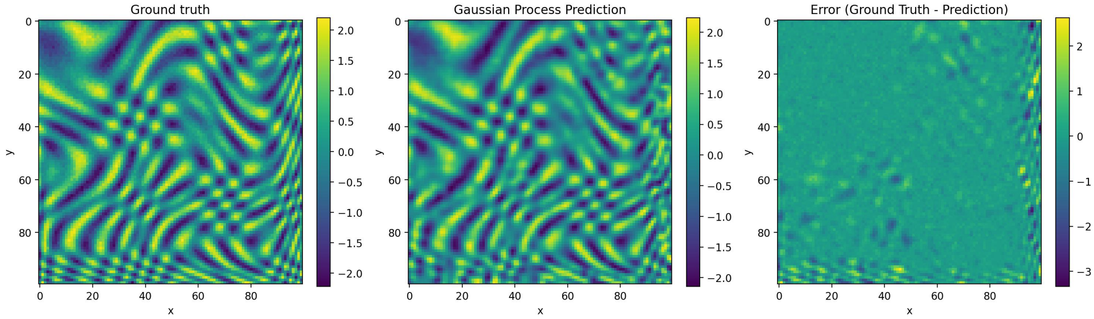
[1] the best length scale, kernel variance, and noise variance.

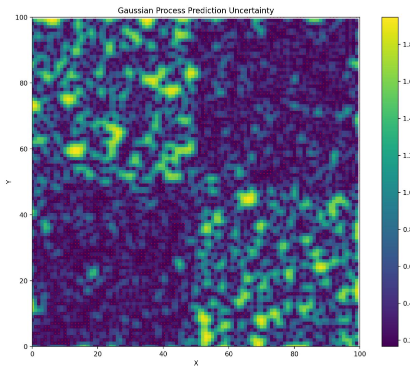
Define a reasonable initial value
l2_init = [0.5, 1, 2] # length scale squared
s2_init = [0.5, 1, 2] # signal variance
e2_init = [0.01, 0.1] # noise variance

Best initial values: l2_init=0.5, s2_init=2, e2_init =0.01

Best optimized parameters: l2=5.55569658, s2=0.98374679,e2= 0.01538258

[2] Plot maps of the ground truth value, the GP estimation, the error, and the standard deviation of GP estimation.





[3] Describe how good is the GP estimation. Discuss whether there is any problem in the current GP model and what potential solutions are.

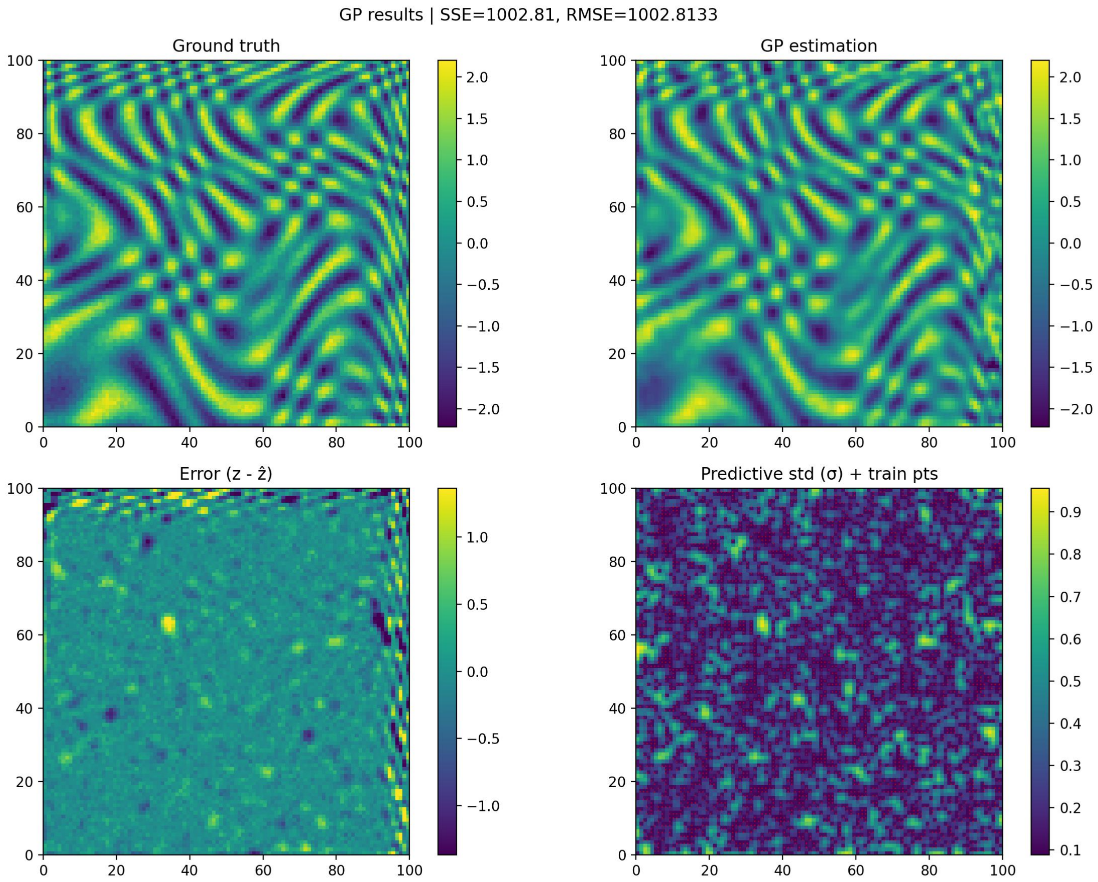
The Gaussian Process (GP) model successfully captures the large-scale spatial trends of the underlying field. However, several patterns in the prediction results indicate limitations that are inherent to both the chosen kernel (squared exponential white noise) and the distribution of training samples.

<u>Key observations from the estimation maps:</u>

- Higher errors occur at the boundaries, where the GP must extrapolate beyond the convex hull of training samples. This is expected, as GP prediction variance increases with distance from training data.

- The standard deviation map shows larger uncertainty in regions where training samples are sparsely distributed, indicating that the GP correctly represents epistemic uncertainty.

- When training samples are uniformly distributed, the GP estimation improves significantly, leading to a substantial reduction in the squared error (from 1399.89 to 1002.8132882045988, i.e., a reduction of 29%), demonstrating the sensitivity of GP performance to sample placement.



<u>Problems in the Current GP Model</u>

- ***Non-uniform training sample distribution:*** Sparse regions lead to large local interpolation errors. The GP model relies on spatial correlation; when samples cluster in certain areas, the model underrepresents other regions, resulting in over-smoothed predictions or high extrapolation bias.GP

- ***Boundary effect:*** Near the grid edges, the GP can only interpolate in one direction (inward), causing higher uncertainty and error. This manifests as low confidence in the predictions and poor fit relative to the center regions.
- ***Kernel limitations:*** The squared exponential kernel assumes stationarity and smoothness across the entire domain. However, the ground truth spatial field exhibits locally varying frequencies and non-stationary behavior, which cannot be fully captured by a single global length scale.

<u>Potential solutions</u>

Problem	***Impact***	***Proposed Solution***
Non-uniform training samples	High error in sparse regions	Use evenly distributed sampling or actively select informative samples (active learning)

Problem	*Impact*	*Proposed Solution*
Boundary effect	Increased extrapolation error	Introduce virtual boundary points or switch to kernels with better extrapolation (e.g., Matérn with finite differentiability)
Stationary kernel assumption	Over-smoothing in regions with high variability	Replace RBF with Matérn or use composite kernels (long-scale + short-scale components)