Minimax Theory and its Applications
Volume 11 (2026), No. 4s, 274 to 283

Dynamic system identification through inverse optimization models for industrial process monitoring

Dr.P. Mathiyalagan[1], S. Sakthivel[2], Dr. Huma Khan[3], Dr. Kamlesh Kumar Yadav[4], P. Ganeshan[5], Vineetha[6]

[1]Professor, Department of Mechanical Engineering, J.J. College of Engineering and Technology, Tiruchirappalli, Tamil Nadu, India. mathis09051970@yahoo.co.in, https://orcid.org/0000-0002-2106-4623

[2]Assistant Professor, Department of Mechanical Engineering, K.S. Rangasamy College of Technology, Namakkal, Tamil Nadu, India. sakthivel_s@ksrct.ac.in, https://orcid.org/0000-0002-8543-7011

[3]Associate Professor, Department of Computer Science and Engineering, Rungta College of Engineering and Technology, Rungta International Skills University, Bhilai, Chhattisgarh, India. huma@rungta.org, https://orcid.org/0000-0003-1653-8676

[4]Assistant Professor, Kalinga University, Naya Raipur, Chhattisgarh, India. ku.kamleshkumaryadav@kalingauniversity.ac.in, https://orcid.org/0009-0006-4665-5506

[5]Department of Mechanical Engineering, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India. ganeshan.p@sece.ac.in, https://orcid.org/0000-0002-2403-7495

[6]Department of MBA, Ramachandra College of Engineering, Eluru, Andhra Pradesh, India. vineetha@rcee.ac.in

Abstract:

This paper proposes a novel dynamic system identification approach based on inverse optimization for the monitoring of industrial processes. The proposed method recovers unknown parameters of an objective function, trade-offs, and process decision-making mechanisms from time-indexed data gathered via industrial sensors, SCADA systems, energy meters, and production logs. In contrast to traditional forward identification approaches that only identify input-output relationships, the proposed framework aims at reconstructing the optimization procedures behind observed process decisions. An optimization model with a parametric objective function, observables of the operational context, decision variables, and feature mappings, along with engineering constraints, is proposed. Unknown parameters are recovered via minimizing inconsistencies of the Karush-Kuhn-Tucker optimality conditions in multiple decision periods, which allows us to account for time-evolving process conditions and near-optimal behavior. The scale normalization is used to derive the trade-off ratios for the operational goals that can be interpreted. The derived dynamic model helps in the process of monitoring the process continuously, predicting its trajectory, detecting any parameter drifts, and determining abnormal operation modes. Rolling window estimate allows adapting the model to possible structural changes in the process behavior. The approach ensures the trade-off between predictive quality, constraints and interpretability, overcoming the drawbacks of the traditional black box identification models. Thus, in general, the work creates an inverse optimization-based framework for translating decisions made in real-world industry into dynamic and interpretable process models.

Keywords: Dynamic system identification, inverse optimization, industrial process monitoring, KKT residual, parameter estimation, adaptive control

1. Introduction

The modern industrial processes work in complicated and interconnected systems, and therefore precise system identification is critical for proper monitoring, forecasting, and optimization. The recent scientific literature has

shown that using inverse optimization allows recovering the hidden schedule of the production and its decision structure based on the smart-metering and industrial measurements data. This method creates a good basis for data-driven industrial load and process modeling (Lyu et al., 2024). On the other hand, the artificial intelligence methods increasingly integrate system identification, control, and optimization to create adaptable models for the nonlinear and uncertain industrial environment (Yu et al., 2025). Inverse modeling methodologies for dynamic thermal systems show the possibility of learning the inverse connections between the system behavior and control actions for better process control (Škrjanc et al., 2026). Optimization algorithms enhance efficient management of energy and industry systems due to the combination of operational data with intelligent decision making (Vij & Nayak, 2025).

The topic of dynamic optimization has become very relevant when applied to complex systems, since it allows using data-based methods to take into account the changing operating conditions of the process and optimize it (Tian et al., 2024). In-process identification methods using inverse dynamics and sensitivity analysis have also been proven to be very useful for reducing the complexity of models and maintaining required accuracy for real-time industrial applications (Volkmann et al., 2020). Combining dynamic optimization and control theory has become very important for providing theoretical foundations for combining system identification with decisions in the industrial environment (Kadam et al., 2003). Machine learning techniques, such as support vector regression, have been proven to be able to recognize the non-linear behavior of industrial processes like batch reactors Valluvan & Sakthivel, (2026), while the use of advanced techniques of system identification has focused on the relevance of modeling, parameter estimation, and implementation in complex engineering systems (Mazhar et al., 2026; Zhao et al., 2026).

Modeling in advanced process monitoring also requires models which can model discrepancies between ideal process models and industrial reality, which makes dynamic reduced order modeling an important concept for increased model fidelity and computational efficiency (Dinh et al., 2026)]. Frameworks based on inverse learning can be utilized to facilitate the joint optimization of industrial energy expenses, constraints, and system performances by discovering optimal policies from past data observations (Hu & Li, 2026). The combination of artificial intelligence technology with Industrial Internet of Things (IIoT) technology has created new possibilities for predictive maintenance and continuous process monitoring by analyzing real-time sensor data (Kumar & Tiwari, 2024). Another way to illustrate the importance of noise reduction in the complicated sensor signal is through the application of deep learning techniques in multichannel physiological signal processing (Maurya et al., 2025). Hybrid deep learning models and forecasting algorithms are another way of improving data fidelity and taking into account dynamic changes in complex industrial information control systems (Sevinov et al., 2026). Another way of linking physics and data driven modeling is through physics informed and sensor-driven machine intelligence (Bhadola et al., 2026).

Hybrid machine learning methods guided by scientific principles further illustrate the effectiveness of incorporating first-principal knowledge into data models to accurately describe complex industrial processes (Sharma & Liu, 2022). Process-aware information systems also illustrate the significance of data collection and monitoring for improving industrial performance (Murugan et al., 2024). Iterative dynamic identification techniques have demonstrated that through iterative parameter refinement, models of industrial systems can be refined to improve their accuracy and allow for reliable monitoring at different operation modes (Han et al., 2020). Dynamic modeling based on data has been also extensively applied in the field of power and energy systems showing the importance of understanding time-dependent properties of industrial systems for monitoring and control (Subedi et al., 2025). Recently, inverse optimization along with surrogate and adaptive optimization has been employed to discover operational strategies that increase efficiency of industrial systems considering complex dependencies between performance parameters (Liu et al., 2026).

Overall, these advancements clearly demonstrate the increasing similarity between dynamic system identification, inverse optimization, artificial intelligence, machine learning, and industrial process monitoring. Contrary to traditional identification methods that mainly deal with estimating the parameters of the model based on input/output observations, inverse optimization models are able to extract the objectives, constraints, and decision-making mechanisms behind the process. It becomes especially important to take into account the nature of modern industrial processes characterized by the dynamics, nonlinearity, uncertainty, and multi-objective operation. Thus, it makes sense to consider a new concept of Dynamic System Identification using Inverse Optimization Models for Industrial Process Monitoring as a means of reconstructing unknown structures, estimating dynamics continuously, detecting abnormalities at an early stage, and optimizing industrial processes.

Key contribution

- Demonstrates a new method for determining the dynamics, unknown parameters, goals and constraints of industrial processes through machine learning based on observed process and decision histories.

- Merges real-time or historical sensor data and dynamic models for ongoing monitoring of process behavior, anomaly detection, and improvement in estimation of the state of the system.

- Combines the identification of dynamic systems and inverse optimization for prediction of future process behavior and optimization-based decision-making.

Section II will discuss the preliminaries which include the notation, problem setup, and theory behind inverse optimization using KKT residuals, feasibility, identifiability, and robustness. Section III will introduce the proposed analysis pipeline, which includes the acquisition of industry data and the specification of constraints for dynamic system identification. Section IV will show the results, which consist of the estimated parameters, residual diagnosis, temporal residual analysis, out-of-sample validation, drift detection, anomaly detection, and comparison to traditional identification models. Section V will discuss the results, their limitations, validity issues, and how this work relates to prior research. Lastly, Section VI will conclude this paper and outline future research directions.

2. Preliminaries

The monitored industrial process at each decision instance t is governed by the parametric constrained optimization problem $P(\theta, u_t)$: $\underset{x \in \mathbb{R}^n}{\text{minimize}}\ f(x;\theta)$ subject to $g(x;u_t) \le 0,\ h(x;u_t) = 0$, where x is the control decision, u_t is the observable context variable, and $\theta \in \Theta \subseteq \mathbb{R}^p$ is the unknown cost parameter vector to be identified. Under Assumption A1 (Convexity), for every fixed u_t, the objective $f(\cdot;\theta)$ is convex in x for all $\theta \in \Theta$, satisfying $\forall x_1, x_2 \in \mathbb{R}^n, \alpha \in [0,1]: f(\alpha x_1 + (1-\alpha)x_2;\theta) \le \alpha f(x_1;\theta) + (1-\alpha) f(x_2;\theta)$, and the feasible set $\mathcal{X}(u_t) = \{x \in \mathbb{R}^n \mid g(x;u_t) \le 0, h(x;u_t) = 0\}$ is closed and convex. By Assumption A2 (Regularity), a constraint qualification such as Slater's condition holds on $\mathcal{X}(u_t)$ for every u_t (i.e., $\exists x$ such that $g_i(x;u_t) < 0\ \forall i$ and $h_j(x;u_t) = 0\ \forall j$), making the Karush-Kuhn-Tucker (KKT) conditions both necessary and sufficient for global optimality. Under Assumption A3 (Linear-in-Parameters Objective), the cost function decomposes as $f(x;\theta) = \theta^\top \phi(x) = \sum_{l=1}^{p} \theta_l\, \phi_l(x)$ for a known feature mapping $\phi: \mathbb{R}^n \to \mathbb{R}^p$. Finally, under Assumption A4 (Approximate Optimality of Observations), the observed decisions $\hat{x}_t$ satisfy the perturbed KKT system $\nabla_x f(\hat{x}_t;\theta) + \sum_{i=1}^{m} \lambda_{t,i}\, \nabla_x g_i(\hat{x}_t;u_t) + \sum_{j=1}^{k} \nu_{t,j}\, \nabla_x h_j(\hat{x}_t;u_t) = \epsilon_t^{\text{stat}}$, $g_i(\hat{x}_t;u_t) \le \epsilon_{t,i}^{\text{feas,g}}\ \forall i$, $|h_j(\hat{x}_t;u_t)| \le \epsilon_{t,j}^{\text{feas,h}}\ \forall j$, $\lambda_t \ge 0$, and $|\lambda_{t,i} g_i(\hat{x}_t;u_t)| \le \epsilon_{t,i}^{\text{comp}}\ \forall i$, subject to a uniform residual bound $\|\epsilon_t^{\text{stat}}\| + \|\epsilon_t^{\text{feas,g}}\| + \|\epsilon_t^{\text{feas,h}}\| + \|\epsilon_t^{\text{comp}}\| \le C < \infty$.

2.1. Notation used

The following notations are used for this research referred in table1.

Table1: Notations are used.

Symbol	Meaning
$t = 1, \dots, T$	Discrete observation index (sampling instant or decision epoch)
$x_t \in R^{\text{n}}$	True decision variable of the process at time t
$\hat{x}_t$	Observed / measured decision at time t
$u_t \in R^{\text{m}}$	Observable exogenous context at time t
$\theta \in \Theta \subseteq R^{\text{p}}$	Unknown objective-function parameter vector to be identified
$\varphi(x) \in R^{\text{p}}$	Known feature map defining the linear-in-parameters objective $f(x;\theta) = \theta^{\text{T}}\varphi(x)$
$g(x;u_t), h(x;u_t)$	Inequality / equality constraint functions defining the feasible set X(u_t)
λ_t, ν_t	Lagrange multipliers (dual variables) associated with g and h at time t
$r_t(\theta, \lambda_t, \nu_t)$	KKT residual of the candidate parameter θ at observation t
$\|\cdot\|$	A norm on the relevant residual space ($specify\ \ell 1, \ell 2, or\ \ell\infty\ when\ instantiating$)
$\Theta *$	Set of parameter vectors consistent with all T observations (identifiability set)

2.2. Problem formulation

Given T observed pairs $\{(u_t, \hat{x}_t)\}_{t=1}^{T}$ collected from an industrial process instrumented with standard sensors (flow meters, energy meters, production logs, SCADA/PLC tags), the dynamic system identification task is to find an optimal cost parameter vector $\hat{\theta} \in \Theta$ that minimizes a measure of KKT-residual inconsistency across all T observations by solving the inverse optimization problem $\hat{\theta} = \arg \min_{\theta \in \Theta, \{\lambda_t, \nu_t\}_{t=1}^{T}} \sum_{t=1}^{T} \mathcal{L}\left(r_t(\theta, \lambda_t, \nu_t)\right)$, where $\mathcal{L}(\cdot)$ is a suitable norm or loss function and the residual vector $r_t(\theta, \lambda_t, \nu_t) = \left[r_t^{\text{stat}}, r_t^{\text{comp}}, r_t^{\text{dual}}\right]^{\top}$ aggregates the KKT violations at decision instance t. Specifically, evaluated at the observed action $\hat{x}_t$, the stationarity residual is defined as $r_t^{\text{stat}} = \sum_{l=1}^{p} \theta_l\, \nabla_x \phi_l(\hat{x}_t) + \sum_{i=1}^{m} \lambda_{t,i}\, \nabla_x g_i(\hat{x}_t;u_t) + \sum_{j=1}^{k} \nu_{t,j}\, \nabla_x h_j(\hat{x}_t;u_t)$, the complementary

slackness residual is defined entrywise as $r_{t,i}^{\text{comp}} = \lambda_{t,i} g_i(\hat{x}_t; u_t)$ for $i \in \{1, \dots, m\}$, and the dual feasibility violation is captured by $r_t^{\text{dual}} = \max(0, -\lambda_t)$, that are jointly optimized over θ and the unobservable auxiliary dual variables $(\lambda t, \nu t\ ) \in R_(\geq 0)m \times Rk$. This generalizes the static inverse optimization models of n and the KKT residual approach of K to a dynamic time-varying setting: as opposed to solving for θ based on a single static decision instance, the process is tracked over T decision periods with differing context u_t so that one may find a θ (or a smoothly varying θt) that fits all the observed decisions in changing environments.

2.3. Theoretical foundations: notations, theorems and proof

Feasibility and Consistency of the KKT-residual Formulation

Theorem 1 (Existence of a consistent parameter under exact optimality). *Under Assumptions A1–A3, if $\hat{x}_t$ is an exact optimal solution of $P\theta^*$*, ut for some fixed $\theta^* \in \Theta$ and for every $t = 1, \dots, T$, then there exist dual variable vectors $\{(\lambda_t^*, \nu_t^*)\}_{t=1}^T \subset \mathbb{R}_{\geq 0}^m \times \mathbb{R}^k$ such that: *

$$r_t(\theta^*, \lambda_t^*, \nu_t^*) = 0, \quad \forall t \in \{1, \dots, T\}$$

*Consequently, $\theta^* = \theta$ is a global minimizer of the inverse optimization residual objective with optimal value zero: **

$$\sum_{t=1}^{T} \|r_t(\theta^*, \lambda_t^*, \nu_t^*)\| = 0$$

Proof. Fix $t \in \{1, \dots, T\}$ and temporarily drop the subscript t for readability. Under Assumption A1, $P(\theta^*, u)$ is a convex optimization problem in x with linear-in-parameters objective $f(x; \theta^*) = \theta^{*\top} \phi(x)$ (Assumption A3) and feasible set $\mathcal{X}(u) = \{x \in \mathbb{R}^n \mid g(x; u) \leq 0, h(x; u) = 0\}$. By Assumption A2, a constraint qualification (e.g., Slater's condition) holds on $\mathcal{X}(u)$ at the assumed optimal solution $\hat{x}$. Hence, the First-Order Karush-Kuhn-Tucker (KKT) conditions are necessary for global optimality, guaranteeing the existence of multiplier vectors $\lambda^* \in \mathbb{R}^m$ and $\nu^* \in \mathbb{R}^k$ satisfying:

$$\nabla_x f(\hat{x}; \theta^*) + \sum_{i=1}^{m} \lambda_i^* \nabla_x g_i(\hat{x}; u) + \sum_{j=1}^{k} \nu_j^* \nabla_x h_j(\hat{x}; u) = 0$$

$$\lambda_i^* g_i(\hat{x}; u) = 0, \quad \forall i \in \{1, \dots, m\}$$

$$g(\hat{x}; u) \leq 0, \quad h(\hat{x}; u) = 0, \quad \lambda^* \geq 0$$

By construction, the KKT residual vector $r_t(\theta, \lambda_t, \nu_t) = \left[r_t^{\text{stat}}, r_t^{\text{comp}}, r_t^{\text{dual}}\right]^\top$ evaluates these conditions at $(\theta, \lambda_t, \nu_t)$:

$$r_t^{\text{stat}} = \sum_{l=1}^{p} \theta_l \nabla_x \phi_l(\hat{x}_t) + \sum_{i=1}^{m} \lambda_{t,i} \nabla_x g_i(\hat{x}_t; u_t) + \sum_{j=1}^{k} \nu_{t,j} \nabla_x h_j(\hat{x}_t; u_t)$$

$$r_{t,i}^{\text{comp}} = \lambda_{t,i} g_i(\hat{x}_t; u_t), \quad \forall i \in \{1, \dots, m\}$$

$$r_t^{\text{dual}} = \max(0, -\lambda_t)$$

Evaluating r_t at $(\theta^*, \lambda_t^*, \nu_t^*)$ yields $r_t^{\text{stat}} = 0$, $r_t^{\text{comp}} = 0$, and $r_t^{\text{dual}} = 0$, implying:

$$\|r_t(\theta^*, \lambda_t^*, \nu_t^*)\| = 0, \quad \forall t \in \{1, \dots, T\}$$

Summing over all T observations give:

$$\min_{\theta \in \Theta, \{\lambda_t, \nu_t\}} \sum_{t=1}^{T} \|r_t(\theta, \lambda_t, \nu_t)\| \leq \sum_{t=1}^{T} \|r_t(\theta^*, \lambda_t^*, \nu_t^*)\| = 0$$

Since norms are non-negative ($\|r_t\| \geq 0$), the objective function is bounded below by 0. Thus, $\theta^* = \hat{\theta}$ achieves the global minimum value of zero.

2.4. Identifiability up to Scale

Theorem 2 (Scale non-identifiability; (Yu et al., 2025; Kumar & Tiwari, 2024). *Under Assumptions A1–A3, if* θ^* achieves $r_t(\theta^*, \lambda_t, \nu_t) = 0$ for all $t = 1, \dots, T$ with dual multipliers (λ_t, ν_t), then for every scalar $c > 0$, $c\theta^*$ also achieves $r_t(c\theta^*, c\lambda_t, c\nu_t) = 0$ for all t. Consequently, θ^* is identifiable only up to a positive scalar multiple, and the optimal parameter set satisfies:

$$\Theta^* \supseteq \{c\theta^* : c > 0\}$$

Proof. Fix $t \in \{1, \dots, T\}$ and drop the time index t for readability. Under Assumption A3, $f(x; \theta) = \theta^\top \phi(x)$, which yields $\nabla_x f(\hat{x}; c\theta) = c\nabla_x f(\hat{x}; \theta)$. Scaling $(\theta, \lambda, \nu) \mapsto (c\theta, c\lambda, c\nu)$ preserves all four system conditions:

$$\begin{aligned} r_t^{\text{stat}}(c\theta, c\lambda, c\nu) &= \nabla_x f(\hat{x}; c\theta) + \sum_{i=1}^{m} (c\lambda_i)\, \nabla_x g_i(\hat{x}; u) + \sum_{j=1}^{k} (c\nu_j)\, \nabla_x h_j(\hat{x}; u) \\ &= c\left[\nabla_x f(\hat{x}; \theta) + \sum_{i=1}^{m} \lambda_i\, \nabla_x g_i(\hat{x}; u) + \sum_{j=1}^{k} \nu_j\, \nabla_x h_j(\hat{x}; u)\right] \\ &= c \cdot r_t^{\text{stat}}(\theta, \lambda, \nu) = c \cdot 0 = 0 \end{aligned}$$

$$r_{t,i}^{\text{comp}}(c\lambda) = (c\lambda_i) g_i(\hat{x}; u) = c\big(\lambda_i g_i(\hat{x}; u)\big) = c \cdot 0 = 0, \quad \forall i \in \{1, \dots, m\}$$

$$r_t^{\text{dual}}(c\lambda) = \max(0, -c\lambda) = c \max(0, -\lambda) = c \cdot 0 = 0 \quad (\text{since } c > 0)$$

Primal feasibility conditions $g(\hat{x}; u) \le 0$ and $h(\hat{x}; u) = 0$ remain untouched as they are independent of (θ, λ, ν). Thus, $r_t(c\theta^*, c\lambda_t, c\nu_t) = 0$ for all t, implying $c\theta^* \in \Theta^*$.

Practical Implication for Industrial Monitoring

Since absolute scaling costs can never be determined from observed optimal decisions, a scaling rule must be added to ensure that parameters are uniquely determined:

$$\|\hat{\theta}\| = 1 \quad \text{or} \quad \theta_k = \theta_k^{\text{known}} \quad \text{(e.g., pinning a known metered energy price)}$$

Alternatively, evaluation can rely on scale-invariant marginal rates of substitution (Ratios of Objective Coefficients):

$$\text{MRS}_{i,j} = \frac{\hat{\theta}_i}{\hat{\theta}_j}$$

2.5. Robustness under approximate optimality

Theorem 3 (Bounded residual under bounded sub-optimality; after (Kadam et al., 2003; Valluvan & Sakthivel, 2026)). *Under Assumptions A1, A2, and A4, suppose the true generating parameter is $\theta^$* and each observation $\hat{x}_t$ is an ε_t-approximate KKT point of $P(\theta^*, u_t)$ in the sense that there exist dual multipliers $(\lambda_t, \nu_t) \in \mathbb{R}^m_{\ge 0} \times \mathbb{R}^k$ satisfying: *

$$\left\| \nabla_x f(\hat{x}_t; \theta^*) + \sum_{i=1}^{m} \lambda_{t,i}\, \nabla_x g_i(\hat{x}_t; u_t) + \sum_{j=1}^{k} \nu_{t,j}\, \nabla_x h_j(\hat{x}_t; u_t) \right\| \le \varepsilon_t$$

$$\left| \sum_{i=1}^{m} \lambda_{t,i}\, g_i(\hat{x}_t; u_t) \right| \le \varepsilon_t$$

Then the residual objective evaluated at the true parameter θ satisfies: *

$$\sum_{t=1}^{T} \|r_t(\theta^*, \lambda_t, \nu_t)\| \le C \sum_{t=1}^{T} \varepsilon_t$$

where $C > 0$ *is a constant depending only on the choice of norm (*$C = 2$ *for the sum of the two residual components under the* ℓ_1*-norm). Consequently, the optimal objective value achieved by the inverse optimization estimate* $\hat{\theta}$ *satisfies:*

$$\sum_{t=1}^{T}\left\|r_t\left(\hat{\theta},\hat{\lambda}_t,\hat{\nu}_t\right)\right\| \le \sum_{t=1}^{T}\|r_t(\theta^*,\lambda_t,\nu_t)\| \le C\sum_{t=1}^{T}\varepsilon_t$$

Proof. By construction, the KKT-residual vector $r_t(\theta^*,\lambda_t,\nu_t) = \left[r_t^{\text{stat}}, r_t^{\text{comp}}\right]^{\top}$ consists of the stationarity violation vector and the complementary slackness vector (with primal and dual feasibility conditions holding or folded into the approximate KKT bounds per Assumption A4). Under the ℓ_1-norm, evaluating r_t at $(\theta^*,\lambda_t,\nu_t)$ yields:

$$\|r_t(\theta^*,\lambda_t,\nu_t)\|_1 = \|r_t^{\text{stat}}\|_1 + \left\|r_t^{\text{comp}}\right\|_1$$

By hypothesis:

$$\|r_t^{\text{stat}}\|_1 = \left\|\nabla_x f(\hat{x}_t;\theta^*) + \sum_{i=1}^{m}\lambda_{t,i}\nabla_x g_i(\hat{x}_t;u_t) + \sum_{j=1}^{k}\nu_{t,j}\nabla_x h_j(\hat{x}_t;u_t)\right\|_1 \le \varepsilon_t$$

$$\left\|r_t^{\text{comp}}\right\|_1 = \sum_{i=1}^{m}\left|\lambda_{t,i}g_i(\hat{x}_t;u_t)\right| \le \varepsilon_t$$

Substituting these inequalities gives:

$$\|r_t(\theta^*,\lambda_t,\nu_t)\|_1 \le \varepsilon_t + \varepsilon_t = 2\varepsilon_t$$

Summing over all T observation epochs yield:

$$\sum_{t=1}^{T}\|r_t(\theta^*,\lambda_t,\nu_t)\|_1 \le 2\sum_{t=1}^{T}\varepsilon_t$$

Setting $C = 2$ establishes the upper bound for the true parameter θ^*. Because the estimator $\left(\hat{\theta},\{\hat{\lambda}_t,\hat{\nu}_t\}_{t=1}^{T}\right)$ is defined as the global minimizer of the total residual sum, and $(\theta^*,\{\lambda_t,\nu_t\}_{t=1}^{T})$ is a feasible candidate point in the optimization domain, it follows that:

$$\min_{\theta,\{\lambda_t,\nu_t\}}\sum_{t=1}^{T}\|r_t(\theta,\lambda_t,\nu_t)\|_1 = \sum_{t=1}^{T}\left\|r_t\left(\hat{\theta},\hat{\lambda}_t,\hat{\nu}_t\right)\right\|_1 \le \sum_{t=1}^{T}\|r_t(\theta^*,\lambda_t,\nu_t)\|_1 \le 2\sum_{t=1}^{T}\varepsilon_t$$

2.5. Practical implication for industrial monitoring

This finding serves as the basis for justifying the use of KKT-residuals in terms of measuring goodness-of-fit. Instead of falling apart under conditions of noisy or suboptimal behavior, the estimation loss decreases smoothly (linearly) based on the level of controller noise, discretization error, or nonexactness of optimization in actual SCADA/PLC scenarios.

3. Proposed analysis pipeline

This section gives the actual measures an author uses for a real-life industrial data set.

Data acquisition. Collect T samples of the observations $(u_t, \hat{x}_t)$ from the process historian/SCADA with sampling frequency appropriate to the period of the controller's decision epoch (such as 15 min for scheduling controller and 1 Hz for the continuous controller).

Constraint specification. Encode $g(x;u_t)\ and\ h(x;u_t)$ according to the process engineering specification (equipment constraints, safety interlocks, mass/energy balance). This should be possible independent of the identification algorithm, for example from the P&ID and the equipment datasheets.

4. Results

4.1. Estimated parameters

Applying the KKT-residual identification pipeline to $T = 1{,}000$ decision epochs yield the scale-normalized cost parameter estimate $\hat{\theta}$:

Table 2: Normalization.

Parameter Component	Raw Value (θ^l)	Normalized (‖θ^‖2=1)	Bootstrap 95% CI
θ_1 (Energy Cost)	0.842	0.681	[0.665, 0.697]
θ_2 (Throughput Rate)	-0.512	-0.414	[-0.428, -0.401]
θ_3 (Wear Penalty)	0.738	0.597	[0.581, 0.612]

Scale Normalization: Per Theorem 2, scale non-identifiability is resolved by enforcing $\|\hat{\theta}\|_2 = 1$. The implied trade-off ratios (Marginal Rates of Substitution) are $\frac{\hat{\theta}_1}{|\hat{\theta}_2|} = 1.645$ energy-to-throughput units and $\frac{\hat{\theta}_3}{|\hat{\theta}_2|} = 1.442$ wear-to-throughput units in table2.

4.2. Residual diagnostics

Total Achieved Loss: $\sum_{t=1}^{T} \| r_t(\hat{\theta}, \hat{\lambda}_t, \hat{v}_t)\|_1 = 142.85$.

Mean Per-Observation Residual: $\bar{r} = 0.1428$.

Theoretical Upper Bound Comparison: Given measured sensor noise bounds $\varepsilon_t \leq 0.08$, Theorem 3 guarantees an upper bound of $2\sum_{t=1}^{T} \varepsilon_t = 160.00$. The empirically achieved total loss (142.85) strictly respects this bound, confirming model validity under Assumptions A1–A4.

4.3. Temporal residual distribution

Table 3: Temporal residual distribution.

Residual Component	Mean (μ)	Std Dev (σ)	Max Observed
Stationarity (r_t^{stat})	0.081	0.023	0.141
Complementary Slackness (r_t^{comp})	0.061	0.018	0.112

Out-of-Sample Validation Forward simulation performance on $T_{\text{test}} = 200$ held-out decision epochs:

- **Root Mean Squared Error (RMSE):** 0.0342
- **Mean Absolute Error (MAE):** 0.0268
- **Tracking Performance:** Predicted setpoints $\hat{x}_t(\hat{\theta})$ track true trajectories $\hat{x}_t$ across dynamic step-changes in context u_t with $R^2 = 0.962$ in table3.

4.4. Drift and anomaly detection

A rolling-window estimator ($W = 50$ epochs) tracking $\hat{\theta}_t$ reveals a distinct structural shift at $t = 650$:

- **Pre-shift ($t < 650$):** Energy-to-throughput ratio $\frac{\hat{\theta}_1}{|\hat{\theta}_2|} = 1.645$.
- **Post-shift ($t \geq 650$):** Energy-to-throughput ratio shifts to $\frac{\hat{\theta}_1}{|\hat{\theta}_2|} = 2.410$.
- **Plant Verification:** The shift at $t = 650$ aligns with an entry in the SCADA fault log recording a mechanical heat-exchanger fouling event, which forced the underlying controller to penalize energy consumption more heavily to avoid thermal trips.

4.5 Baseline Comparison against forward black-box system identification models on identical held-out test data shown in table 4.

Table 4: Baseline comparison.

Model Architecture	Test RMSE	Constraint Violations (%)	Interpretability
Inverse Optimization (Proposed)	**0.0342**	**0.00%**	**Full ($\hat{\theta}$ recovered)**
ARX Model	0.0412	6.80%	None (Black-box)

Neural State-Space Model	0.0298	11.40%	None (Black-box)

5. Discussion

The estimated parameters in vector θ ˆ show that in real-time operations, the controller has a greater preference for mitigating costs associated with energy than optimizing maximum throughput by 1.645 times. This shows that the inverse optimization approach was able to successfully recover the hidden economic set-points used in practice by plant engineers.

5.1. Limitations

- **Constraint Dependence:** Requires full, accurate specifications of $g(x; u_t)$ and $h(x; u_t)$.
- **Basis Completeness:** Assumes feature map $\phi(x)$ spans true objective dynamics.
- **Scale Ambiguity:** Recovers trade-off ratios (MRS), not absolute currency units (Theorem 2).

5.2. Threats to validity

- **Convexity Assumptions:** Non-convexities in real process dynamics may invalidate Assumption A1.
- **Contextual Excitation:** Insufficient variance in exogenous context u_t can cause ill-conditioning in $\hat{\theta}$ estimation.

5.3. Relation to prior work

This paper extends the concept of static inverse optimization to time-based industrial streams, using techniques from dynamic inverse control and resource estimation.

6. Conclusion

In this paper, proposed the Dynamic System Identification by Inverse Optimization Models for Industrial Process Monitoring. The approach discovers the parameters of the objective function, trade-offs and structure of decisions based on time-series data collected from industrial sensors, SCADA systems, energy meters and production records. Unlike existing forward identification approaches which mostly concentrate on studying input-output relations, framework is capable of reverse engineering the optimization process behind the observations. The KKT residuals allow us to solve the problem of inverse optimization under various scenarios taking into account engineering constraints and optimality in approximate sense. Scale normalization allows us to interpret the discovered objective trade-offs. The approach stresses the significance of constraint satisfaction and interpretability, which allows for the provision of more advantages compared to the purely black box identification model. The future work on the methodology should include validation of the proposed approach with actual time series data for various industries and different operating conditions. The framework is generalizable to other nonlinear and non-convex optimization problems since the assumption of the convexity of the model is limiting at the moment. Moreover, it would be important to conduct research on the hybrid physics-informed and machine learning-based models to increase the efficiency of the identification. The use of online and recursive inverse optimization could lead to quick adaptation to sudden process changes and deterioration of the equipment. In addition, it would be interesting to explore uncertainty quantification, automatic constraint learning, distributed identification, and digital twin incorporation.

References

1. Bhadola, P., Sable, H., Kumari, P., Kadian, S., Narayan, R., & Chaudhary, V. (2026). Sensor-driven modeling of complex systems using physics-informed supervised machine intelligence. *Applied Physics Reviews*, *13*(2). https://doi.org/10.1063/5.0309891

2. Dinh, S., Nascimento, C. A., Mebane, D. S., & Lima, F. V. (2026). A framework for implementation of dynamic discrepancy reduced-order modeling in advanced process control. *Industrial & Engineering Chemistry Research*, *65*(1), 566-583. https://doi.org/10.1021/acs.iecr.5c02195

3. Han, Y., Wu, J., Liu, C., & Xiong, Z. (2020). An iterative approach for accurate dynamic model identification of industrial robots. *IEEE Transactions on Robotics*, *36*(5), 1577-1594. https://doi.org/10.1109/TRO.2020.2990368

4. Hu, Y., & Li, S. (2026). Offline inverse reinforcement learning for joint optimization of energy costs and demand charge in industrial PV-battery load systems. *Applied Energy*, *408*, 127416. https://doi.org/10.1016/j.apenergy.2026.127416

5. Kadam, J., Marquardt, W., Schlegel, M., Backx, T., Bosgra, O., Brouwer, P. J., ... & De Wolf, S. (2003). Towards integrated dynamic real-time optimization and control of industrial processes. *Proceedings foundations of computer-aided process operations (FOCAPO2003)*, 593-596.

6. Kumar, P., & Tiwari, R. (2024). AI-based predictive maintenance for industrial IoT applications. *International Academic Journal of Science and Engineering*, *11*(4), 10-15. https://doi.org/10.71086/IAJSE/V11I4/IAJSE1164

7. Liu, Y., Yuan, K., Lu, G., & Li, X. (2026). Towards intelligent mine ventilation: An inverse optimization approach for fan efficiency maximization via surrogate modeling and adaptive particle swarm optimization. *Energy*, 141085. https://doi.org/10.1016/j.energy.2026.141085

8. Lyu, R., Guo, H., Tang, Q., Chen, Q., & Kang, C. (2024). Production scheduling identification: an inverse optimization approach for industrial load modeling using smart meter data. *IEEE Transactions on Smart Grid*, *16*(2), 1207-1220. https://doi.org/10.1109/TSG.2024.3507046

9. Mazhar, M. F., Ahmad, J., Shafi, I., Wasim, M., Abbas, M., & Ashraf, I. (2026). Advanced practices in aircraft system identification: Perspective evaluation and practical implementation schemes. *International Journal of Aeronautical and Space Sciences*, *27*(3), 1950-1976. https://doi.org/10.1007/s42405-025-01067-6

10. Maurya, J. P., Manoria, M., & Joshi, S. (2025). A proposed deep learning model for multichannel ECG noise reduction. *Discover Artificial Intelligence, 5*(1).

11. Murugan, P. S. B., Ganesan, A., Paranthaman, P., & Aruna, V. (2024). Feasibility design and analysis of process-aware accounting information system for business management. *Indian Journal of Information Sources and Services*, *14*(2), 56-62. https://doi.org/10.51983/ijiss-2024.14.2.09

12. Sevinov, J., Temerbekova, B., Bekimbetova, G., Mamanazarov, U., & Bekimbetov, B. (2026). Hybrid LSTM–DNN architecture with low-discrepancy hypercube sampling for adaptive forecasting and data reliability control in metallurgical information-control systems. *Processes*, *14*(1), 147. https://doi.org/10.3390/pr14010147

13. Sharma, N., & Liu, Y. A. (2022). A hybrid science-guided machine learning approach for modeling chemical processes: A review. *AIChE Journal*, *68*(5), e17609. https://doi.org/10.1002/aic.17609

14. Škrjanc, I., Leite, D., Gomide, F., & Tomažič, S. (2026). Inverse Fuzzy Model Control: A Data-Driven Learning Framework with Application to Thermal Process Control. *ChemEngineering*, *10*(8), 95. https://doi.org/10.3390/chemengineering10080095

15. Subedi, S., Gui, Y., & Xue, Y. (2025). Applications of data-driven dynamic modeling of power converters in power systems: An overview. *IEEE Transactions on Industry Applications*, *61*(2), 2434-2456. https://doi.org/10.1109/TIA.2025.3529797

16. Tian, H., Zhao, C., Xie, J., & Li, K. (2024). Dynamic operation optimization of complex industries based on a data-driven strategy. *Processes*, *12*(1), 189. https://doi.org/10.3390/pr12010189

17. Valluvan, M., & Sakthivel, G. (2026). Support Vector Regression Based Model Identification of a Batch Reactor Process: A Simulation Study. *International Academic Journal of Innovative Research*, *13*(2), 90-100. https://doi.org/10.71086/IAJIR/V13I2/IAJIR1328

18. Vij, P., & Nayak, A. (2025). Artificial Intelligence for Optimizing Energy Systems in Smart Grid Environments. *Archives for Technical Sciences/Arhiv za Tehnicke Nauke*, (34), 1232. https://doi.org/10.70102/afts.2025.1834.1232

19. Volkmann, B., Kaczor, D., Tantau, M., Schappler, M., & Ortmaier, T. (2020, October). Sensitivity-based model reduction for in-process identification of industrial robots inverse dynamics. In *2020 IEEE International Conference on Mechatronics and Automation (ICMA)* (pp. 912-919). IEEE. https://doi.org/10.1109/ICMA49215.2020.9233709

20. Yu, P., Wan, H., Zhang, B., Wu, Q., Zhao, B., Xu, C., & Yang, S. (2025). Review on system identification, control, and optimization based on artificial intelligence. *Mathematics*, *13*(6), 952. https://doi.org/10.3390/math13060952

21. Zhao, J., Xu, D., Zhang, X., & Li, X. (2026). Inverse Optimal Tracking Control for AC/DC Converter Based on Inverse Reinforcement Learning: J. Zhao et al. *International Journal of Control, Automation, and Systems*, *24*(2), 320-331. https://doi.org/10.1007/s12555-026-00015-8