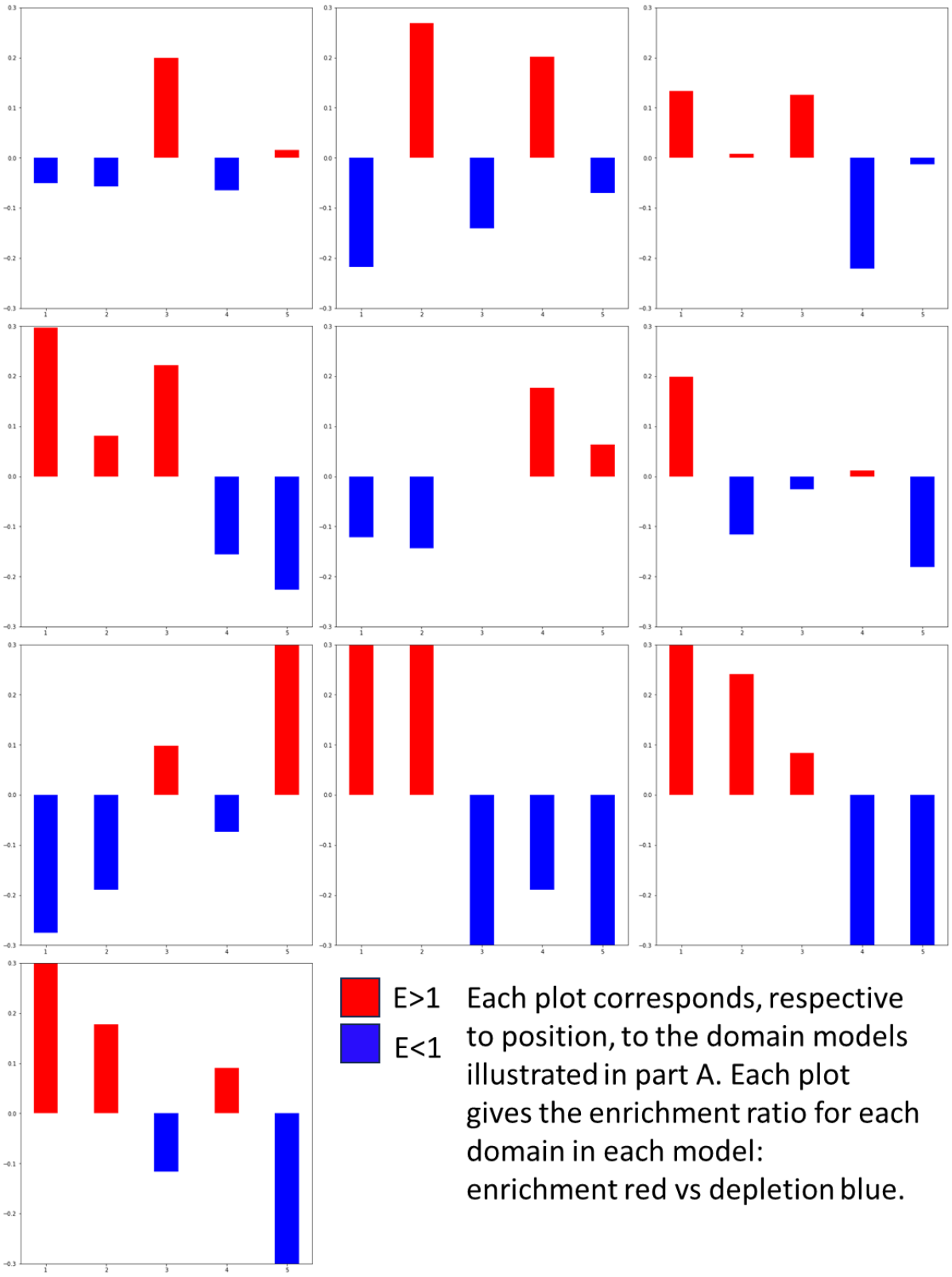


Appendix 1A: Random indictor fields generated with different beta factors producing a range of domain models with white to red noise signals. All random fields are segmented to preserve uniform category populations.

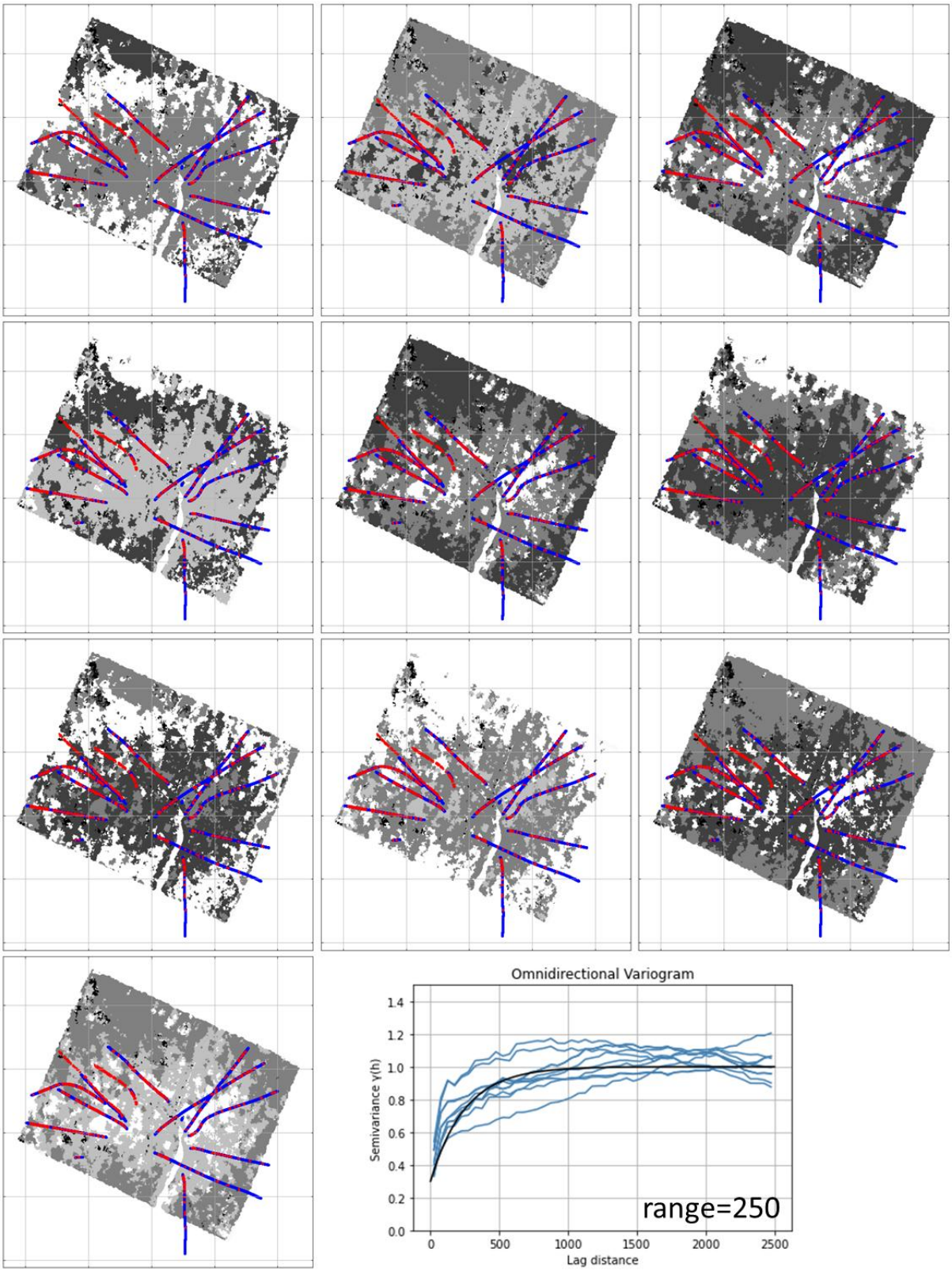


Each plot corresponds, respective to position, to the domain models illustrated in part A. Each plot gives, for each domain in each model: P(C|F) in red and P(C) in blue.

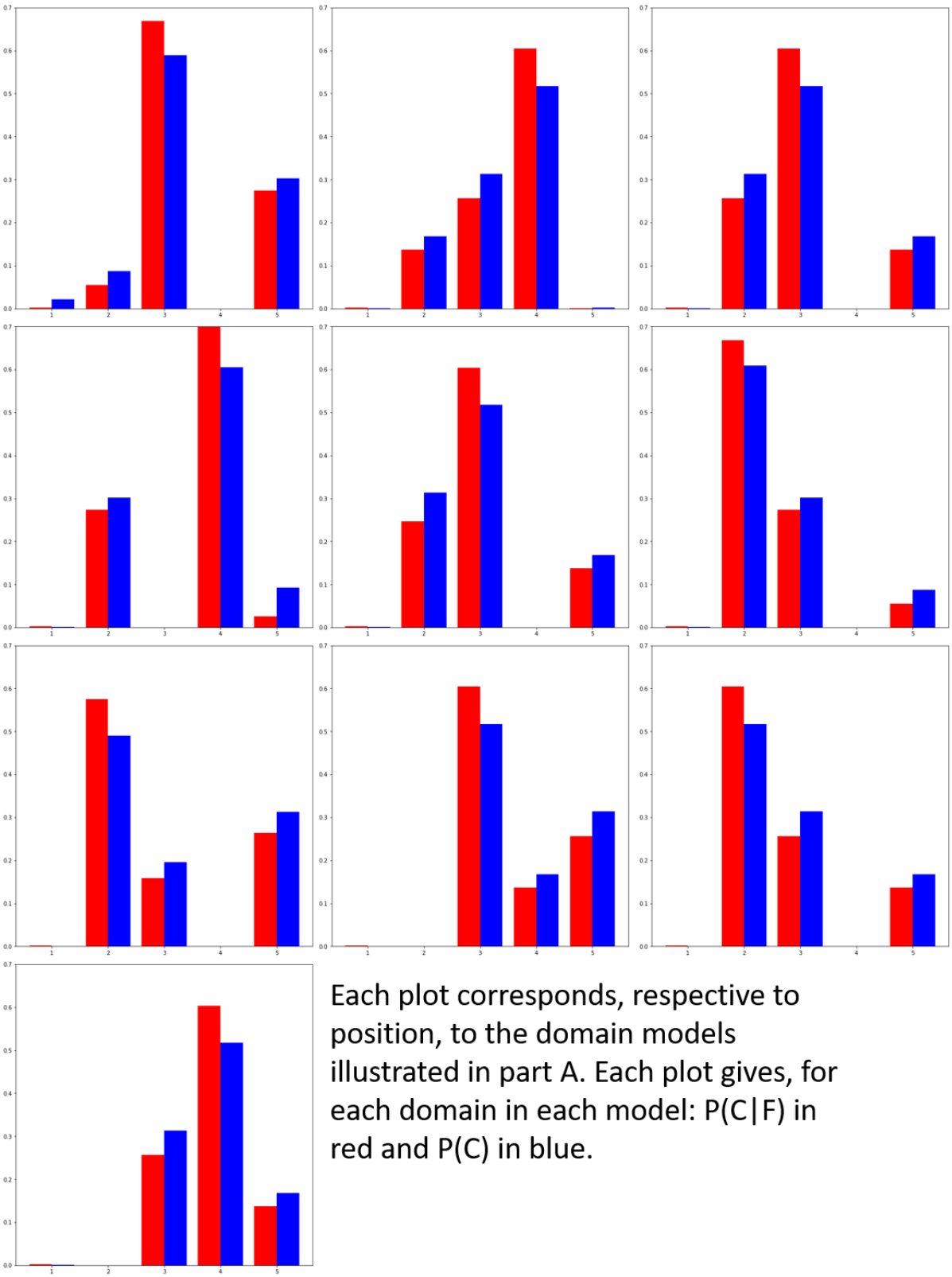
Appendix 1B: Distribution of *P(C|A)* and *P(C)* computed for AI stochastic models presented in App. 1A.



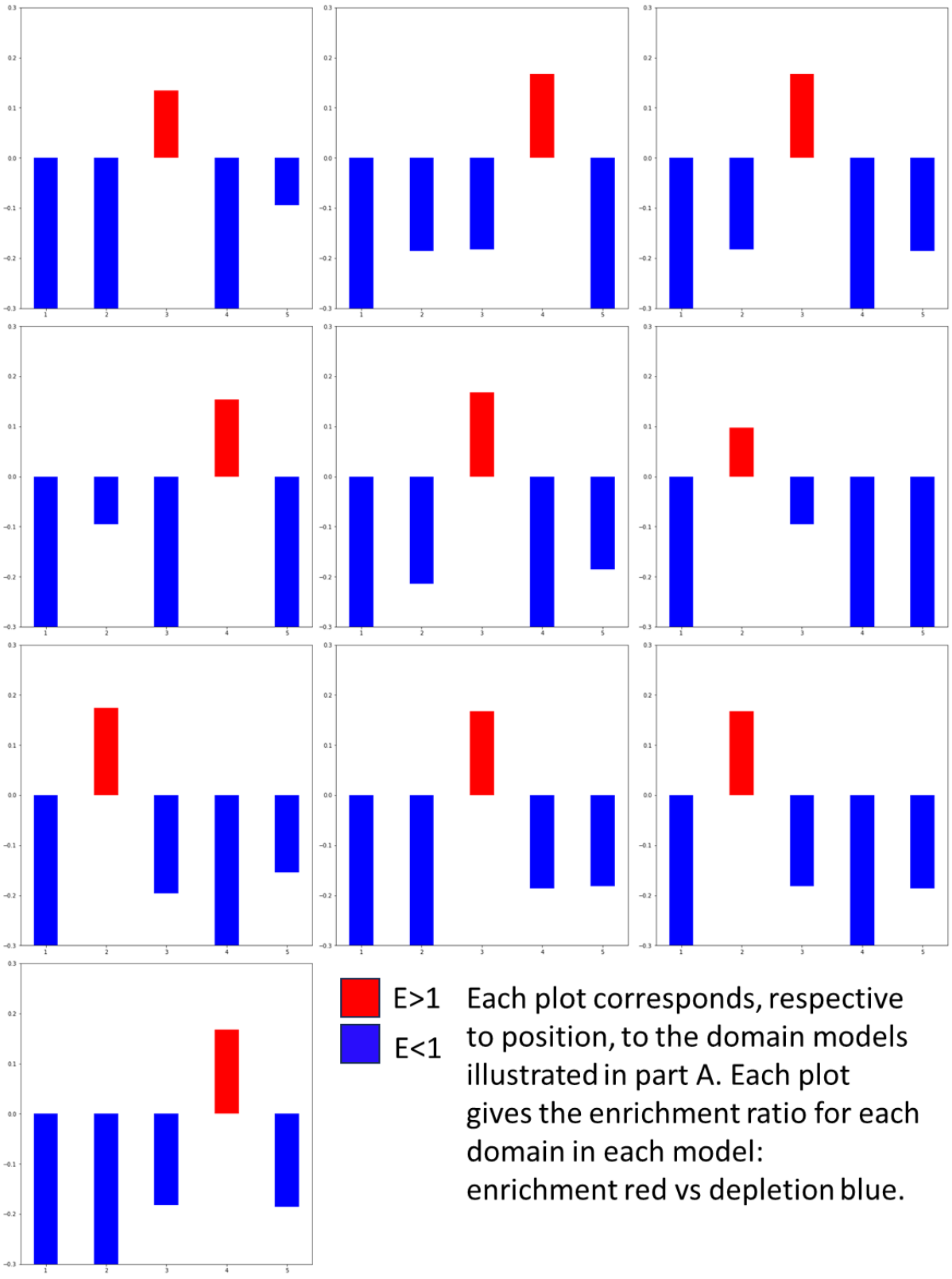
Appendix 1C: Computed *E* for stochastic domain models presented in App. 1A.



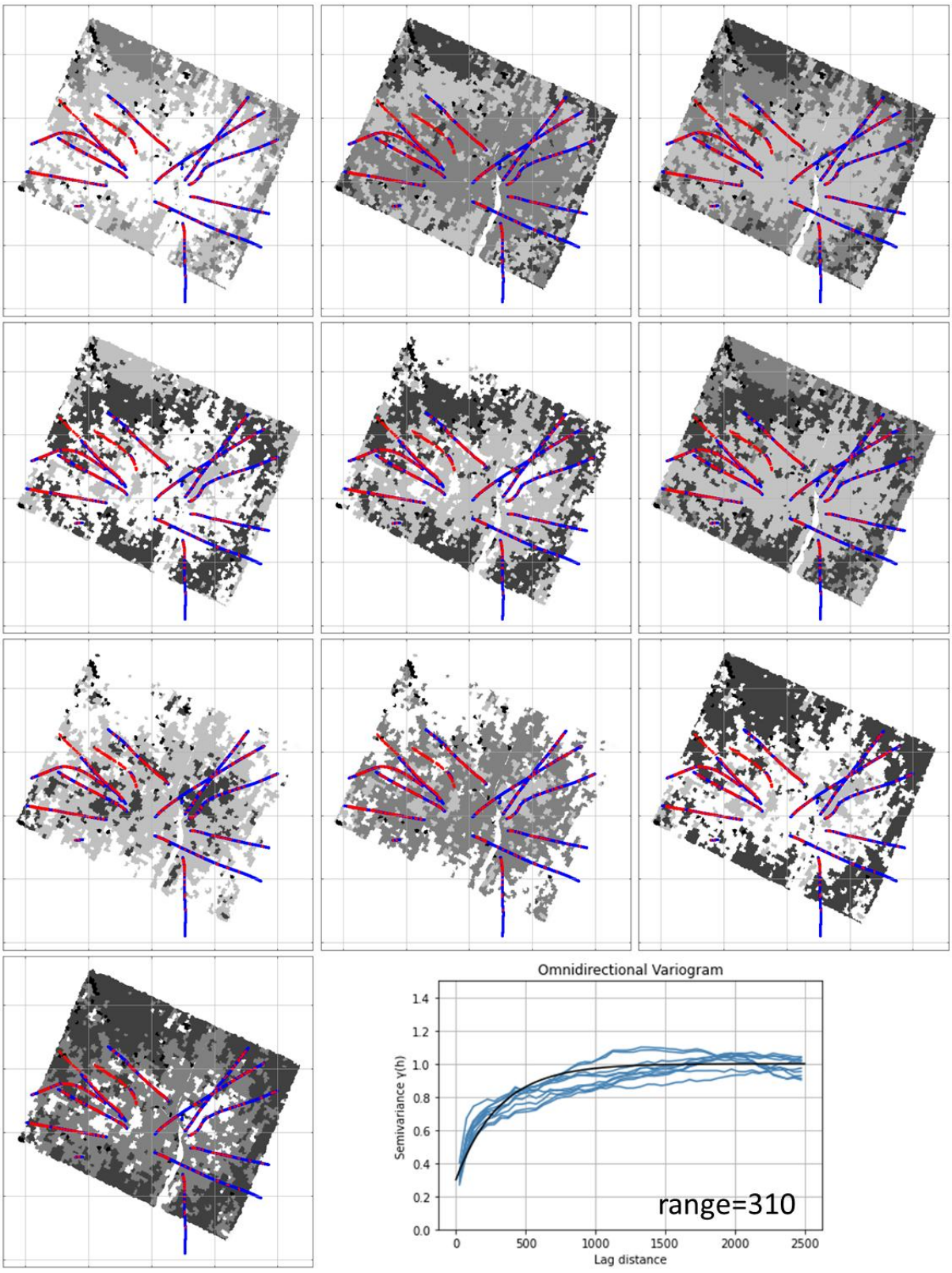
Appendix 2A: Series of AI seismic domain models produced at a high resolution with inset variogram (range is for exponential model fitted to experimental variograms produced for all models).



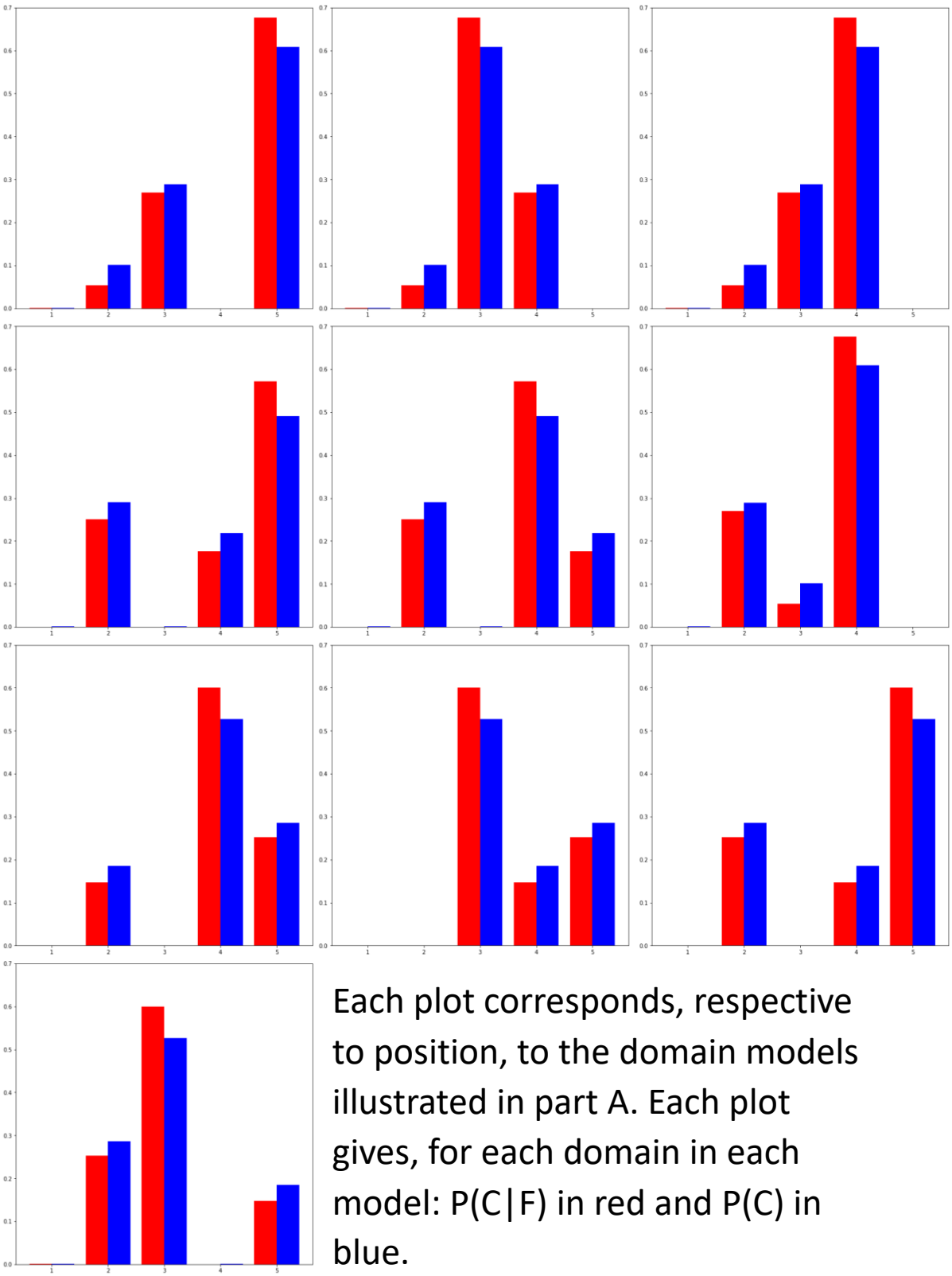
Appendix 2B: Distribution of *P(C|A)* and *P(C)* computed for AI domain models presented in App. 2A.



Appendix 2C: Computed *E* for AI domain models presented in App. 2A.

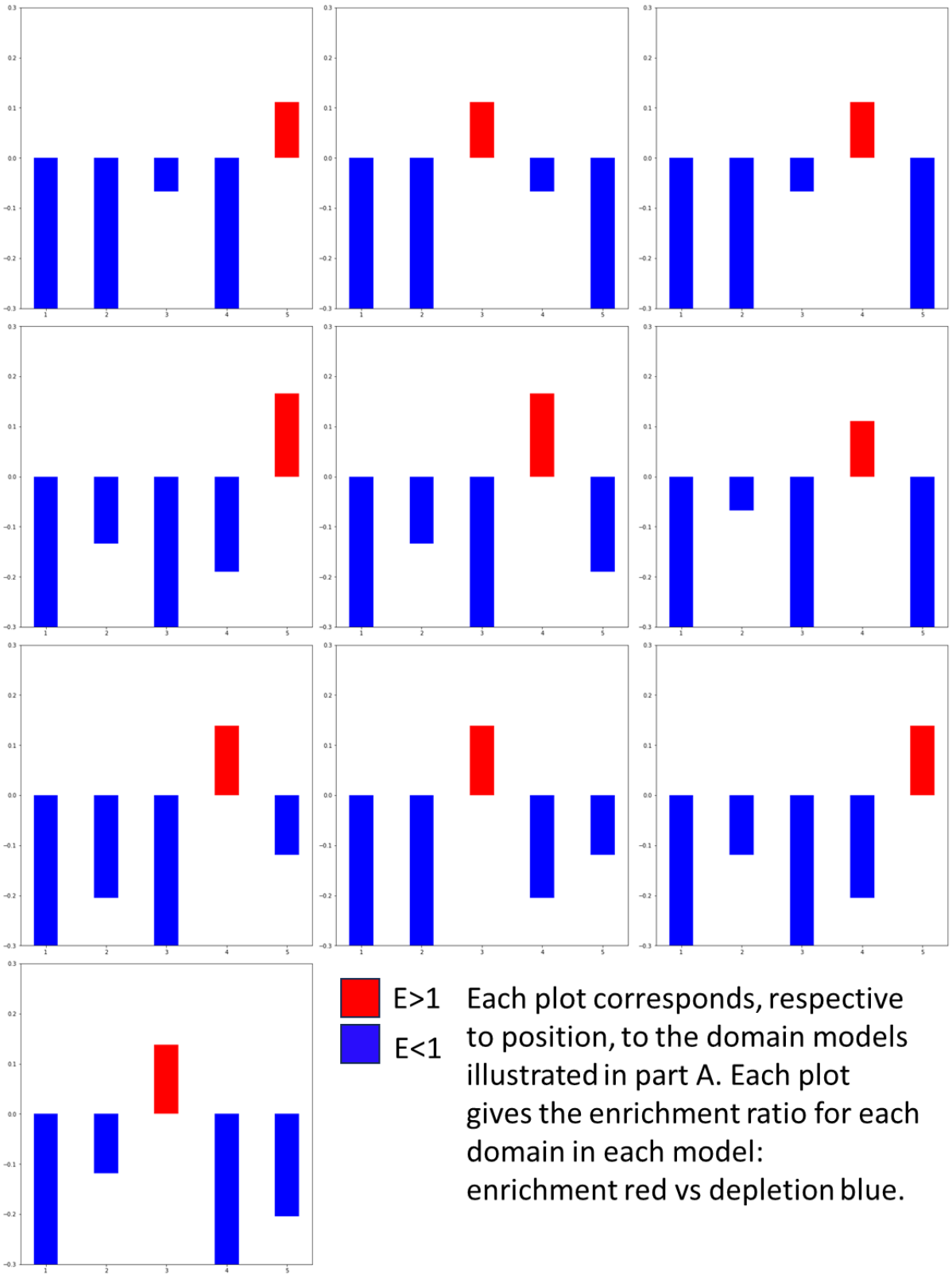


Appendix 3A: Series of AI seismic domain models produced at a low resolution with inset variogram (range is for exponential model fitted to experimental variograms produced for all models).

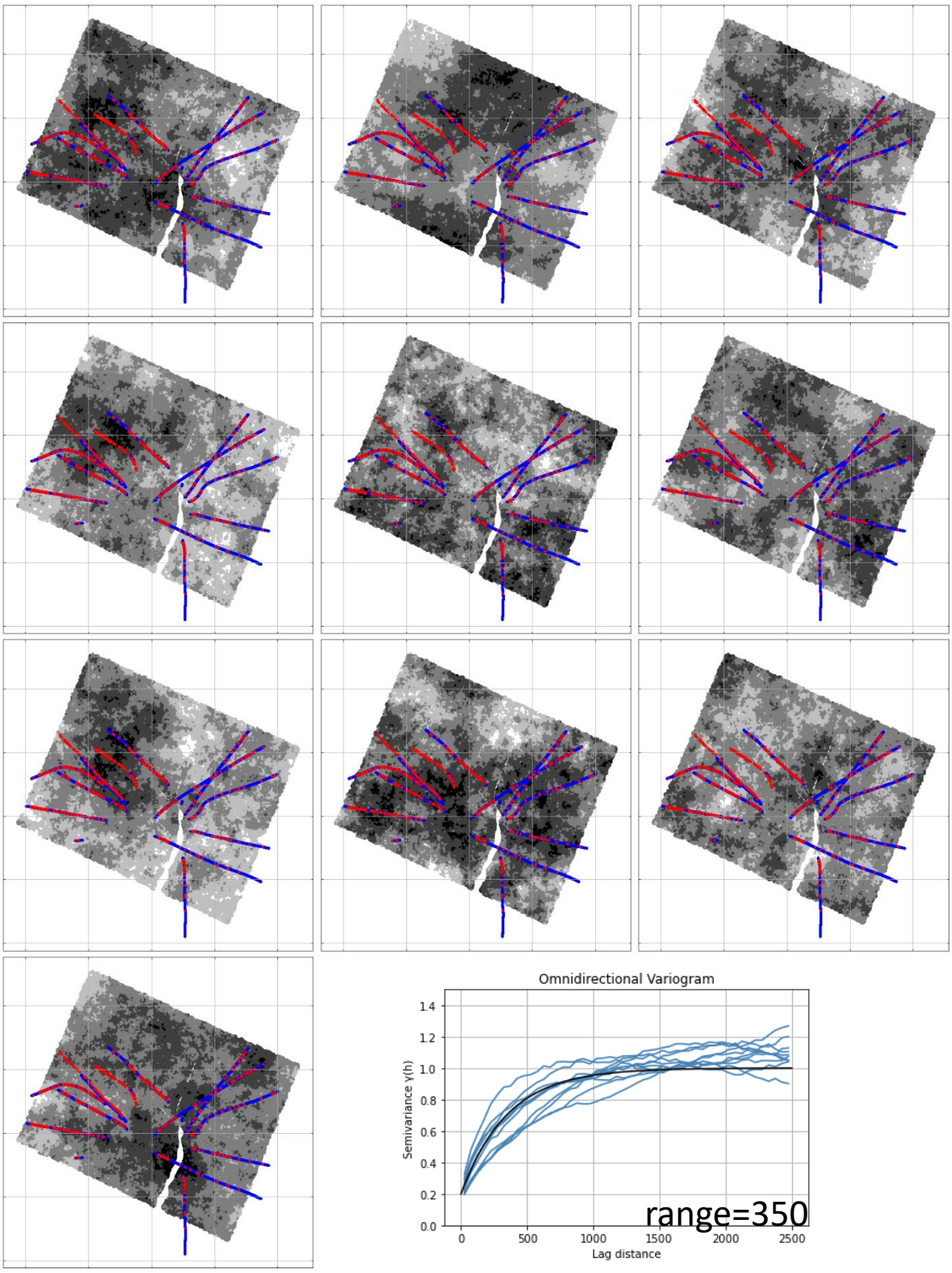


Each plot corresponds, respective to position, to the domain models illustrated in part A. Each plot gives, for each domain in each model: P(C|F) in red and P(C) in blue.

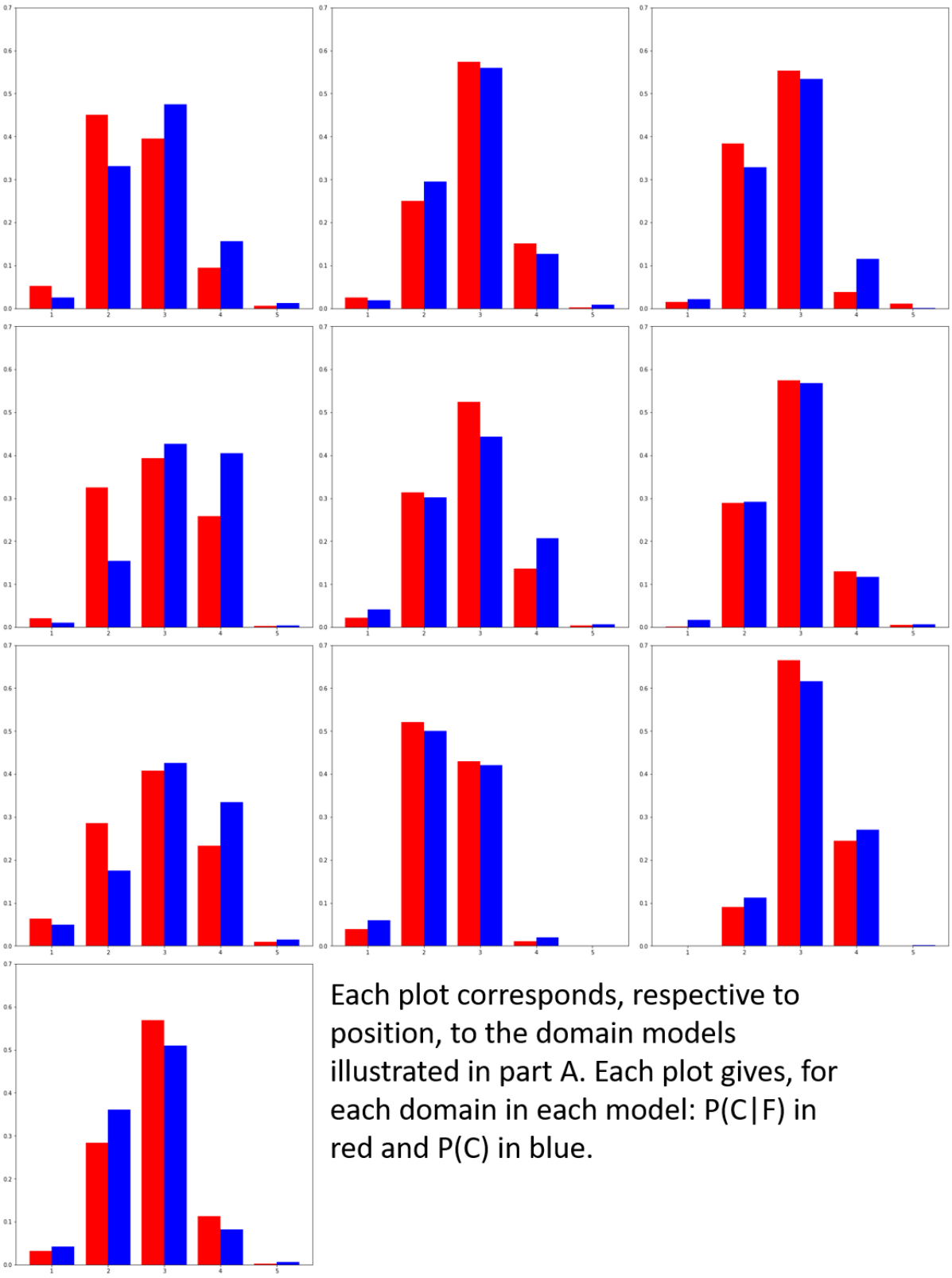
Appendix 3B: Distribution of *P(C|A)* and *P(C)* computed for AI domain models presented in App. 3A.



Appendix 3C: Computed *E* for AI domain models presented in App. 3A

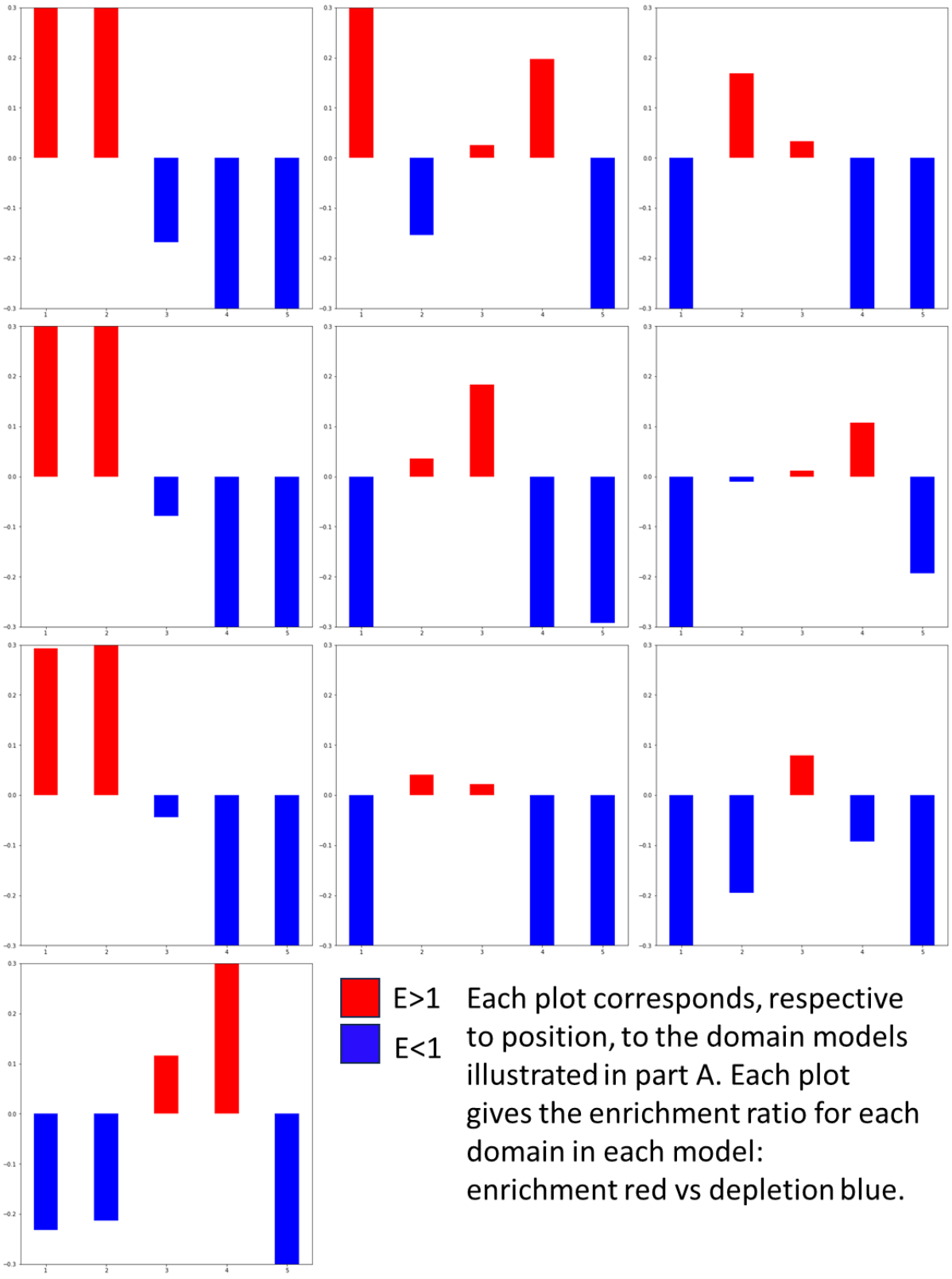


Appendix 4A: Random indictor fields generated with beta = 2.5 producing a range of domain models. All random fields are segmented to preserve uniform category ranges. Variogram exponential model with a range of 350 is defined for the corresponding experimental set.

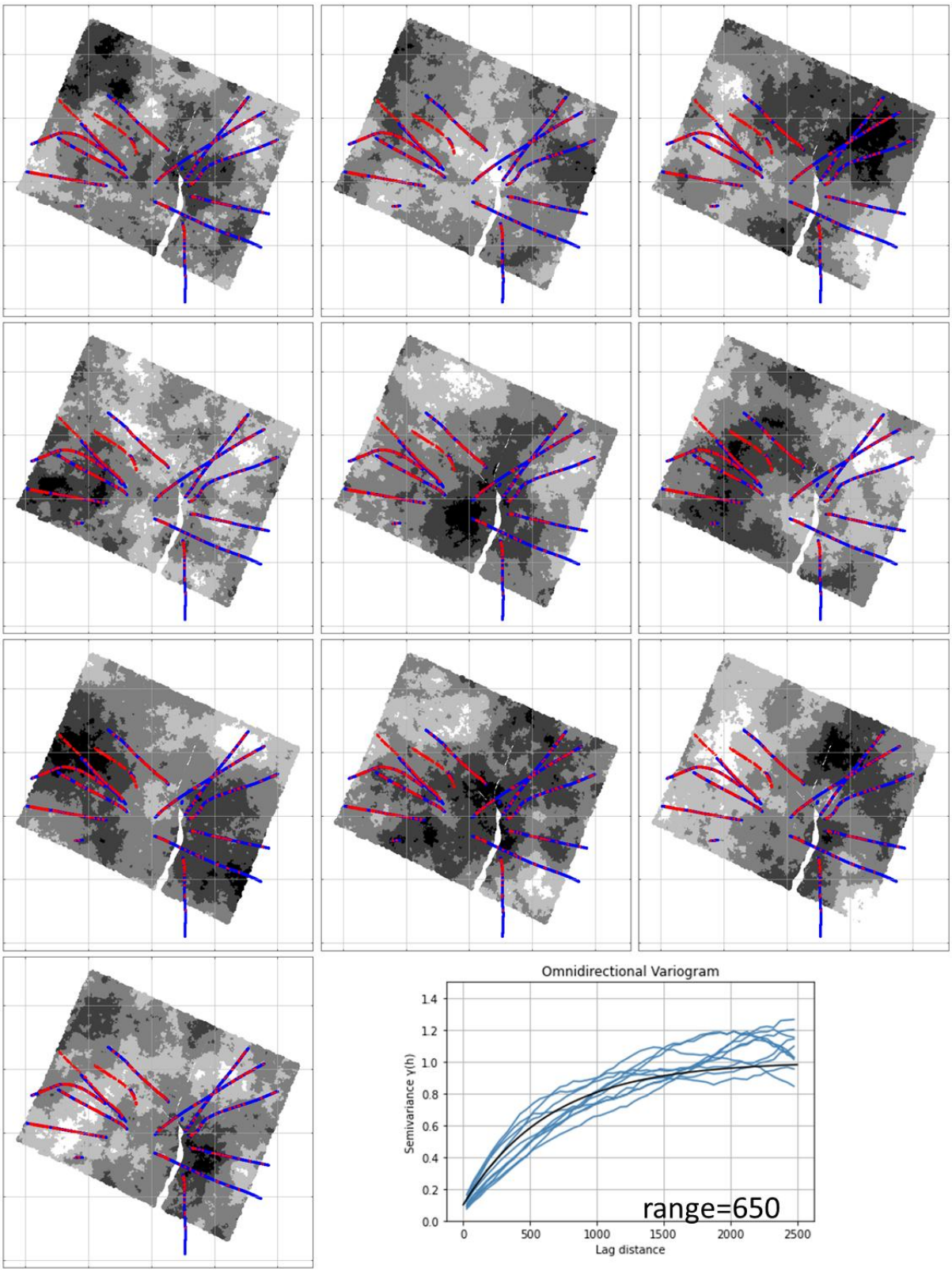


Each plot corresponds, respective to position, to the domain models illustrated in part A. Each plot gives, for each domain in each model: P(C|F) in red and P(C) in blue.

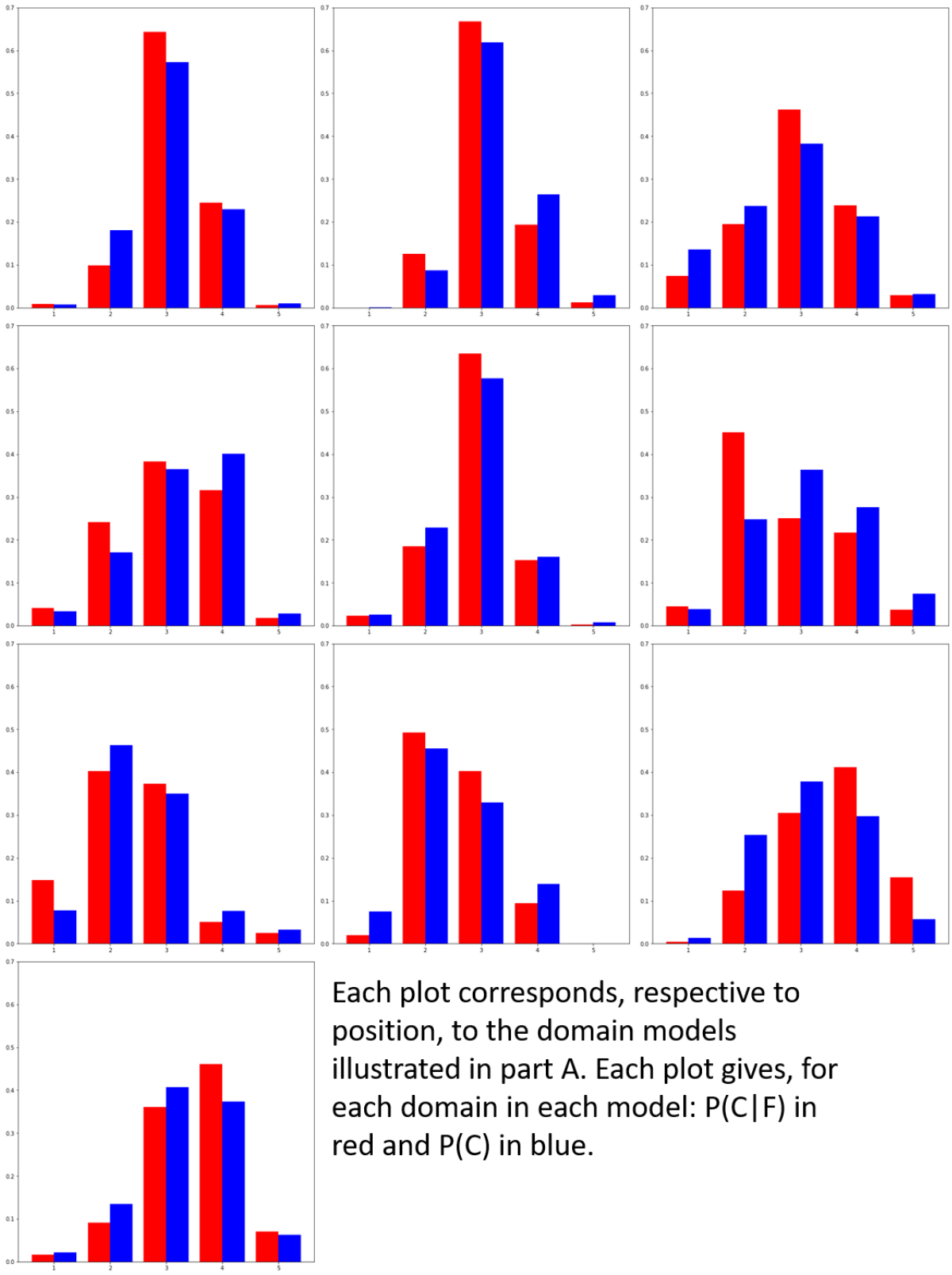
Appendix 4B: Distribution of *P(C|A)* and *P(C)* computed for AI stochastic models presented in App. 4A.



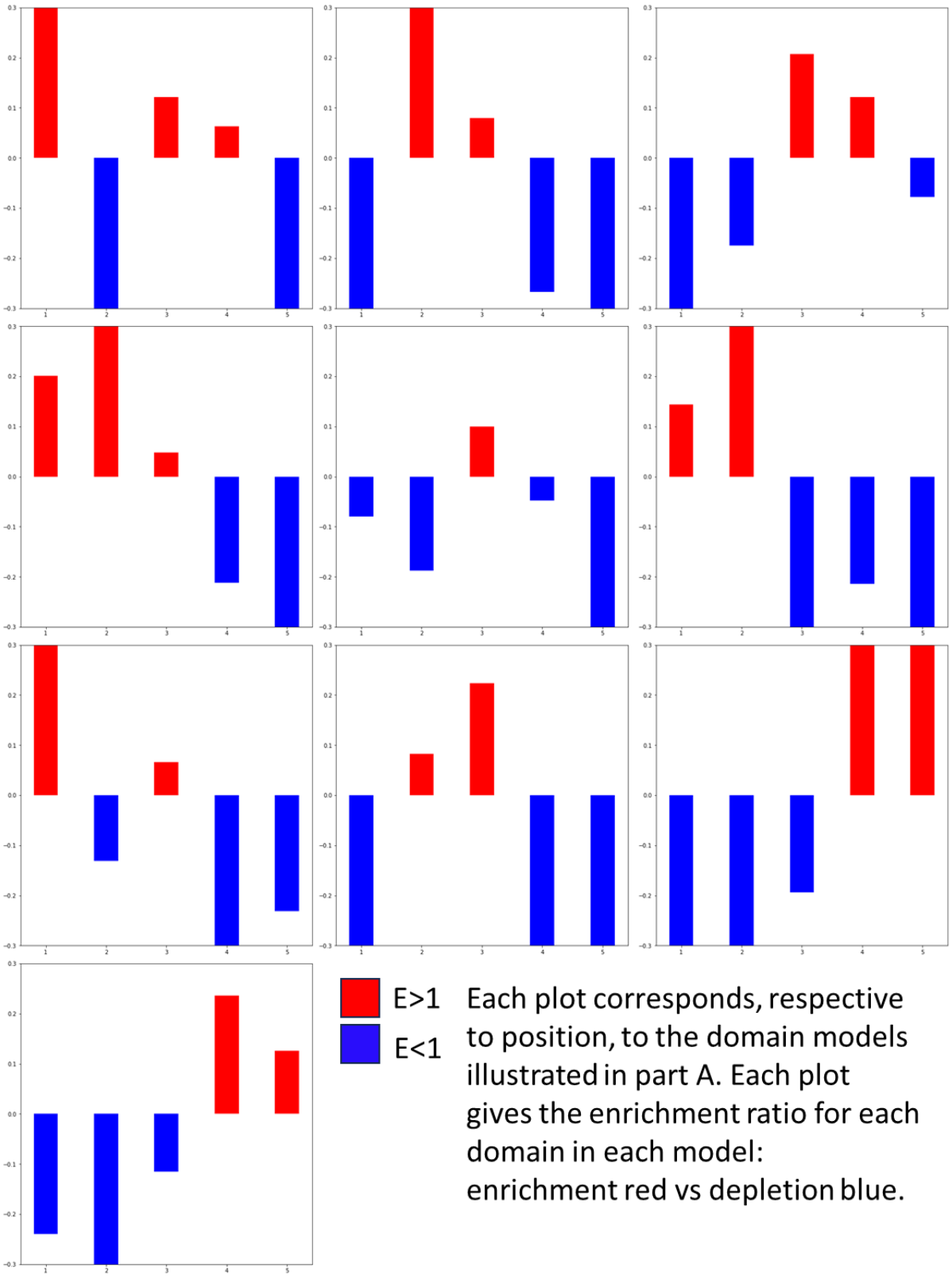
Appendix 4C: Computed *E* for stochastic domain models presented in App. 4A.



Appendix 5A: Random indictor fields generated with beta = 3.0 producing a range of domain models. All random fields are segmented to preserve uniform category ranges. Variogram exponential model with a range of 650 is defined for the corresponding experimental set.



Appendix 5B: Distribution of *P(C|A)* and *P(C)* computed for AI stochastic models presented in App. 5A.



Appendix 5C: Computed *E* for stochastic domain models presented in App. 5A.

Using the Bayes framework:

$$P(C_i|F) = \frac{P(F|C_i)\,P(C_i)}{P(F)}$$

Now divide both sides by $P(C_i)$:

$$\frac{P(C_i|F)}{P(C_i)} = \frac{P(F|C_i)\,P(C_i)}{P(F)\,P(C_i)}$$

Now cancelling the passive terms:

$$\frac{P(C_i|F)}{P(C_i)} = \frac{P(F|C_i)}{P(F)}$$

In general terms:

$$\frac{P(C|F)}{P(C)} = \frac{P(F|C)}{P(F)}$$

Appendix 6: Equivalence between the complementary conditional arrangements for the enrichment ratio.