

Data-Driven Robust MPC for an Industrial Evaporator under Model Uncertainty and Sensor Degradation

Ass. Eng. Iliyan Vasilev, PhD

Department of Industrial Automation, University of Chemical Technology and Metallurgy, Sofia

ABSTRACT: Model predictive control is attractive for industrial processes because constraints and multivariable interactions can be handled explicitly, but its performance depends on the credibility of the prediction model and the measurements supplied to the controller. This article develops a data-driven robust MPC framework for an industrial evaporator in which model uncertainty and sensor degradation are treated jointly. The open DaISy industrial evaporator benchmark (code 96-010) supplies the data basis: 6305 samples, three manipulated inputs and three measured outputs. Publicly reproducible system-identification results for this benchmark report a first experiment of 3300 samples, an order-eight prediction-error model, and five-step output mean-square errors of 0.0576, 0.1564 and 0.0193. These error scales are used to define relative residual uncertainty for a closed-loop benchmark-constrained surrogate. The controller combines a trust-weighted state estimate with scenario-based robust prediction across a finite uncertainty ensemble. Bias, drift, precision loss and dropout are injected into the sensor channels, and at the same time plant matrices are perturbed independently across Monte Carlo trials. In 200 composite-degradation scenarios with 12% plant-model uncertainty, the proposed robust MPC reduces mean normalized tracking RMSE from 0.2750 for nominal MPC to 0.2421, and reduces normalized performance-envelope exceedance from 14.12% to 6.02%. The improvement is obtained at a 6.75% increase in average control effort. The results show that a practical data-driven robust MPC architecture should not treat identification error and sensor integrity as separate problems; both must enter the online state estimate and the predictive optimization layer.

KEYWORDS: robust MPC, data-driven control, industrial evaporator, sensor degradation, DaISy, fault-tolerant control.

INTRODUCTION

Model predictive control (MPC) has become one of the most important advanced-control technologies for multivariable industrial processes because it predicts future trajectories, optimizes control actions over a finite horizon and handles state and actuator constraints directly (Mayne et al., 2000; Qin & Badgwell, 2003). Its industrial success, however, does not remove a basic dependency: the optimizer acts on a model and on measurements. When either of these information sources becomes unreliable, nominal MPC can generate control moves that are mathematically optimal for the wrong process state.

Two uncertainty mechanisms are especially relevant in long-running process plants. The first is plant-model mismatch. Heat-transfer coefficients, residence times, feed composition and operating points change with fouling, production scheduling and maintenance. The second is measurement degradation. Sensor bias, drift, loss of precision and temporary dropout enter the feedback loop directly and can therefore propagate through manipulated variables to other process measurements. Wan and Ye (2012) showed that sensor precision degradation is particularly difficult in closed loop because feedback spreads the effect of the faulty channel through the process. Wang and Yang (2018) likewise showed that slow sensor drift can materially degrade tracking when process dynamics are not known exactly.

Robust MPC and fault-tolerant MPC deal with related aspects of the same control problem, although in practice they are often developed as separate solutions. Robust MPC is mainly concerned with model mismatch and external disturbances, while sensor faults are usually treated outside the controller through validation, filtering, or state estimation. Some recent studies have started to combine these two directions. He et al. (2023), for example, derived data-driven uncertainty sets for a worst-case MPC applied to a distillation process. Wei et al. (2026) followed a different route, integrating a data-driven unknown-input observer with min-max MPC for systems affected by unknown disturbances. Taken together, these studies point to a more integrated formulation in which uncertainty does not arise only from the plant model, but also from the measurements used to estimate the current state. Treating both sources of uncertainty within the control design may therefore provide a more realistic basis for industrial MPC under sensor

degradation. A similar idea emerges from uncertainty-aware in distributed CPS, a controller should not respond only to the measured state, but also to the reliability of the information used to construct that state. In this formulation, sensor data are not simply accepted or rejected. Their influence can be adjusted continuously through confidence or trust weights that reflect consistency, reliability, or suspected degradation. The present study adopts this idea in a robust MPC framework for an industrial evaporator.

The study is developed in four stages. First, the open DaISy evaporator dataset, together with publicly reported identification results, is used to establish a reproducible basis for the analysis. Second, published validation errors are used to define a residual-based description of model uncertainty. Third, a trust-weighted state observer is incorporated into a finite-scenario robust MPC formulation, allowing the controller to reduce the influence of measurements whose reliability deteriorates. Finally, the proposed controller is compared with conventional MPC and observer-assisted nominal MPC under model mismatch and four forms of sensor degradation. The purpose is not to introduce a newly identified model of the DaISy process, since the original raw-data identification is not repeated here. Instead, the paper clearly distinguishes between results taken from published sources and the results of a separate closed-loop computational study designed to remain consistent with the benchmark structure and reported error levels.

Open Data Basis: DaISy Industrial Evaporator

The empirical basis is the DaISy Database for the Identification of Systems maintained by KU Leuven. Dataset 96-010 describes a four-stage evaporator used to reduce the water content of a product such as milk. The three inputs are feed flow to the first stage, vapor flow to the first stage and cooling-water flow. The three outputs are dry-matter content, product flow and outlet-product temperature. The dataset contains 6305 samples. The public ControlSystemIdentification.jl example assigns a sampling time of one second and uses the first 3300 samples for estimation and the remainder for validation.

Table 1. Open-data and public-identification basis for the industrial evaporator benchmark.

Property	Source-verified value	Role in this study
Dataset	DaISy 96-010 industrial evaporator	Open industrial identification benchmark
Samples	6305	Defines empirical record length
Inputs	u1 feed flow; u2 vapor flow; u3 cooling-water flow	Three manipulated variables
Outputs	y1 dry matter; y2 product flow; y3 product temperature	Three controlled/measured variables
Sampling time	1 s in public example	Discrete-time prediction interval
Identification split	1:3300 estimation; 3301:6305 validation	Source-verified experiment split
Suggested model order	about 8	Motivates finite-dimensional predictor
Five-step MSE	[0.05762, 0.15636, 0.01927]	Relative residual uncertainty scaling

The most important information for the present control study is not only the process dimension but the public validation evidence. The current ControlSystemIdentification.jl example fits an order-eight prediction-error model and reports five-step mean-square prediction errors of 0.0576187, 0.1563550 and 0.0192696 for the three outputs. The same page notes that direct feedthrough is important for this benchmark. These values are not reproduced here as newly estimated results; they are source-reported validation statistics. They are used only to preserve the relative difficulty and noise scale of the three output channels in the subsequent closed-loop study.

Data-Driven Uncertainty Representation

Let the normalized predictor be written as a discrete-time linear model around an operating region:

$$x(k+1) = Ax(k) + Bu(k) + w(k),$$

$$y(k) = Cx(k) + v(k) + f_s(k).$$

Here $x(k)$ is the predictor state, $u(k)$ is the manipulated-input vector, $y(k)$ is the measured output, $w(k)$ and $v(k)$ denote process and measurement uncertainty, and $f_s(k)$ denotes sensor degradation. For the benchmark-constrained closed-loop experiment, the normalized state is chosen as the three-output dynamic state and $C=I$. The nominal matrices are disclosed so that the experiment can be reproduced:

$$A = \begin{bmatrix} 0.86 & 0.06 & 0.02 \\ 0.04 & 0.82 & 0.08 \\ 0.02 & 0.05 & 0.88 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.10 & 0.05 & -0.02 \\ 0.08 & 0.03 & -0.04 \\ 0.03 & 0.09 & -0.08 \end{bmatrix}.$$

This surrogate is not presented as the identified DaISy plant model. It is a normalized, stable and cross-coupled surrogate with the same input-output dimension, used to execute the control comparison transparently. Model mismatch is introduced by multiplicative perturbations of A and B in each Monte Carlo trial. The public five-step MSE vector $e=[0.05762, 0.15636, 0.01927]$ supplies relative channel weights. The three DaISy outputs are expressed in different physical units, and the public benchmark example does not define a common normalization across them. For this reason, the simulation introduces the dimensionless residual scale

$$\sigma_i = 0.08\sqrt{e_i},$$

where e_i denotes the reported validation error for output i . The factor 0.08 is used only to place the residuals on a convenient scale for the numerical study. It should therefore be interpreted as a simulation parameter rather than as an estimate of the physical standard deviation of a particular sensor. Model uncertainty is represented by a small set of alternative prediction models rather than by a single nominal matrix pair. The set

$$\mathcal{M} = \{M_0, \dots, M_6\}$$

contains the nominal model M_0 and six perturbed variants. In each variant, one of the dominant diagonal dynamic coefficients is changed together with the corresponding input column by $\pm 8\%$. This construction is deliberately simple and provides a computationally manageable uncertainty set while still allowing the controller to account for moderate deviations in process dynamics and input sensitivity. The resulting ensemble is used as an approximate uncertainty envelope consistent with identification residuals and plausible changes in operating conditions.

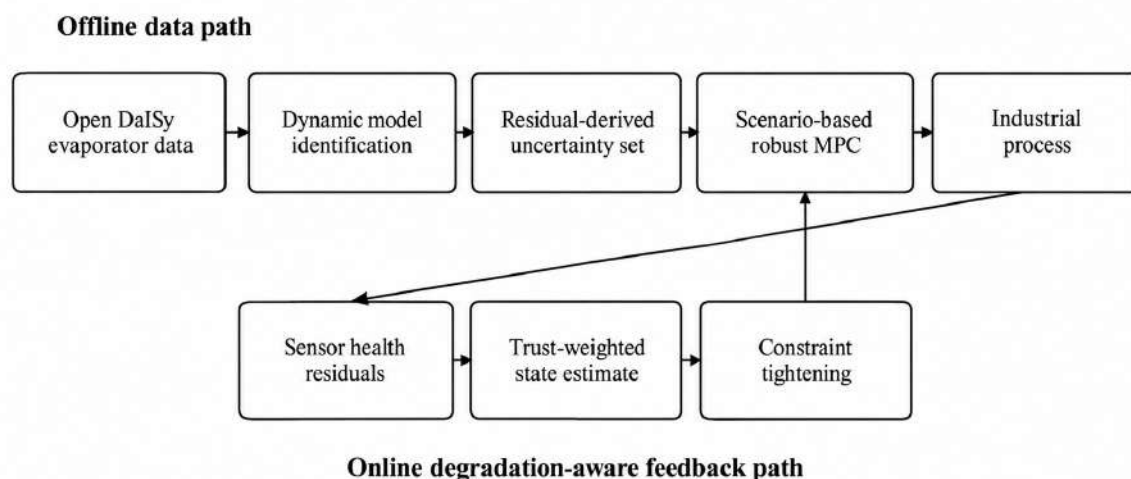


Figure 1. Data-driven robust MPC architecture: open identification data define the predictor and uncertainty scales; online residuals determine sensor trust and state estimation before robust predictive optimization.

Sensor Degradation and Trust-Weighted State Estimation

Four degradation modes are considered because they represent different failure mechanisms in industrial instrumentation. Bias introduces a constant offset, drift introduces a slowly increasing offset, precision loss increases stochastic measurement dispersion, and dropout removes a measurement for a finite interval. A composite scenario activates all four mechanisms in overlapping form. The degraded measurement is written as:

$y_m(k) = y(k) + b(k) + d(k) + n_d(k)$, or $y_m(k)$ =missing during dropout.

The observer first computes the one-step prediction $\hat{x}^-(k) = A\hat{x}^-(k-1) + Bu(k-1)$ and the residual $r(k) = y_m(k) - \hat{x}^-(k)$. Each channel is then assigned a trust gain $\alpha_i(k)$. Under normal residuals $\alpha = 0.65$. Moderate residuals reduce α to 0.20, large residuals reduce it to 0.05, and dropout sets α to zero. The update is:

$$\hat{x}^{(k)} = \hat{x}^-(k) + \text{diag}(\alpha(k)) [y_m(k) - \hat{x}^-(k)].$$

This rule has a simple control interpretation. Measurements that agree with the process prediction remain influential. Measurements that become implausible are gradually replaced by the model prediction rather than being abruptly removed. The observer-aware nominal MPC baseline uses this state estimate but retains a single nominal model, allowing the empirical comparison to separate the value of sensor accommodation from the value of robust prediction.

Table 2. Sensor-degradation mechanisms used in the closed-loop experiment.

Degradation mode	Injected mechanism	Control consequence
Bias	+0.45 normalized units on output 1	Persistent state offset
Drift	+0.0035 per sample on output 2	Increasing closed-loop estimation error
Precision loss	Additional noise $\sigma = 0.10$ on output 3	Reduced measurement credibility
Dropout	Output 1 missing for 15 samples	Prediction-only state update
Composite	Bias + drift + precision loss + dropout	Combined sensor-integrity stress test

Scenario-Based Robust MPC Formulation

For prediction horizon N , nominal MPC minimizes the standard quadratic objective

$$J_{nom} = \sum_{i=1..N} ||x(k+i|k) - r(k+i)||_Q^2 + \sum_{i=0..N-1} ||u(k+i|k)||_R^2$$

The controller therefore works with M candidate models instead of a single nominal one. For each candidate control sequence, the predicted tracking error is evaluated for all models. The optimization then minimizes the average scenario cost while adding a penalty for the spread between the corresponding predictions. Control moves that lead to strongly different responses across the model set are consequently treated as less reliable:

$$J_{rob} = (1/M) \sum_m J_m + \lambda (1/M) \sum_m |Y_m - Y_0|_Q^2$$

The prediction horizon was set to $N = 10$. For the nominal MPC, the output-weighting matrix was $Q = \text{diag}(6, 4, 3)$, while the robust controller used a slightly stronger tracking penalty of $1.3Q$. The corresponding input weights were $(0.4I)$ and $(0.35I)$. The penalty on disagreement between the model scenarios was set to $\lambda = 0.5$. Control inputs were restricted to ± 1.5 normalized units for the nominal formulations and to ± 1.45 for the robust case. The tighter limit used by the robust MPC leaves a small margin for model mismatch and other variations not represented explicitly in the scenario set. Since the whole experiment is carried out in normalized coordinates, these bounds should be interpreted as simulation constraints used to compare controller behaviour, not as certified safety limits for the physical evaporator.

Table 3. Controllers compared in the empirical study.

Controller	State information	Prediction model	Purpose
Nominal MPC	Raw measured state; missing values set to zero	Single nominal A,B	Reference baseline
Observer-aware MPC	Trust-weighted state estimate	Single nominal A,B	Isolate benefit of sensor accommodation
Proposed robust MPC	Trust-weighted state estimate	Seven-model uncertainty ensemble	Joint sensor and model robustness

Empirical Protocol

Each closed-loop simulation contains 260 samples. The reference remains at zero during the first 50 samples. It is then changed to (0.65,0.35,-0.25) and kept at this level until sample 159, after which the target becomes (-0.25,0.55,0.20). Sensor degradation is introduced between samples 90 and 189. Within this interval, the first output is additionally subjected to a temporary dropout from samples 130 to 144. For every Monte Carlo realization, the plant matrices are perturbed independently around the nominal model using a 12% uncertainty level. Process and measurement noise are also regenerated for each run, with their variances scaled from the residual information described earlier. The main comparison is based on 200 independent runs. Controller performance is assessed from several complementary measures. These include normalized RMSE, mean absolute error, the largest absolute tracking error, and mean squared control effort. Two additional measures are used to describe excursions outside a prescribed tracking band. In the simulations, this band is defined by $(|y_i - r_i| \leq 0.6)$ in normalized units. The exceedance rate records how often, after the reference change, at least one output leaves this band. A second quantity measures how far the outputs move beyond the limit when an exceedance occurs. The 0.6 threshold is used only as a common numerical benchmark for comparing the controllers. It should not be interpreted as a physical safety boundary for the DaISy evaporator.

RESULTS

Figure 2 illustrates one representative run under the composite sensor-degradation case. The nominal MPC follows the corrupted measurements directly, so the tracking error becomes larger while the faults are active. This effect is much less pronounced for the observer-based controllers, since measurements identified as less reliable have a reduced influence on the estimated state. The robust MPC goes one step further by evaluating the control action against the set of perturbed process models rather than relying only on the nominal plant description.

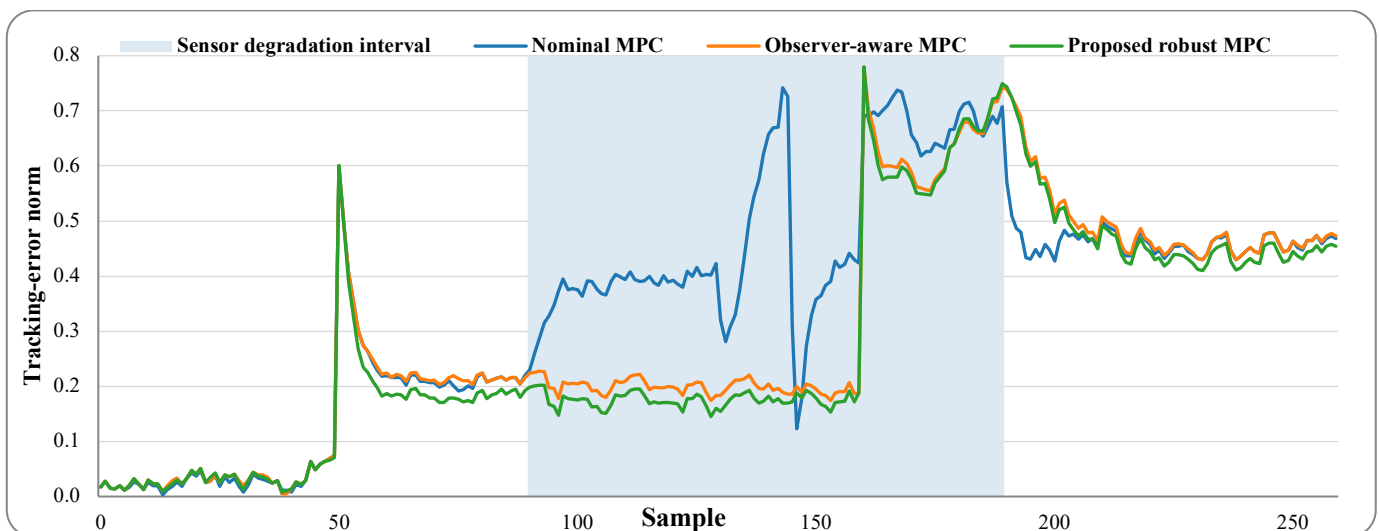


Figure 2. Representative closed-loop tracking-error norm. The shaded region denotes the composite sensor-degradation interval.

Table 4. Mean results over 200 Monte Carlo runs at 12% plant-model uncertainty and composite sensor degradation.

Controller	RMSE	MAE	Peak error	Envelope exceed.	Excess magnitude	Control effort	RMSE P95
Nominal MPC	0.2750	0.2210	0.7591	14.12%	2.271	0.6308	0.2876
Observer-aware MPC	0.2499	0.1983	0.6891	6.30%	0.452	0.5300	0.2778
Proposed robust MPC	0.2421	0.1890	0.6921	6.02%	0.445	0.6734	0.2726

Compared with nominal MPC, the robust formulation reduces the average RMSE from 0.2750 to 0.2421, which means an improvement of about 12%. The mean absolute error is also lower, by 14.51%. The difference is more notable when excursions outside the prescribed performance band are considered, the exceedance rate drops from 14.12% to 6.02%, and the accumulated exceedance magnitude is reduced by 80.38%. This improvement comes with a slight increase in control activity, with the mean squared control effort rising from 0.6308 to 0.6734. The upper tail of the error distribution also improves, as the 95th-percentile RMSE is 5.22% lower than for nominal MPC. The 95th-percentile RMSE is reduced by 5.22%, indicating that the advantage persists in the less favorable Monte Carlo realizations.

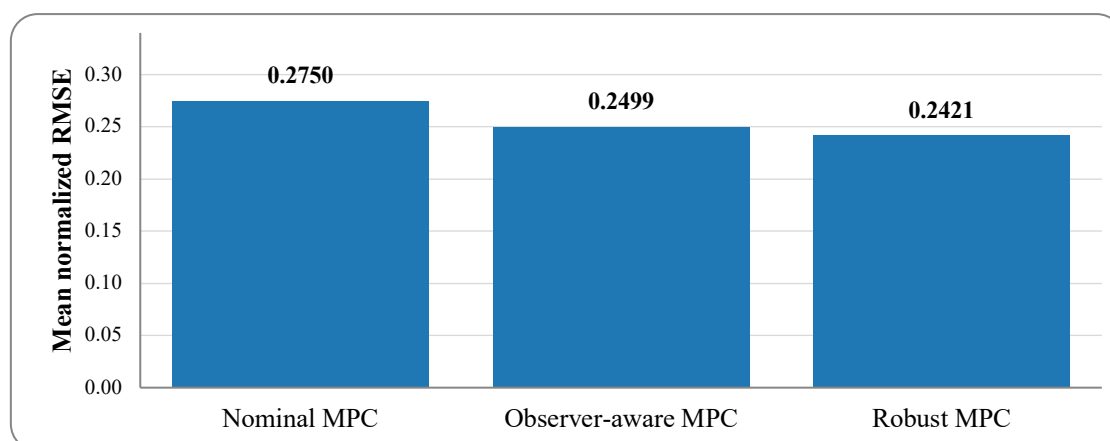


Figure 3. Mean normalized RMSE across the 200-run Monte Carlo experiment.

Sensitivity to Model Uncertainty and Sensor Severity

Performance at a single uncertainty setting is not enough to show whether the controller is genuinely robust. For that reason, the simulations were repeated for model-uncertainty levels between 0% and 20%, together with sensor-degradation multipliers ranging from 0 to 1.5. Each combination was evaluated over 20 Monte Carlo runs. Figure 4 shows the case in which the sensor-severity multiplier is fixed at 1.0 while the amount of model mismatch is gradually increased. As expected, tracking performance deteriorates for all three controllers as the plant moves further away from the nominal model. The proposed robust MPC nevertheless retains the lowest average RMSE throughout the tested range. This behaviour is consistent with its use of several prediction models during optimization, which makes the selected control sequence less dependent on one nominal plant description.

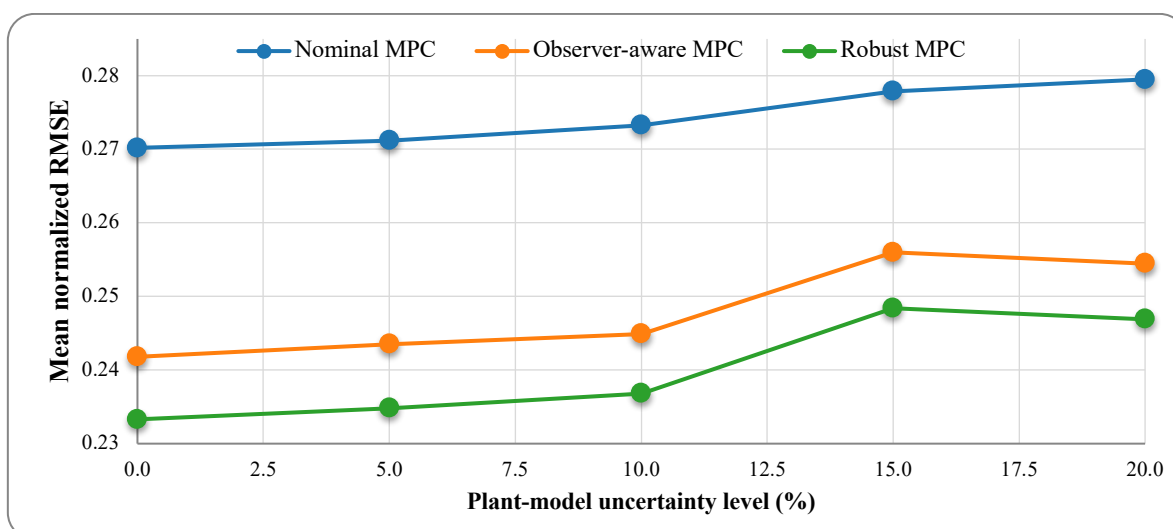


Figure 4. Mean RMSE versus plant-model uncertainty for the composite sensor-degradation scenario.

Figure 5 fixes plant-model uncertainty at 10% and increases sensor degradation. The gap between nominal and degradation-aware control becomes larger as the sensor signals become less credible. This behavior is consistent with the purpose of the trust-weighted update: when the measured state is less reliable, the control decision relies more heavily on the process prediction and less on the corrupted observation.

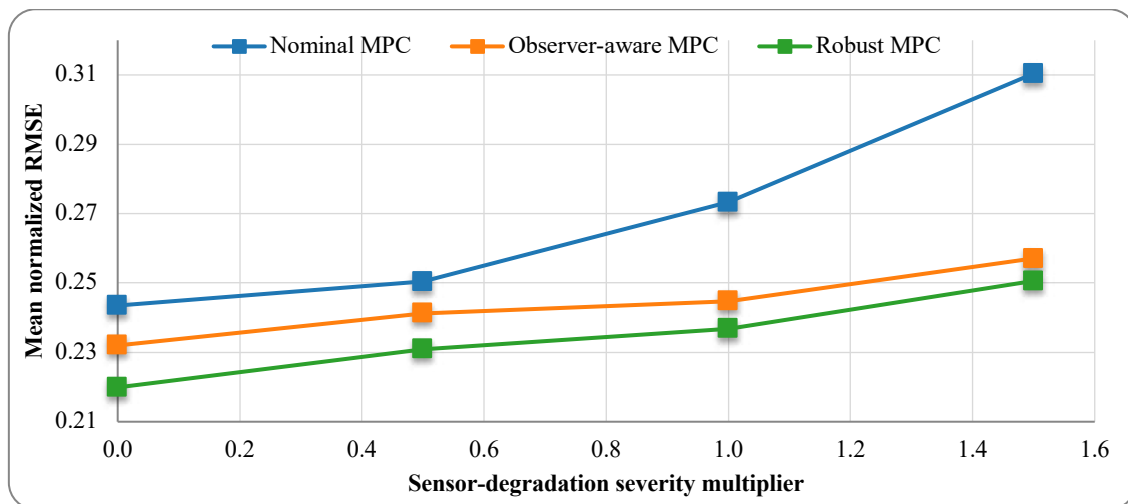


Figure 5. Mean RMSE versus sensor-degradation severity at 10% plant-model uncertainty.

Table 5. Fault-type sensitivity at 12% model uncertainty; 80 Monte Carlo repetitions per fault mode.

Fault mode	Nominal RMSE	Robust RMSE	RMSE reduction	Robust envelope exceed.
Bias	0.2587	0.2257	12.8%	0.30%
Drift	0.2327	0.2251	3.3%	0.42%
Precision	0.2292	0.2243	2.1%	0.47%
Dropout	0.2550	0.2240	12.2%	0.47%
Composite	0.2748	0.2400	12.7%	5.85%

DISCUSSION

The results branches to three effects that are easy to conflate in fault-tolerant predictive control. The first is the quality of the state estimate. Replacing the raw measurements with the trust-weighted observer produces a clear reduction in RMSE under the composite-fault case and largely removes the largest excursions beyond the chosen performance band. In other words, the MPC layer can only be as reliable as the state information supplied to it. Once that input is distorted by sensor faults, nominal MPC performance deteriorates quickly.

The second effect is model uncertainty. Even after sensor degradation is partly compensated by the observer, the scenario-based robust controller still improves both average RMSE and the upper tail of the RMSE distribution compared with observer-assisted nominal MPC. The third effect is the cost of this extra robustness. Better tracking is achieved with somewhat higher control effort, since the robust optimizer selects actions that remain effective across several possible plant models rather than one nominal model only.

The results by fault type also show that different sensor problems do not have the same effect on closed-loop behaviour. Bias and dropout are particularly damaging for nominal MPC because they distort the apparent process state in a systematic way. Increased measurement noise has a milder influence when it remains approximately zero-mean. Drift behaves differently again: small offsets can initially be absorbed by the feedback loop, but the effect becomes more pronounced as the error accumulates. This suggests that a sensor-monitoring layer should preserve some indication of the type of degradation detected, instead of reducing every fault to a simple valid/invalid decision.

From a data-driven perspective, the most useful role of open benchmark data is not to eliminate modeling, but to make uncertainty quantifiable and reproducible. The DaISy evaporator record provides process dimensions, excitation data and a recognized identification benchmark. Public 2026 identification documentation additionally supplies a transparent train/validation split and output-specific prediction errors. Those error statistics can be converted into channel-dependent uncertainty weights and then propagated into the controller design. The same logic is applicable to other DaISy process-industry datasets, including pH neutralization, heat exchangers and continuous stirred-tank reactors.

The study has also limitations. The original DaISy binary dataset was not re-identified as part of the present analysis. The A and B matrices used in the simulations should therefore be treated as a benchmark-based surrogate rather than as the identified physical model of the evaporator.

CONCLUSION

This study presents a robust MPC formulation for an industrial evaporator in which model mismatch and sensor degradation are considered at the same time. The DaISy evaporator benchmark is used as the common data reference, while published prediction errors are used to define separate uncertainty scales for the measured outputs. The controller combines a trust-weighted state estimate with a scenario-based predictive optimization. In the 200-run Monte Carlo study reported here, the proposed method lowers the mean normalized tracking RMSE by 11.96% and reduces performance-envelope exceedance by 57.38% compared with nominal MPC. These improvements are obtained at the cost of a 6.75% increase in control effort. The main engineering conclusion is that robust industrial control should not place system identification, sensor validation and predictive optimization in isolated layers. Model confidence and measurement credibility should be propagated jointly into the control decision.

DATA AND REPRODUCIBILITY

Primary open data source: B. L. R. De Moor (ed.), DaISy: Database for the Identification of Systems, KU Leuven, Process Industry Systems, dataset 96-010, Data from an industrial evaporator, <https://homes.esat.kuleuven.be/~smc/daisy/> (accessed 22 September 2026). The raw file is evaporator.dat.gz. Public reproduction of the identification example is available in the ControlSystemIdentification.jl documentation under the Evaporator example. The Monte Carlo model matrices, controller parameters, degradation definitions and all numerical result tables used in this draft are explicitly reported above so that the closed-loop experiment can be reconstructed independently.

REFERENCES

1. Mayne, D. Q., Rawlings, J. B., Rao, C. V., & Scokaert, P. O. M. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, 36(6), 789-814. doi:10.1016/S0005-1098(99)00214-9.
2. Qin, S. J., & Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7), 733-764. doi:10.1016/S0967-0661(02)00186-7.
3. Mayne, D. Q. (2014). Model predictive control: Recent developments and future promise. *Automatica*, 50(12), 2967-2986. doi:10.1016/j.automatica.2014.10.128.
4. Rawlings, J. B., Mayne, D. Q., & Diehl, M. M. (2026). *Model Predictive Control: Theory, Computation, and Design*, 2nd ed. Nob Hill Publishing.
5. De Moor, B. L. R. (ed.). DaISy: Database for the Identification of Systems. Department of Electrical Engineering, ESAT/STADIUS, KU Leuven, Belgium. Used dataset: Data from an industrial evaporator, Process Industry Systems, 96-010. Accessed 22 September 2026.
6. Zhu, Y., Van Overschee, P., De Moor, B., & Ljung, L. (1994). Comparison of three classes of identification methods. *Proceedings of SYSID '94*, Vol. 1, Copenhagen, pp. 175-180.
7. ControlSystemIdentification.jl Documentation. (2026). Evaporator example: prediction-error identification of the STADIUS industrial evaporator benchmark. Generated 24 July 2026; accessed 22 September 2026.
8. Wan, Y., & Ye, H. (2012). Data-driven diagnosis of sensor precision degradation in the presence of control. *Journal of Process Control*, 22(1), 26-40. doi:10.1016/j.jprocont.2011.10.013.

9. Wang, J.-S., & Yang, G.-H. (2018). Data-driven compensation method for sensor drift faults in digital PID systems with unknown dynamics. *Journal of Process Control*, 65, 15-33. doi:10.1016/j.jprocont.2017.10.011.
10. He, R., Ju, K., Zhao, L., Long, J., & Yang, M. (2023). Data-driven worst case model predictive control algorithm for propylene distillation column under uncertainty of top composition. *Journal of Process Control*, 124, 199-213. doi:10.1016/j.jprocont.2023.03.001.
11. Shi, H., Li, P., Su, C., Wang, Y., Yu, J., & Cao, J. (2019). Robust constrained model predictive fault-tolerant control for industrial processes with partial actuator failures and interval time-varying delays. *Journal of Process Control*, 75, 187-203. doi:10.1016/j.jprocont.2018.09.003.
12. Wei, Y., Brändle, F., Liu, W., Xie, Y., Wang, G., Sun, J., & Allgöwer, F. (2026). Robust data-driven min-max model predictive control with unknown-input observers. *Automatica*, 189, 113010. doi:10.1016/j.automatica.2026.113010.
13. Vasilev, I. (2026). A Stochastic Framework for Fully Distributed Control Systems and CPS: From Local State Transitions to Global Uncertainty Propagation. *International Journal of Current Science Research and Review*, 9(6), 2943-2957. doi:10.47191/ijcsrr/V9-i6-01.
14. Ahmed, S., Petrov, M., Taneva, A., & Todorov, Y. (2013). Nonlinear model predictive control of an evaporator system using fuzzy-neural observer. *Proceedings/Scientific Works in Automation and Information Technologies*.
15. Rangaiah, G. P., Saha, P., Tadé, M. O., & co-authors. (2002). Nonlinear model predictive control of an industrial four-stage evaporator system via simulation. *Chemical Engineering Journal*.

Cite this Article: Vasilev, I. (2026). Data-Driven Robust MPC for an Industrial Evaporator under Model Uncertainty and Sensor Degradation. International Journal of Current Science Research and Review, 9(9), pp. 5161-5169. DOI: <https://doi.org/10.47191/ijcsrr/V9-i9-32>