Dear Authors,

I believe your work has potential. However, there are several issues that need to be discussed in more detail and with greater clarity. I have included my comments in the text file.

Here are some main concerns:

How do you choose the parameters for the objective function (Equation 14), and what criteria do you use to select these parameters?

How do you choose the parameters for clustering, and what criteria do you use to select them?

Could you please provide results of the inversion with different noise levels and different starting models, and compare these results with those obtained using Hampson-Russell?

Kind Regards

==

Dear Reviewer,

Thank you for your thorough review and insightful comments on our manuscript. We appreciate the time and effort you have invested in providing detailed feedback. Below, we address each of your observations and comments.

We hope the following information and explanations meet your expectations.

1. How do you choose the parameters for the objective function (Equation 14), and what criteria do you use to select these parameters?

Answer:
As stated in Equation 14, "Each objective term is weighted using a denominator, $\theta(\cdot)$, to normalize the contribution of the terms." The specific forms of θ are given by:

$$\theta(d) = \frac{norm(\boldsymbol{d}.2)}{length(\boldsymbol{d})}. \quad \theta(\boldsymbol{x}) = \frac{norm(\boldsymbol{x}.2)}{length(\boldsymbol{x})}. \quad \theta(c) = \frac{norm(\boldsymbol{u}^q \hat{O}.2)}{length(\boldsymbol{x}) \times C}$$

Additionally, lines 219-222 state, "Weighting factors balance the misfit or data fidelity term and the other terms (such that ($w\_d + w\_x + w\_s + w\_C = 1$)). This eliminates the need for high-dimensional L-curve, L-surface, or Generalized Cross-Validation methods to find optimum regularization parameters."

The weighting factors (w_d), (w_x), (w_s), and (w_C) can have equal contributions, or the operator can adjust them based on initial information about the model.

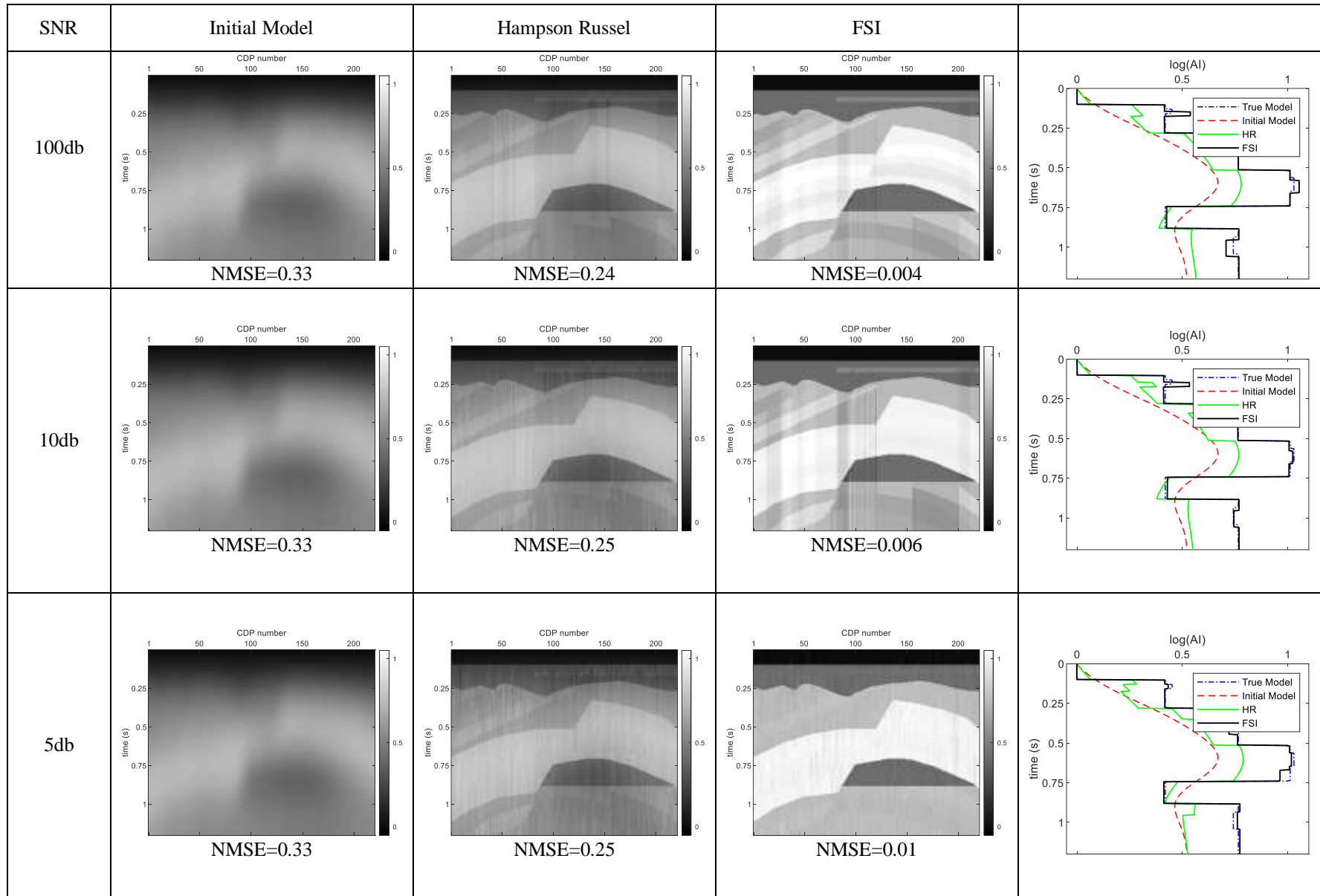- -

2. How do you choose the parameters for clustering, and what criteria do you use to select them?
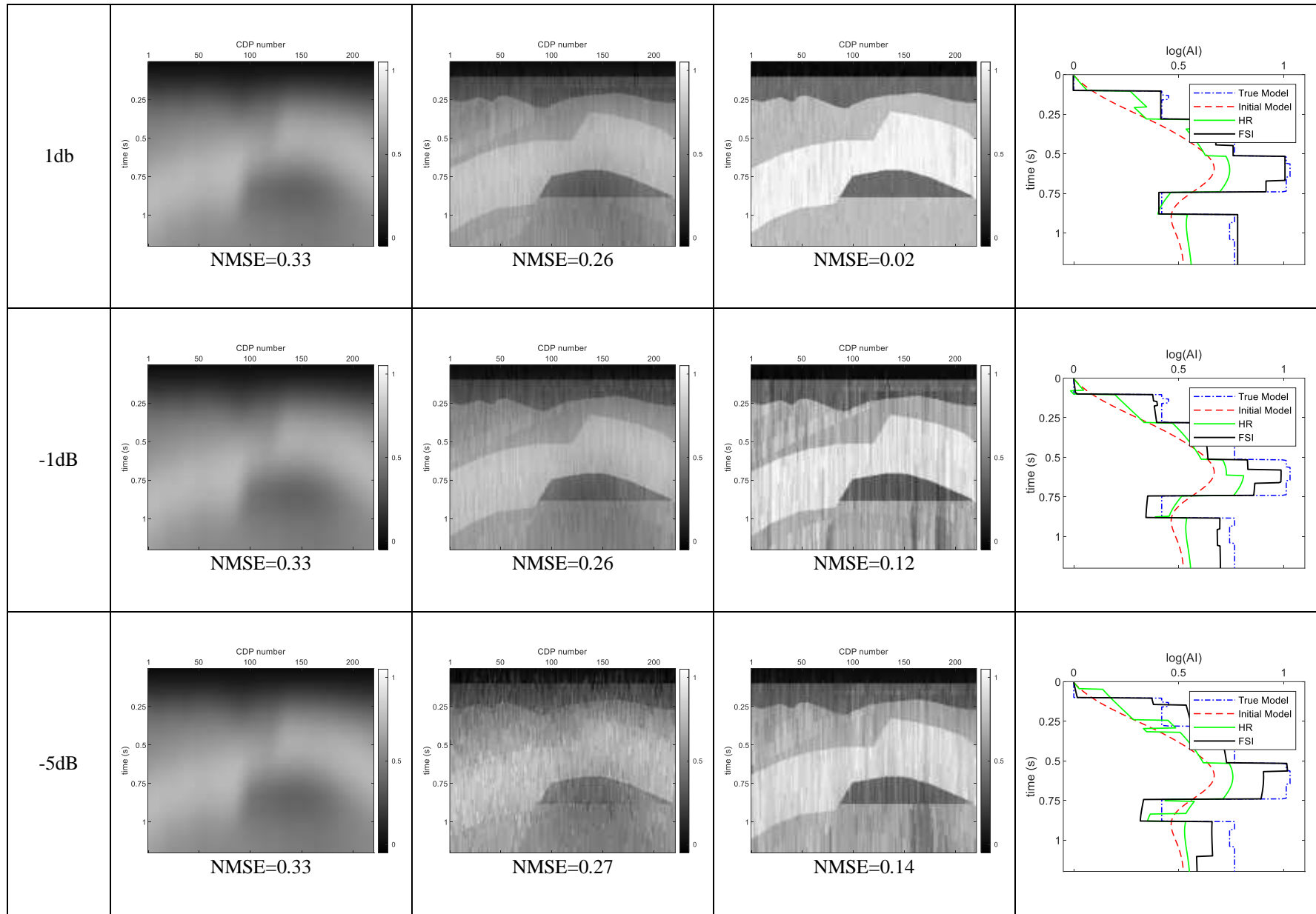
The fuzziness parameter ((q)) controls the clustering result. As stated in Figure 8 and lines 322-323, "the lower fuzziness tends the fuzzy clustering toward hard clustering and decreases the resolution of the membership sections."
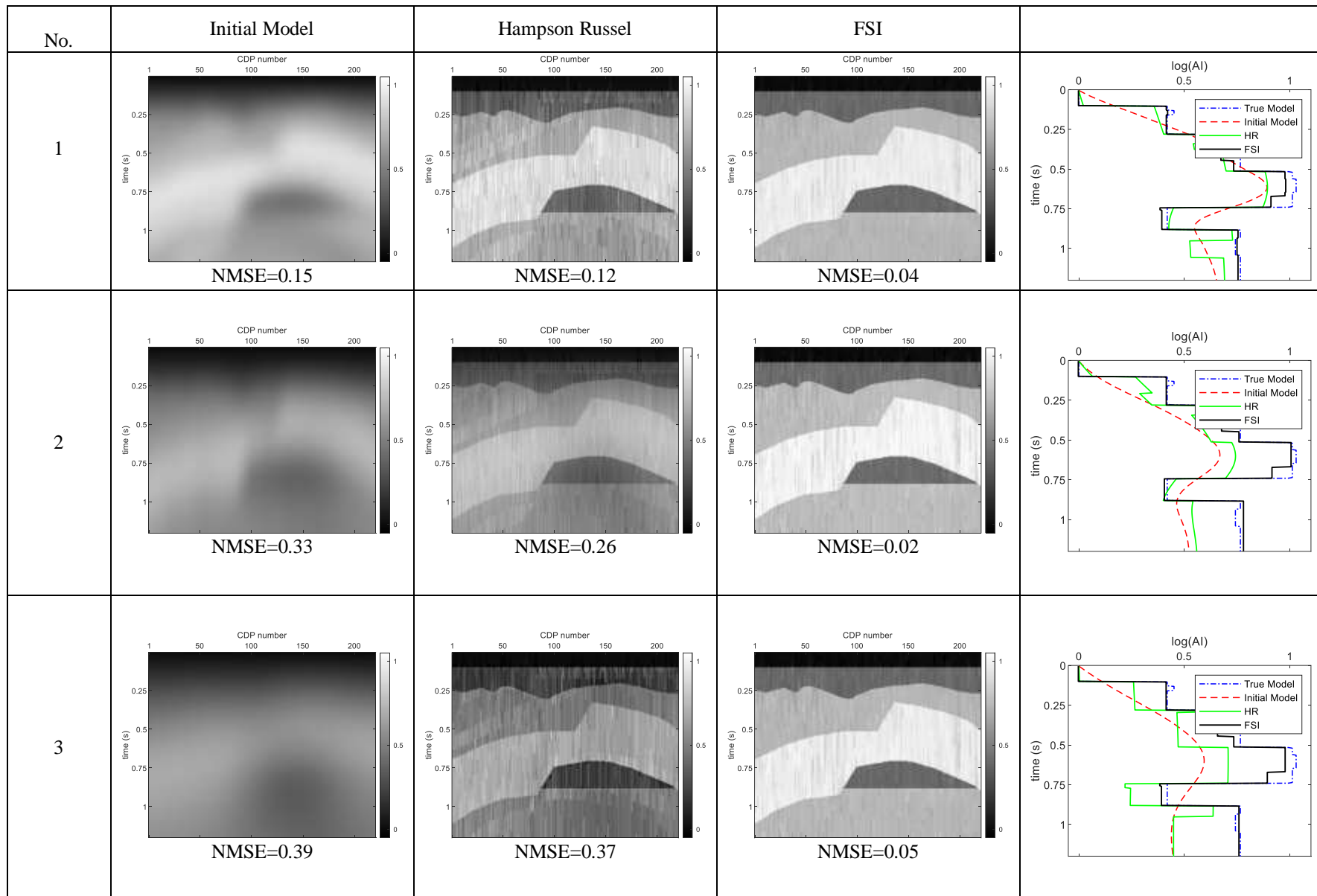
By default, (q = 2). However, to create more distinct clustering membership sections (e.g., to distinguish main sequences), one could decrease (q). Conversely, to create less distinct clustering memberships (e.g., to distinguish within-layer lithologies), one could increase (q). Furthermore, the role of (q) in the interpretation of the results could be an interesting topic for further research.

- -

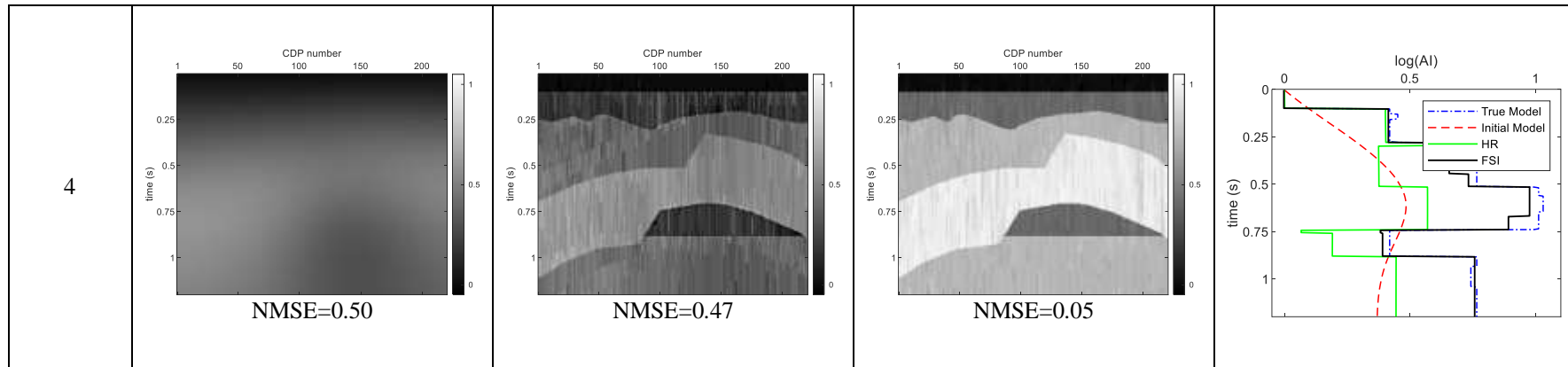3. Could you please provide results of the inversion with different noise levels and different starting models, and compare these results with those obtained using Hampson-Russell?

To maintain the simplicity of the manuscript and avoid presenting extra information and results, we prefer not to add other examples to the manuscript. However, here we present the requested results for different noise levels and different initial models:

SNR	Initial Model	Hampson Russel	FSI	
100db	NMSE=0.33	NMSE=0.24	NMSE=0.004	
10db	NMSE=0.33	NMSE=0.25	NMSE=0.006	
5db	NMSE=0.33	NMSE=0.25	NMSE=0.01	

1db	NMSE=0.33	NMSE=0.26	NMSE=0.02
-1dB	NMSE=0.33	NMSE=0.26	NMSE=0.12
-5dB	NMSE=0.33	NMSE=0.27	NMSE=0.14

No.	Initial Model	Hampson Russel	FSI	
1	NMSE=0.15	NMSE=0.12	NMSE=0.04	
2	NMSE=0.33	NMSE=0.26	NMSE=0.02	
3	NMSE=0.39	NMSE=0.37	NMSE=0.05	

4	NMSE=0.50	NMSE=0.47	NMSE=0.05	

Please note that for the inversion process using the less accurate initial model, we have reduced the weight of the initial model constraint (w_m)

- -

Kind regards