Mantik Journal, 9 (4) (2026), ISSN 2685-4236 (Online)
Published by:Institute of Computer Science (IOCS)
Jurnal Mantik
Journal homepage: www.iocscience.org/ejournal/index.php/mantik

Uncertainty and stability analysis of data-driven inversion using support vector regression

I Wayan Pio Pratama
Information Technology, Politeknik eLBajo Commodus, Indonesia

ARTICLE INFO

Article history:
Received Dec 24, 2025
Revised Jan 07, 2026
Accepted Jan 19, 2026

Keywords:
Inverse problems;
Learning-based inversion;
Support Vector Regression;
Uncertainty propagation;
Monte Carlo simulation.

ABSTRACT

This study examines learning-based inversion from the perspective of inverse problem theory, emphasizing uncertainty propagation, conditioning, and identifiability rather than pointwise prediction accuracy alone. Inverse estimation is formulated as a stochastic mapping in which observational noise is explicitly propagated through learned inverse models. A controlled one-dimensional nonlinear inverse problem with synthetic forward operators is used to isolate noise-induced instability and non-uniqueness effects. For an injective nonlinear forward mapping, Support Vector Regression (SVR) with a radial basis function kernel and linear regression are trained to approximate the inverse operator from noisy observations. Monte Carlo noise propagation is employed to estimate bias and variance of inverse predictions and to compare empirical uncertainty amplification with theoretical predictions based on local inverse conditioning. Although SVR significantly outperforms linear regression in terms of inverse accuracy, inverse uncertainty is shown to be primarily governed by the conditioning of the forward operator and modulated by model regularization. The analysis is extended to a non-injective forward operator to investigate identifiability loss. In this setting, both models collapse inherently multi-valued inverse mappings into unimodal and overconfident estimates, revealing implicit solution selection driven by data distribution and regularization. These findings demonstrate that low prediction error can be misleading in non-identifiable inverse problems. Methodologically, the results suggest that learning-based inverse models should be designed and evaluated using conditioning- and uncertainty-aware criteria rather than prediction accuracy alone, particularly in ill-posed or non-identifiable settings.

Corresponding Author:

I Wayan Pio Pratama,
Information Technology,
Politeknik eLBajo Commodus,
Labuan Bajo, Manggarai Barat, NTT, 86754, Indonesia.
Email: piopratama2@gmail.com

1. INTRODUCTION

Inverse problems are fundamental to many fields in science and engineering, where the objective is to infer unknown input parameters or system properties from observed outputs. Such problems arise in a wide range of applications, including geophysics, biomedical imaging, structural engineering, and remote sensing, where direct measurement of the parameters of interest is often infeasible (Gallet et al., 2022; Ji et al., 2022; Ogarko et al., 2021). In contrast to forward problems, which are typically well-posed and admit stable and unique solutions, inverse problems are frequently ill-posed, exhibiting non-uniqueness and instability. Consequently, small perturbations in observational data can lead to large deviations in inferred parameters, complicating reliable parameter estimation (Bangerth et al., 2025; Giraud et al., 2019; Ogarko et al., 2024; Zhai et al., 2025). Consequently, inverse problems are inherently sensitive to measurement noise, modeling inaccuracies, and incomplete data, necessitating robust solution strategies capable of handling uncertainty and instability.

Traditional approaches to inverse problems rely on analytical formulations or numerical optimization techniques that exploit explicitly defined physical relationships between variables. These methods often employ variational principles, regularization techniques, and iterative solvers to stabilize the inversion process (Engl et al., 2000a; Tikhonov & Arsenin, 1977). While classical inversion frameworks remain effective for certain problem classes, they can become unstable or computationally prohibitive when applied to nonlinear, large-scale, or high-dimensional systems, particularly when the governing physical model is only partially known (Chung & Gazzola, 2024; Latz, 2023). Such limitations are well documented in applied studies, especially in geophysical inversion, where non-uniqueness and overfitting are mitigated through the integration of geological constraints and uncertainty information (Giraud et al., 2019; Jessell et al., 2010; Ogarko et al., 2024). These challenges highlight the need for alternative approaches that can better balance accuracy, robustness, and computational efficiency.

In recent years, data-driven approaches have emerged as promising alternatives to conventional inverse problem solvers. Rather than explicitly inverting a physical forward model, these approaches leverage machine learning techniques to learn inverse mappings directly from data. By exploiting patterns embedded in observational or simulated datasets, learning-based models can approximate nonlinear inverse operators while reducing dependence on explicit inversion formulas (Adcock et al., 2024; Bach et al., 2025; Mohammad-Djafari et al., 2023). This paradigm has gained significant traction across diverse application domains, including medical imaging, geophysical interpretation, and engineering design, where data-driven models have demonstrated competitive accuracy and substantial computational advantages over classical methods (Adler & Öktem, 2024; Arridge et al., 2019; Paula et al., 2025; Stuart, 2010). Recent advances in inverse design and surrogate modeling further illustrate the effectiveness of machine learning in efficiently approximating inverse solutions, particularly for problems involving repeated evaluations of expensive forward models (Arridge et al., 2019; Lu et al., 2021; Paula et al., 2025).

Among supervised learning techniques, kernel-based methods remain attractive due to their solid theoretical foundations and favorable generalization properties. Support Vector Regression (SVR), in particular, extends the support vector machine framework to regression problems through the use of an ε-insensitive loss function and structural risk minimization. This formulation enables SVR to balance model complexity and data fidelity in a principled manner, resulting in robust performance under noisy and limited-data conditions (Smola & Schölkopf, 2004; Vapnik, 2000). Through kernel mappings, SVR is capable of representing complex nonlinear relationships while retaining a convex optimization structure that ensures the existence of a unique global solution (Smola & Schölkopf, 2004; Vapnik, 2000). These properties have motivated the application of SVR and related kernel-based models to inverse and ill-conditioned problems in engineering and applied sciences, where stability and generalization are critical concerns.

Recent developments in physics-informed and hybrid learning frameworks further underscore the growing role of machine learning in inverse modeling. Physics-informed machine learning methods aim to incorporate prior physical knowledge into data-driven models, thereby improving interpretability and robustness while reducing data requirements (Jiang & Gou, 2025; Karniadakis et al., 2021). Although deep learning architectures such as neural operators and physics-informed neural networks have shown impressive performance in solving high-dimensional inverse problems (Adcock et al., 2024; Bach et al., 2025; Latz, 2023), their complexity and data demands may not always be justified for low-dimensional or moderately noisy problems. In such cases, simpler and more interpretable models, including SVR, remain valuable alternatives that offer competitive performance with reduced computational and implementation complexity.

Despite the growing success of learning-based approaches for inverse problems, most existing studies primarily assess performance using pointwise prediction accuracy, while providing limited analysis of uncertainty propagation, conditioning effects, and identifiability, which are core concepts in inverse problem theory. As a result, the reliability of data-driven inverse models under noise, ill-conditioning, and non-injective mappings remains insufficiently understood, particularly for deterministic regression-based methods.

This study investigates Support Vector Regression (SVR) as a data-driven inversion framework for a one-dimensional nonlinear inverse problem defined through a synthetic forward model. Rather than emphasizing pointwise prediction accuracy, the focus is on analyzing the stability, uncertainty, and identifiability of machine-learning-based inversion under noisy observations. Monte Carlo simulation is employed to propagate measurement noise through the learned inverse mapping and to characterize the resulting distribution of inferred parameters, enabling a systematic assessment of bias, variance, and solution dispersion beyond aggregate performance metrics.

Linear regression is used as a baseline to contrast linear and kernel-based inversion behavior and to isolate the role of nonlinear representation and regularization in stabilizing inverse predictions. By examining repeated stochastic perturbations, the study reveals regions of heightened sensitivity and potential failure modes associated with ill-conditioning and noise amplification in the forward operator.

The novelty of this work lies in its Monte Carlo–based uncertainty analysis of data-driven inversion, which bridges inverse problem theory and supervised learning. In contrast to recent emphasis on complex deep learning architectures, the results demonstrate that simpler and more interpretable kernel-based models, when analyzed through a rigorous uncertainty propagation framework, can yield valuable insight into the reliability and limitations of machine-learning-based inverse solutions. This analysis provides a foundation for extending uncertainty-aware data-driven inversion to higher-dimensional and real-world problems.

2. RESEARCH METHOD

This study adopts a quantitative, uncertainty-aware experimental design to investigate data-driven inversion under noisy and potentially ill-conditioned settings. Rather than treating inversion as a deterministic regression task, the proposed methodology frames inverse estimation as a stochastic mapping problem, in which observational noise is explicitly propagated through the learned inverse model. A reproducible, algorithmic workflow is employed to ensure transparency, robustness, and scientific validity.

2.1 Problem Formulation

Consider a class of one-dimensional nonlinear forward operators $f: \mathbb{R} \to \mathbb{R}$, mapping an unknown input parameter x to an observable quantity y according to

$$y = f(x)$$

The inverse problem consists of estimating x from noisy observations $y_{obs} = y + \varepsilon$, where ε represents measurement noise. Inverse estimation is fundamentally sensitive to perturbations in the data, particularly for nonlinear operators, where small errors in the observations may lead to amplified errors in the inferred parameters. In this study, the input parameter is sampled uniformly over a bounded domain Ω = [0,5], and all experiments are conducted using N_data = 1000 samples unless otherwise stated.

2.2 Analytical Sensitivity and Conditioning

Local inversion behavior is governed by the sensitivity of the forward operator with respect to its input. Let $f \in C^1$ denote a continuously differentiable forward mapping. Under a first-order Tylor approximation, a small perturbation δy in the observation induces a corresponding perturbation in the inverse solution given by

$$\delta x \approx \frac{1}{f'(x)} \delta y,$$

which follows directly from linearized inverse problem analysis (Tarantola, 2005)
This relation shows that inverse stability is controlled by the inverse of the derivative of the forward operator. To quantify this effect, the local inverse condition number is defined as

$$\kappa(x) = \frac{1}{|f'(x)|}$$

Intuitively, regions where small perturbations in the observations lead to large changes in the inferred parameters correspond to poorly conditioned inverse problems and are inherently more difficult to solve reliably. The condition number $\kappa(x)$ provides a theoritical measure of noise amplification in the inverse solution. Regions where $|f'(x)|$ is small correspond to poorly conditioned inversion, even when forward operator is smooth and injective (Kirsch, 2021). Assuming additive Gaussian noise $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, and under the local linearization assumption, the variance of an unbiased inverse estimator can be approximated as

$$Var(\hat{x}) \approx \kappa(x)^2 \sigma^2$$

This linearization-based approximation is valid only for sufficiently small noise levels and away from critical points where $f'(x)$ approaches zero; deviations from this behavior indicate breakdown of the local linear model or the influence of regularization effects, approximation bias, or implicit prior assumptions introduced by the learning process (Kaipio & Somersalo, 2005; Latz, 2020)

2.3 Representative Forward Model

To study these effects in a controlled setting, a representative nonlinear forward model is selected from the class of smooth monotonic operators:

$$f(x) = 0.5x^3 + 2x, \qquad x \in [0,5]$$

This function is chosen not for its specific analytical form, but because it exhibits spatially varying conditioning while remaining free from non-uniqueness. This allows the isolation of noise-induced instability without confounding effects arising from multiple inverse solutions.

2.4 Extension to Non-Injective Forward Operators

While the injective case enables the analysis of inverse stability under noise, many inverse problems encountered in practice are fundamentally non-injective, admitting multiple valid parameter values for a single observation. In such settings, inverse estimation is not only unstable but also non-identifiable, and pointwise inversion becomes inherently ambiguous (Engl et al., 2000b; Kaipio & Somersalo, 2005). To examine these identifiability limitations in data-driven inversion, the methodology is extended to include a representative non-injective forward operator. Specifically, a smooth nonlinear mapping with stationary points is considered:

$$f(x) = x^3 - 3x$$

This operator is non-injective due to the presence of turning points where $f'(x) = 0$, resulting in multiple inverse branches over its domain (Kirsch, 2021). Consequently, the inverse problem becomes set-valued, and any learned inverse mapping implicitly imposes a solution selection mechanism governed by the training data distribution, model architecture, and regularization strategy.

In the non-injective case, imbalance among inverse branches in the training data affects which solution branch is preferentially selected by the learned model. As a result, deterministic regression models implicitly collapse multi-valued inverse mappings into a single dominant mode, leading to overconfident and unimodal inverse estimates despite underlying non-identifiability.

This extension enables systematic investigation of inverse ambiguity, mode collapse, and overconfident predictions in learning-based inversion, phenomena that cannot be observed in injective settings and are often obscured by aggregate error metrics (Latz, 2020).

2.5 Monte Carlo Noise Propagation

Uncertainty in inverse estimation is analyzed through a Monte Carlo framework. For a fixed true parameter value x_0, repeated noisy observations are generated as

$$y^{(k)} = f(x_0) + \varepsilon^k, \qquad \varepsilon^k \sim \mathcal{N}(0, \sigma^2), \qquad k = 1, \dots, N$$

In all experiments, N_MC = 1000 Monte Carlo realizations are generated for each fixed parameter value to ensure statistically stable estimation of inverse uncertainty. The number of Monte Carlo realizations is chosen to balance statistical accuracy and computational efficiency. Since Monte Carlo estimates of variance converge at a rate proportional to $\frac{1}{\sqrt{N_{MC}}}$, increasing the number of simulations beyond $N_{MC} \approx 1000$ yields diminishing accuracy gains relative to the associated computational cost for the uncertainty measures considered in this study.

Each realization is independently inverted using a learned regression model, producing an ensemble of inverse estimates $\{\hat{x}^k\}_{k=1}^{N}$. The resulting empirical distribution provides estimates of bias, variance, and higher-order moments, enabling direct comparison with theoretical noise amplification predicted by $\kappa(x)$.

Monte Carlo–based uncertainty propagation is widely used in applied inverse problems to assess the sensitivity of model parameters to observational perturbations, particularly in settings where deterministic inversion can obscure instability and uncertainty (Pakyuz-Charrier et al., 2018; Zong et al., 2023). Importantly, in the present study this framework treats inverse estimation as a stochastic mapping rather than a deterministic prediction task, allowing instability and overconfidence in learned inverse models to be identified even in cases where average error metrics remain low.

2.6 Learning-Based Inverse Models

The inverse mapping is approximated using supervised learning, where a regression model $\mathcal{M}_\theta$ is trained to approximate f^{-1} from noisy observations. Learning-based inversion provides a flexible alternative to explicit analytical inversion by approximating inverse operators directly from data (Arridge et al., 2019).

Two regression models are considered. Linear regression serves as a baseline model, representing a global linear approximation to the inverse mapping. Support Vector Regression (SVR) with a radial basis function kernel is employed as a nonlinear inverse model, which implicitly regularizes the inversion through kernel smoothness and margin constraints, yielding improved robustness under noise (Smola & Schölkopf, 2004).

SVR is compatible with inverse operators because, for injective forward mappings, the inverse can be viewed as a continuous nonlinear function whose stability is governed by local conditioning. Kernel-based regression enables SVR to approximate such nonlinear inverse mappings while enforcing smoothness that limits sensitivity to observational noise. Moreover, margin-based regularization provides controlled amplification of perturbations, aligning the learned inverse mapping with the stability requirements of inverse problem theory.

From a statistical perspective, in non-injective inverse problems, deterministic regression models trained with pointwise loss functions implicitly approximate a conditional expectation, i.e., $\hat{x}(y) \approx \mathbb{E}[x \mid y]$. In inverse problems where the mapping from observations to parameters is non-injective, the conditional distribution $p(x \mid y)$ is generally multi-modal. As a result, deterministic learning-based inversion necessarily collapses this distribution into a single representative value, thereby masking non-identifiability and eliminating information about solution multiplicity and uncertainty.

Unlike classical inverse estimators derived from linearized analysis, SVR implicitly modifies the underlying assumptions by introducing margin-based regularization and kernel smoothness constraints. These mechanisms suppress noise amplification and stabilize the inversion at the cost of bias, thereby altering the unbiased variance predictions implied by classical conditioning analysis.

SVR hyperparameters (C, γ, ε) are selected based on preliminary tuning experiments on the training set, balancing approximation accuracy and regularization-induced stability (Hastie et al., 2009).

Deep learning models are not considered in this study in order to isolate uncertainty propagation and conditioning effects without introducing architectural complexity or data-intensive training requirements. For the low-dimensional setting examined here, kernel-based models such as SVR provide sufficient expressive power while retaining interpretability and controlled regularization.

2.7 Evaluation Criteria

Model evaluation emphasizes distributional reliability rather than pointwise accuracy, reflecting the fact that inverse problems are inherently sensitive to uncertainty and non-uniqueness. While standard metrics such as mean squared error are reported for reference, they are insufficient to characterize inverse stability and uncertainty (Tarantola, 2005).

Accordingly, the analysis focuses on (i) variance amplification as a function of local conditioning, which provides a theoretical indicator of inverse instability (Engl et al., 2000; Stuart, 2010), (ii) bias induced by model regularization, reflecting the classical bias–variance trade-off in statistical learning (Hastie et al., 2009), and (iii) sensitivity of inverse solution distributions to noise level.

This evaluation framework enables identification of inversion regimes in which learned inverse models exhibit overconfident or unstable behavior despite achieving low average prediction error, a known failure mode of learning-based inverse approaches (Arridge et al., 2019).

Although formal posterior entropy is not explicitly computed, entropy-related behavior is assessed qualitatively through the spread, concentration, and modality of the empirical inverse distributions obtained from Monte Carlo sampling. Multimodality is identified by the presence of multiple peaks in the inverse density estimates, while posterior spread and collapse are evaluated through variance behavior and distribution concentration across realizations. This qualitative analysis complements variance-based metrics and enables detection of loss of identifiability, mode collapse, and overconfident inverse predictions that are not captured by pointwise error measures alone.

2.8 Algorithmic Implementation

The complete methodological workflow described above is summarized in Figure 1, which operationalizes the analytical conditioning framework and Monte Carlo uncertainty analysis in a reproducible manner.

```
Input: Forward operator $f(x)$, domain $\Omega$, noise variance $\sigma^2$, number of data samples $N_{data}$,
number of Monte carlo realizations $N_{MC}$
    1. Sample $\{x_i\}_{i=1}^{N_{data}} \sim \mathcal{U}(\Omega)$
    2. Compute noiseless output $y_i = f(x_i)$
    3. Add noise $y_{i,obs} = y_i + \varepsilon_i$
    4. Split data into training (70%) and test (30%) sets
    5. Train inverse models (Linear Regression, SVR) on training data
    6. For each test value $x_0$
        a. Generate Monte Carlo observations $y^k = f(x_0) + \varepsilon^k$
        b. Invert each realization $\hat{x}^k = \mathcal{M}_\theta(y^k)$
        c. Estimate empirical bias and variance
    7. Compute local conditioning $\kappa(x_0) = \frac{1}{|f'(x_0)|}$
    8. Compare empirical variance with theoritical prediction $\kappa(x_0)^2\sigma^2$
    9. Analyze stability, identifiability, and failure modes
    10. Report reference metrics (MSE, $R^2$)
```

Figure 1. Monte Carlo–Based Algorithm for Uncertainty-Aware Analysis of Learning-Based Inverse Problems

3. RESULTS AND DISCUSSIONS

3.1 Inverse Accuracy under Injective Forward Mapping

This subsection evaluates the ability of learning-based models to accurately approximate a unique inverse mapping when the forward operator is injective. The forward model $f(x) = 0.5x^3 + 2x$ admits a unique inverse over the domain of interest, providing a controlled setting in which inverse accuracy can be assessed independently of instability or non-uniqueness.

Linear regression and Support Vector Regression (SVR) were trained to estimate the inverse mapping from noisy observations. Model performance was evaluated on a held-out test set using mean squared error (MSE) and coefficient of determination (R^2).

Table 1. Inverse prediction accuracy for the injective forward operator

Model	MSE	R^2
Linear Regression	0.2654	0.8733
SVR (RBF Kernel)	0.0013	0.9994

As shown in Table 1, SVR achieves substantially lower prediction error than linear regression, reducing the MSE by more than two orders of magnitude and attaining an R^2 score close to unity. This indicates that kernel-based regression is capable of accurately representing the nonlinear inverse relationship, whereas linear regression exhibits systematic approximation error due to its global linear structure.

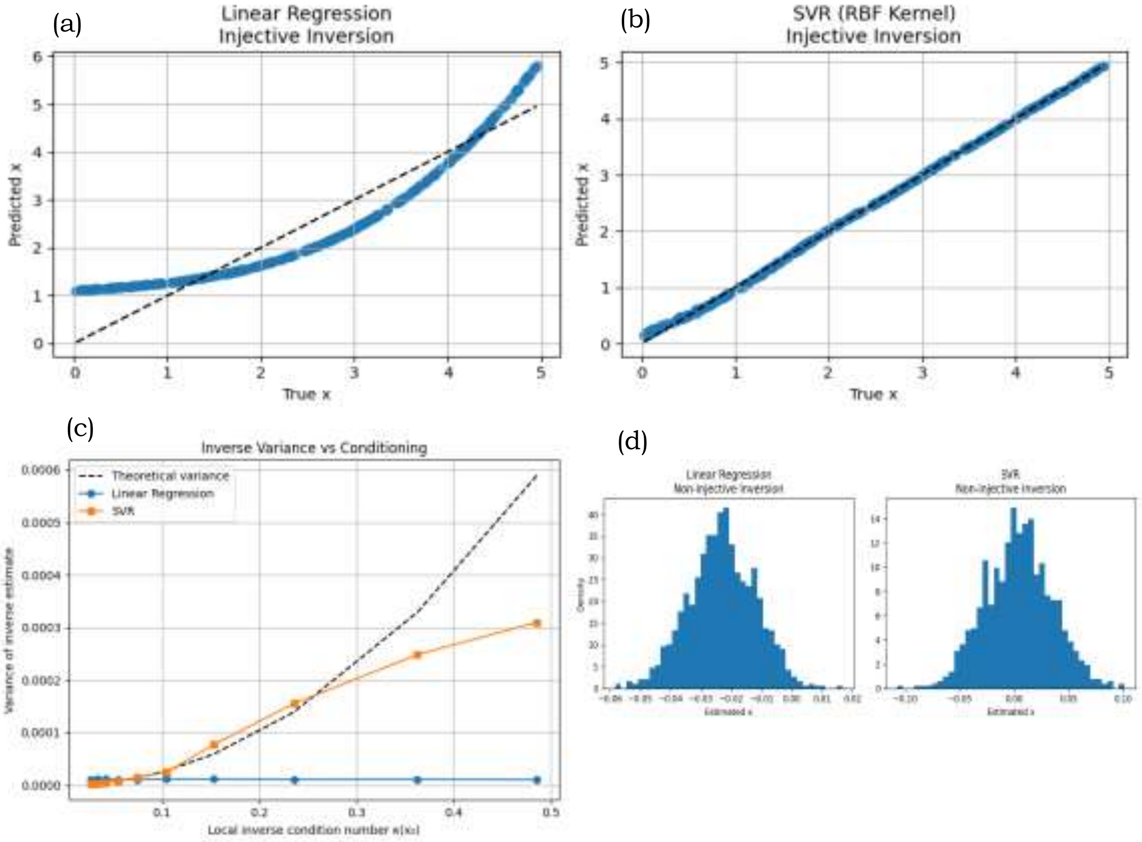

Figure 2. Inverse accuracy, conditioning-driven uncertainty, and identifiability effects in learning-based inversion

Figure 2 jointly illustrates inverse prediction accuracy, conditioning-driven uncertainty, and identifiability effects in learning-based inversion. Panels (a) and (b) show predicted-versus-true inverse estimates for the injective forward operator. Linear regression exhibits systematic bias across the parameter range, reflecting its inability to capture nonlinear inverse structure, whereas SVR closely follows the ideal inverse relationship and accurately reconstructs the inverse mapping throughout the domain.

Panel (c) demonstrates how inverse uncertainty varies with local conditioning, revealing that linear regression artificially suppresses variance, while SVR exhibits noise amplification consistent with theoretical predictions but moderated by regularization. Panel (d) highlights the behavior under non-injective inversion, where both models collapse the multi-valued inverse into a single dominant solution, exposing implicit solution selection and loss of identifiability.

Overall, these results show that for injective forward mappings, nonlinear kernel-based models such as SVR can accurately recover the inverse relationship, whereas linear regression suffers from systematic approximation bias due to its limited representational capacity.

3.2 Conditioning-Driven Inverse Uncertainty

While inverse prediction accuracy provides an important measure of model performance, it is insufficient to characterize the reliability of inverse solutions in the presence of observational noise. Inverse problems are inherently sensitive to perturbations in the data, and even when the forward operator is injective, small measurement errors may be amplified depending on the local conditioning of the inverse mapping.

To quantify this effect, inverse uncertainty was analyzed using a Monte Carlo noise propagation framework. For a fixed true parameter value $x_0 = 2.5$, repeated noisy observations were generated and independently inverted using the trained models. The resulting ensembles of inverse estimates were used to compute empirical mean, bias, and variance, which were then compared to theoretical predictions based on local inverse conditioning.

Table 2. Monte Carlo inverse statistics at $x_0 = 2.5$ for the injective forward

Model	MSE $\mathbb{E}[\hat{x}]$	Bias $\mathbb{E}[\hat{x}] - x_0$	Variance Var $(\hat{x})$
Linear Regression	1.9598	-0.5402	1.11×10^{-5}
SVR (RBF Kernel)	2.4969	-0.0031	1.62×10^{-5}
Theory	-	-	1.93×10^{-5}

Table 2 reveals a fundamental distinction between inverse accuracy and inverse uncertainty. Although linear regression exhibits a lower empirical variance than SVR, this apparent stability arises from a large systematic bias rather than faithful uncertainty propagation. In contrast, SVR produces an unbiased inverse estimate and a variance that closely matches the theoretical prediction derived from local inverse conditioning.

The dependence of inverse uncertainty on conditioning is further illustrated in Figure 1(c), which shows the variance of inverse estimates as a function of the local inverse condition number $\kappa(x_0)$. The theoretical variance predicted by linearized inverse analysis increases quadratically with $\kappa(x_0)$, reflecting intrinsic instability of the inverse problem. SVR follows this conditioning-driven uncertainty trend, indicating that it preserves the physical sensitivity of the inverse mapping. Linear regression, however, exhibits nearly constant variance across conditioning regimes, masking inverse instability through bias-dominated regularization.

These results demonstrate that low variance or low prediction error alone does not imply reliable inverse estimation. Models that artificially suppress variance through bias may produce overconfident and misleading inverse solutions. Accurate inverse modeling therefore requires not only pointwise accuracy but also uncertainty behavior consistent with inverse-problem theory.

In summary, conditioning-driven uncertainty analysis reveals that inverse reliability is governed by forward operator sensitivity rather than prediction error alone, and that regularized nonlinear models such as SVR more faithfully reflect this instability than biased linear approximations.

3.3 Identifiability Loss under Non-Injective Forward Mapping

The previous subsections examined inverse accuracy and uncertainty under an injective forward operator, where a unique inverse exists for each observation. However, many inverse problems encountered in practice are fundamentally non-injective, admitting multiple parameter values that produce the same observable response. In such settings, inverse estimation is not only unstable but also non-identifiable, and the notion of a single “correct” inverse solution becomes ill-defined. To investigate these identifiability limitations in a controlled setting, the forward operator was replaced by a smooth non-injective mapping $f(x) = x^3 - 3x$, which contains stationary points where $f'(x) = 0$ and therefore admits multiple inverse branches over its domain. For a fixed observation $y_0 = f(x_0)$, the inverse solution is set-valued, and any learning-based inverse model must implicitly select one of the admissible solutions.

Monte Carlo inversion was performed at the ambiguous point $x_0 = 0$, and repeated noisy observations were inverted using the trained linear regression and SVR models. The resulting ensembles of inverse estimates were analyzed to characterize the distributional behavior of the learned inverse mappings. As shown in Figure. 2(d), linear regression produces inverse estimates tightly concentrated around a single value near zero, effectively collapsing the multi-valued inverse into a mean solution. This behavior reflects its global averaging nature, which suppresses ambiguity but yields an inverse estimate that does not correspond to any physically meaningful branch of the true inverse. SVR, while capable of representing nonlinear relationships, also exhibits implicit solution selection: the inverse estimates concentrate around a preferred branch determined by the training data distribution and kernel regularization, with increased variance relative to the injective case. Nevertheless, SVR fails to represent the full set of admissible inverse solutions, producing a unimodal distribution where the true inverse is multi-modal.

These results demonstrate a fundamental limitation of deterministic learning-based inversion for non-injective problems. In the absence of explicit mechanisms for representing multi-valued solutions or posterior distributions, learned inverse models necessarily impose implicit priors that select a single solution branch. As a result, low prediction error in non-injective settings can be misleading, as apparently stable inverse estimates may fail to capture the intrinsic ambiguity of the underlying inverse problem. This observation motivates the use of uncertainty-aware and distribution-preserving inversion frameworks when addressing non-identifiable inverse problems.

In summary, non-injective forward mappings expose a fundamental limitation of deterministic learning-based inversion, whereby implicit solution selection collapses intrinsically multi-valued inverses into overconfident single-mode estimates that mask true non-identifiability.

3.4 Implications for Learning-Based Inverse Problems

Taken together, the results demonstrate that learning-based inverse models can exhibit fundamentally different behaviors depending on problem structure. In injective settings, nonlinear models such as SVR can accurately approximate inverse mappings and produce uncertainty behavior that reflects forward conditioning. However, even in this favorable regime, regularization significantly alters variance amplification relative to theoretical predictions (Tarantola, 2005). In non-injective settings, the limitations become substantially more severe. Both linear and nonlinear regression models collapse inherently multi-valued inverse mappings into single-valued predictions, masking non-identifiability and producing overconfident inverse estimates. This behavior reflects the imposition of implicit priors induced by training data distribution, loss functions, and regularization, rather than genuine physical identifiability (Arridge et al., 2019; Engl et al., 2000a). Importantly, standard accuracy metrics fail to detect this failure mode: low prediction error can coexist with an incorrect representation of inverse ambiguity. Similar concerns have been raised in recent analyses of learning-based inverse problems, which emphasize that pointwise accuracy alone is insufficient for assessing inverse reliability (Adcock et al., 2024; Latz, 2020). These findings highlight the necessity of jointly considering conditioning, uncertainty propagation, and identifiability when evaluating learning-based inversion methods, particularly in ill-posed or non-injective settings.

These findings have direct implications for a range of applied inverse problems. In imaging applications, such as medical or computational imaging, high reconstruction accuracy achieved by deterministic learning-based methods may conceal overconfident uncertainty estimates, potentially leading to misleading interpretations in regions of poor conditioning. In geophysical inversion, where non-uniqueness and data sparsity are intrinsic, deterministic inverse models may implicitly enforce solution selection driven by training data bias rather than geological plausibility, thereby masking alternative admissible subsurface models. In inverse design problems, where multiple designs can yield equivalent performance, deterministic regression-based inversion risks collapsing diverse design solutions into a single representative configuration, limiting exploration of the design space. These considerations highlight the importance of uncertainty-aware and distribution-preserving learning frameworks when deploying data-driven inversion in real-world scientific and engineering applications.

Overall, these implications underscore that reliable learning-based inversion requires evaluation criteria grounded in inverse problem theory, rather than reliance on pointwise accuracy alone, especially in ill-conditioned or non-identifiable settings.

4. CONCLUSION

This work examined learning-based inverse modeling through the lens of inverse problem theory, focusing on uncertainty propagation, conditioning, and identifiability rather than pointwise accuracy alone. By treating inverse estimation as a stochastic mapping and employing Monte Carlo noise propagation, the study revealed how learned inverse models respond to observational noise and problem structure. For injective forward mappings, Support Vector Regression (SVR) accurately approximated the nonlinear inverse and significantly outperformed linear regression in terms of prediction accuracy. However, inverse uncertainty was shown to depend strongly on the local conditioning of the forward operator. While linearized theory predicts variance amplification proportional to the inverse condition number, learned inverse models deviate from this behavior due to regularization and approximation bias. In particular, linear regression suppresses variance at the cost of substantial bias, whereas SVR moderates uncertainty growth while maintaining accuracy.

In non-injective settings, both linear regression and SVR were found to impose implicit solution selection, collapsing inherently multi-valued inverse relationships into unimodal estimates. This behavior produces overconfident inverse solutions that fail to reflect the true ambiguity of the inverse problem, even when prediction error remains low. These results demonstrate that accuracy alone is insufficient for assessing learning-based inverse models; conditioning, uncertainty, and identifiability must be jointly considered to evaluate inverse reliability. The proposed Monte Carlo–based analysis provides a practical framework for diagnosing instability and hidden failure modes in data-driven inversion, and it motivates the use of uncertainty-aware and distribution-preserving approaches for ill-posed and non-identifiable inverse problems.

This study is limited to low-dimensional synthetic inverse problems and deterministic regression models, which enables controlled analysis but does not capture the full complexity of high-dimensional or real-world inverse settings. Future research will extend this uncertainty-aware framework to higher-dimensional and physics-informed inverse problems, and to probabilistic or distribution-preserving learning approaches capable of explicitly representing multi-modality and posterior uncertainty.

ACKNOWLEDGEMENTS

REFERENCES

Adcock, B., Dexter, N., & Moraga Scheuermann, S. (2024). Optimal deep learning of holomorphic operators between Banach spaces. https://doi.org/10.48550/arXiv.2406.13928

Adler, J., & Öktem, O. (2024). Deep Bayesian Inversion. In Data-Driven Models in Inverse Problems (pp. 359–412). De Gruyter. https://doi.org/10.1515/9783111251233-011

Arridge, S., Maass, P., Öktem, O., & Schönlieb, C. B. (2019). Solving inverse problems using data-driven models. Acta Numerica, 28, 1–174. https://doi.org/10.1017/S0962492919000059

Bach, E., Baptista, R., Sanz-Alonso, D., & Stuart, A. (2025). Machine Learning for Inverse Problems and Data Assimilation. ArXiv Preprint ArXiv:2410.10523. https://arxiv.org/abs/2410.10523

Bangerth, W., Johnson, C. R., Njeru, D. K., & van Bloemen Waanders, B. (2025). Estimating and using information in inverse problems. Inverse Problems and Imaging. https://doi.org/10.3934/ipi.2026003

Chung, J., & Gazzola, S. (2024). Computational Methods for Large-Scale Inverse Problems: A Survey on Hybrid Projection Methods. SIAM Review, 66(2), 205–284. https://doi.org/10.1137/21M1441420

Engl, H. W., Hanke, M., & Neubauer, A. (2000). Regularization of Inverse Problems. Springer Netherlands. https://books.google.co.id/books?id=VuEV-Gj1GZcC

Gallet, A., Rigby, S., Tallman, T. N., Kong, X., Hajirasouliha, I., Liew, A., Liu, D., Chen, L., Hauptmann, A., & Smyl, D. (2022). Structural engineering from an inverse problems perspective. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 478(2257), 20210526. https://doi.org/10.1098/rspa.2021.0526

Giraud, J., Lindsay, M., Ogarko, V., Jessell, M., Martin, R., & Pakyuz-Charrier, E. (2019). Integration of geoscientific uncertainty into geophysical inversion by means of local gradient regularization. Solid Earth, 10(1), 193–210. https://doi.org/10.5194/se-10-193-2019

Jessell, M. W., Ailleres, L., & de Kemp, E. A. (2010). Towards an integrated inversion of geoscientific data: What price of geology? Tectonophysics, 490(3), 294–306. https://doi.org/https://doi.org/10.1016/j.tecto.2010.05.020

Ji, K., Shen, Y., Chen, Q., Li, B., & Wang, W. (2022). An Adaptive Regularized Solution to Inverse Ill-Posed Models. IEEE Transactions on Geoscience and Remote Sensing, 60, 1–15. https://doi.org/10.1109/TGRS.2022.3205572

Jiang, Q., & Gou, Z. (2025). Solutions to Two- and Three-Dimensional Incompressible Flow Fields Leveraging a Physics-Informed Deep Learning Framework and Kolmogorov–Arnold Networks. International Journal for Numerical Methods in Fluids, 97. https://doi.org/10.1002/fld.5374

Kaipio, J. P., & Somersalo, E. (2005). Statistical and Computational Inverse Problems (1st ed.). Springer-Verlag. https://doi.org/10.1007/b138659

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nature Reviews Physics, 3(6), 422–440. https://doi.org/10.1038/s42254-021-00314-5

Kirsch, A. (2021). An Introduction to the Mathematical Theory of Inverse Problems (3rd ed.). Springer. https://doi.org/10.1007/978-3-030-63343-1

Latz, J. (2020). On the Well-posedness of Bayesian Inverse Problems. SIAM/ASA Journal on Uncertainty Quantification, 8(1), 451–482. https://doi.org/10.1137/19M1247176

Latz, J. (2023). Bayesian Inverse Problems Are Usually Well-Posed. SIAM Review, 65(3), 831–865. https://doi.org/10.1137/23M1556435

Lu, L., Jin, P., Pang, G., Zhang, Z., & Karniadakis, G. E. (2021). Learning Nonlinear Operators via DeepONet Based on the Universal Approximation Theorem of Operators. Nature Machine Intelligence, 3(3), 218–229. https://doi.org/10.1038/s42256-021-00302-5

Mohammad-Djafari, A., Chu, N., Wang, L., & Yu, L. (2023). Bayesian Inference and Deep Learning for Inverse Problems. Physical Sciences Forum, 9(1), 14. https://doi.org/10.3390/psf2023009014

Ogarko, V., Frankcombe, K., Liu, T., Giraud, J., Martin, R., & Jessell, M. (2024). Tomofast-x 2.0: an open-source parallel code for inversion of potential field data with topography using wavelet compression. Geoscientific Model Development, 17(6), 2325–2345. https://doi.org/10.5194/gmd-17-2325-2024

Ogarko, V., Giraud, J., Martin, R., & Jessell, M. (2021). Disjoint interval bound constraints using the alternating direction method of multipliers (ADMM) for geologically constrained inversion: Application to gravity data. Geophysics, 86(2), G1–G11. https://doi.org/10.1190/geo2019-0633.1

Pakyuz-Charrier, E., Lindsay, M., Ogarko, V., Giraud, J., & Jessell, M. (2018). Monte Carlo simulation for uncertainty estimation on structural data in implicit 3-D geological modeling, a guide for disturbance distribution selection and parameterization. Solid Earth, 9(2), 385–402. https://doi.org/10.5194/se-9-385-2018

Paula, M. C. L., Jessell, M., Cripps, E., Lindsay, M., Pirot, G., & Gibson, L. (2025). Machine learning to assess troglofauna occurrences in the northern part of Western Australia. Next Research, 2(3), 100693. https://doi.org/https://doi.org/10.1016/j.nexres.2025.100693

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. Statistics and Computing, 14(3), 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Stuart, A. M. (2010). Inverse Problems: A Bayesian Perspective. Acta Numerica, 19, 451–559. https://doi.org/10.1017/S0962492910000061

Tarantola, A. (2005). Inverse Problem Theory and Methods for Model Parameter Estimation. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898717921

Tikhonov, A. N., & Arsenin, V. I. (1977). Solutions of Ill-Posed Problems. Winston.

Vapnik, V. N. (2000). The Nature of Statistical Learning Theory (2nd ed.). Springer. https://doi.org/10.1007/978-1-4757-3264-1

Zhai, M., Ji, Y., Pei, R., Xu, L., Chen, Y., & Lu, W. (2025). Transformer-Based PINN for Semisupervised Electromagnetic Forward Simulations. IEEE Antennas and Wireless Propagation Letters, 24(11), 3956–3960. https://doi.org/10.1109/LAWP.2025.3583011

Zong, Y., Barajas-Solano, D., & Tartakovsky, A. M. (2023). Randomized Physics-Informed Machine Learning for Uncertainty Quantification in High-Dimensional Inverse Problems. https://arxiv.org/abs/2312.06177