

Time-Dependent Action-based Prediction of Parameters for Operator Decision Support in Abnormal Situations at Nuclear Power Plants

Hyojin Kim^a, Jonghyun Kim^{a*}

^aDepartment of Nuclear and Quantum Engineering, Korea Advanced Institute of Science and Technology (KAIST),
291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea

*Corresponding author: jonghyun.kim@kaist.ac.kr

***Keywords :** time-dependent action-conditioned prediction, uncertainty quantification, operator decision support

1. Introduction

Maintaining safe operation of a nuclear power plant (NPP) requires operators to manage abnormal situations in which the plant departs from normal operation. In some of these situations, only a limited time remains before a reactor-protection setpoint is reached, and operators must execute the actions prescribed in abnormal operating procedures (AOPs) to arrest the abnormal progression and stabilize the plant. The AOP specifies the required actions and their sequence, but it does not quantify how a delay in an action changes the future plant response or how much time remains before the action loses its effectiveness. For timely mitigation, operators must therefore judge whether the prescribed action can prevent an undesirable event, such as a reactor trip, and how soon it must be performed [1].

To help operators make this judgment, previous studies have developed methods for predicting the future plant response. Within Endsley's three-level framework, such prediction corresponds to Level 3 situation awareness, which enables operators to anticipate future plant states and determine an appropriate course of action [2]. Performing this projection mentally is demanding because NPPs are nonlinear, multivariate, and tightly coupled systems and because time pressure degrades decision quality. However, most existing prediction models forecast the plant response from past measurements alone and therefore cannot represent the operator actions that will reshape it [3, 4]. The few action-aware studies condition the forecast on recently observed device states or on a static representation of a selected action. Conditioning a prediction on a time-resolved future-action trajectory that describes how a planned action unfolds over time has received limited attention in the NPP literature [5]. The existing formulations are consequently unable to distinguish the plant responses that arise when the same action is initiated at different times or executed through different command sequences.

An operator-support prediction method should be evaluated by its forecast accuracy and by its ability to support the determination of an action plan that achieves the procedural objective. This study therefore proposes a time-dependent action-conditioned prediction framework that forecasts the multivariate plant response over a prediction horizon from the observed plant state and a planned future action trajectory, enabling plans that

differ in initiation time, magnitude, and temporal sequence to be compared from the same observed state before execution.

This study makes four contributions. First, NPP response forecasting is formulated as time-dependent action-conditioned prediction in which the intended operator action is represented as a time-resolved trajectory over the prediction horizon. Second, reversible instance normalization (RevIN), an inverted Transformer (iTransformer) encoder, and a causal Transformer-temporal convolutional network (TCN) decoder are integrated into a multivariate architecture that conditions the long-horizon forecast on the planned action trajectory. Third, a heteroscedastic output head and a deep ensemble are combined to quantify the data-conditional predictive spread and the model uncertainty, respectively. Fourth, the decision-support use of the framework is evaluated in an abnormal-situation case study of a chemical and volume control system (CVCS) letdown-line leak using data collected with a full-scope simulator of a 1400 MWe-class pressurized water reactor (PWR). The case study compares alternative action plans from the same observed state and identifies the latest evaluated isolation timing that averts the reactor trip.

2. Time-dependent action-conditioned prediction framework

The framework comprises forward RevIN, the iTransformer encoder, the causal Transformer-TCN decoder, the heteroscedastic output head, inverse RevIN, and the deep ensemble.

2.1 Problem formulation

Let the past plant-measurement window be $X_p \in \mathbb{R}^{L \times C_x}$, where L is the observation-window length and C_x is the number of input plant variables. The planned future action trajectory is $U_f \in \mathbb{R}^{H \times C_u}$, where H is the prediction horizon and C_u is the number of action variables. The prediction target is the future plant response $Y_f \in \mathbb{R}^{H \times C_y}$, where C_y is the number of target plant parameters.

As expressed in Eq. (1), the framework is a mapping f_θ , parameterized by the trainable weights θ , that returns from X_p and U_f the predictive mean $\hat{\mu}_f$ and variance $\hat{\sigma}_f^2$

of every output parameter at every horizon step after inverse RevIN:

$$(\hat{\mu}_f, \hat{\sigma}_f^2) = f_\theta(X_p, U_f), \quad \hat{\mu}_f, \hat{\sigma}_f^2 \in \mathbb{R}^{H \times C_y} \quad (1)$$

The complete action trajectory U_f is supplied in a single forward pass. Within this pass, the causal decoder constrains the prediction at horizon step h to depend only on the action steps from 1 through h . An action planned after step h therefore cannot affect an earlier prediction. This temporal-causality constraint prevents future-action leakage, whereas the explicit action input enables the trained framework to evaluate alternative action plans repeatedly from the same observed state.

2.2 Algorithm design

Fig. 1 illustrates the overall architecture, which specializes the encoder to the variable axis and the decoder to the time axis. Forward RevIN first standardizes each plant variable within the observation window and applies a learnable variable-specific affine transformation, reducing sensitivity to window-dependent changes in level and scale. The iTransformer encoder then embeds the history of each input variable as one token and applies self-attention across the variable axis. This representation captures dependencies among plant quantities associated with inventory, pressure, temperature, and flow.

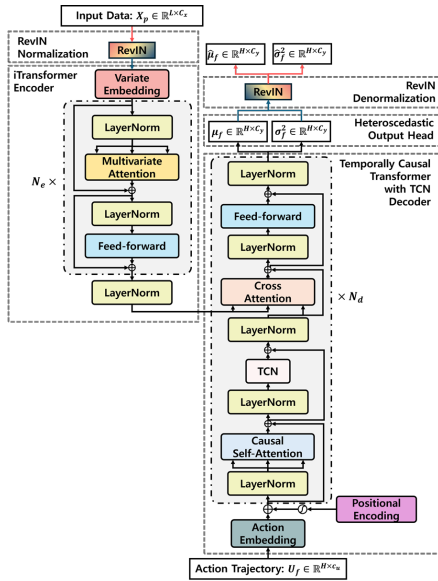


Fig. 1. Architecture of the proposed framework.

The decoder represents the temporal structure of the planned action trajectory and enforces temporal causality. Masked self-attention captures long-range dependencies across the prediction horizon. A dilated causal TCN extracts local and multiscale temporal patterns associated with changes in the action commands. Cross-attention to the encoder memory then conditions each horizon-step representation on the observed plant state. This ordering first organizes the temporal action features and then retrieves the corresponding plant-state context. The

causal attention mask and the left-padded causal convolutions ensure that the prediction at each horizon step is independent of actions planned for later steps.

2.3 Uncertainty quantification

The heteroscedastic output head maps the decoded representation to a diagonal Gaussian for every target parameter and horizon step. It returns the predictive mean $\tilde{\mu}_f$ and positive variance $\tilde{\sigma}_f^2$ in the RevIN-transformed space. Inverse RevIN subsequently restores these predictions to the common min-max scale, giving $\hat{\mu}_f$ and $\hat{\sigma}_f^2$.

Each framework instance was trained by minimizing the Gaussian negative log-likelihood (NLL) on the restored scale. Omitting the additive constant $\ln(2\pi)/2$, the loss is:

$$\mathcal{L}_{NLL} = \frac{1}{NHC_y} \sum_{w=1}^N \sum_{h=1}^H \sum_{c=1}^{C_y} \frac{1}{2} \left[\ln \hat{\sigma}_{w,h,c}^2 + \frac{(y_{w,h,c} - \hat{\mu}_{w,h,c})^2}{\hat{\sigma}_{w,h,c}^2} \right] \quad (2)$$

where N is the number of training windows and w , h , and c index the training window, horizon step, and target parameter, respectively. The target, mean, and variance in Eq. (2) are all expressed on the common min-max scale after inverse RevIN.

To estimate model uncertainty, S independently initialized framework instances are combined as a deep ensemble. Because each member has its own learned RevIN parameters, the member distributions are combined only after inverse RevIN. Eqs. (3) and Eq. (4) show the ensemble mean and total predictive variance, respectively:

$$\tilde{\mu}_f = \frac{1}{S} \sum_{s=1}^S \hat{\mu}_f^{(s)}, \quad (3)$$

$$\hat{\sigma}_{tot}^2 = \frac{1}{S} \sum_{s=1}^S \left[\hat{\sigma}_f^{2,(s)} + (\hat{\mu}_f^{(s)} - \tilde{\mu}_f)^2 \right]. \quad (4)$$

The ensemble mean provides the point prediction, and the total predictive standard deviation $\hat{\sigma}_{tot}$ determines the prediction interval.

3. Data collection

The target abnormal situation was a letdown-line leak simulated with a full-scope generic PWR simulator developed using the 3KEYMASTER code. This scenario was selected because mitigation effectiveness depends on action timing. The leak decreases the reactor coolant system (RCS) inventory, which reduces the pressurizer level and RCS pressure and drives the departure-from-nucleate-boiling ratio (DNBR) toward its low trip setpoint. Without mitigation, the low-DNBR signal initiates an automatic reactor trip within approximately 4–6 min of leak onset, depending on the leak severity. The corresponding AOP prescribes two mitigating actions: isolating the letdown line and regulating the charging flow to restore the pressurizer level.

The framework encodes the letdown-line isolation as a binary trajectory that changes from the open state to the

closed state at the scheduled isolation time. The charging action is represented by two trajectories: the control mode, which is either automatic or manual, and the commanded valve position. Together, these three action channels form a single time-resolved action trajectory (U_f), which represents plans with different isolation times, valve openings, and command sequences, including successive manual adjustments of the charging valve. This encoding corresponds to the control variables through which the two mitigating actions are implemented in the simulator.

The database comprises 1,367 scenarios divided among five scenario types, as summarized in Table I. Leak severity serves as the disturbance factor, whereas isolation timing and the charging-control trajectory serve as the operator-action factors. The no-action reference enumerates the severity range without mitigation. The two isolation types vary the isolation timing on a fixed post-leak grid and a severity-specific trip-boundary grid. The static and dynamic charging types apply a single valve-position command and a sequence of command changes, respectively. These two types were sampled using Latin hypercube sampling (LHS) because full enumeration of their combined condition spaces was impractical. Each scenario spans 60 min and was recorded at intervals of 1 s.

Table I: Scenario families of the collected database.

Scenario family	Varied conditions	Sampling	Count
No-action reference	Leak severity	Grid	41
Letdown isolation	Severity and isolation time	Grid	36
Trip-boundary isolation	Severity and isolation time near the trip	Grid	54
Static charging control	Severity, isolation time, and fixed commanded opening	LHS	300
Dynamic charging control	Severity, isolation time, and multi-step commanded openings	LHS	936
Total			1,367

4. Model training and results

4.1 Implementation and training

The 20 output parameters comprise pressurizer pressure and level, DNBR, reactor coolant temperatures, and the principal CVCS flows. Input variables were screened from the full simulator variable set with a k-nearest-neighbor mutual information estimator, retaining 355 variables. The horizon spans 40 min at 10 s resolution (240 steps).

Sliding the forecast origin in 1 s steps produced 610 origins per scenario and 833,870 samples in total, divided into 533,750 for training, 133,590 for validation, and 166,530 for testing. The split was performed at the scenario level such that no scenario contributes windows to more than one subset, preventing leakage between

overlapping samples. All variables were min–max scaled to the unit interval using padded bounds computed from the training set only.

Hyperparameters were tuned with Optuna under a Hyperband pruner. Training used a one-cycle learning-rate schedule and an exponential moving average of the weights, and model selection and early stopping were governed by the validation NLL. The ensemble comprises $S = 5$ instances trained from independent random initializations. Accuracy is reported by the normalized mean absolute error (nMAE), and the normalized mean absolute percentage error (nMAPE), computed on the min–max normalized scale to be comparable across output parameters; lower nMAE and nMAPE indicate more accurate point predictions. Predictive intervals are assessed by the prediction interval coverage probability (PICP), i.e., the fraction of samples falling inside the interval, at the 95 % nominal level and the mean prediction interval width (MPIW); the uncertainty is well estimated when PICP is close to the nominal level while MPIW remains small, meaning the intervals are both reliable and sharp. Table II shows the validation performance of the ensemble members and the deep ensemble.

Table II: Validation performance of the ensemble members and the deep ensemble.

Model	nMAE	nMAPE	PICP	MPIW
Seed 1	0.00156	0.356	0.891	0.00871
Seed 2	0.00157	0.376	0.902	0.00992
Seed 3	0.00166	0.397	0.890	0.00931
Seed 4	0.00167	0.394	0.889	0.00899
Seed 5	0.00163	0.388	0.898	0.00984
Ensemble	0.00153	0.359	0.920	0.0100

4.2 Predictive performance

The test set contains 273 held-out scenarios (166,530 samples), none of which contributed to training or validation. On this set, the deep ensemble achieved an nMAE of 0.00145, and an nMAPE of 0.333 %. Its 95 % prediction intervals attained a PICP of 0.925 at an MPIW of 0.0100. These results indicate high point accuracy and predictive intervals with modest undercoverage relative to the nominal coverage of 0.95. Across the 20 output parameters, the per-parameter nMAE and nMAPE range from 0.00009 to 0.0063 and from 0.08 % to 1.25 %, respectively, with log power the least accurately predicted parameter on both measures; even for this parameter the mean error remains below 0.7 % of its normalized range.

Fig. 2 presents the action-conditioned predictions for a representative dynamic-command test scenario. The ensemble mean closely follows the simulated response over the 40 min horizon and reproduces the transients induced by the successive valve commands; in particular, the charging flow drops to zero whenever the charging valve is fully closed in both the simulated and predicted trajectories, which overlap almost exactly. The simulated trajectories remain largely within the 95 % prediction intervals.

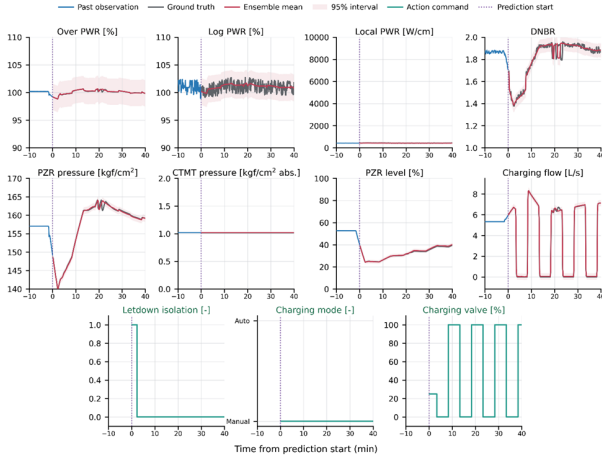


Fig. 2. Prediction example for a test scenario with the commanded action trajectory.

4.3 Delayed-isolation analysis for trip avoidance

The practical value of action-conditioned prediction lies in comparing alternative plans from the same observed state. Fig. 3 compares the predicted responses to letdown-isolation delays of 0.5–4.0 min with the no-isolation response for a 49 gpm letdown line leak.

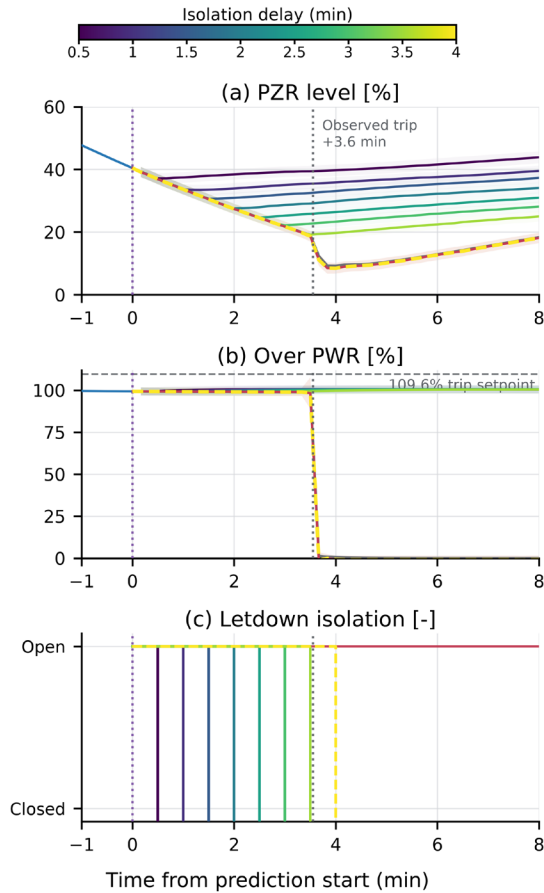


Fig. 3. Predicted results for isolation delays of 0.5–4.0 min compared with the no-isolation response for a 49 gpm letdown line leak.

As the isolation delay increased, the predicted pressurizer level decreased further before recovery. Delays through 3.5 min were predicted to avert the reactor trip, whereas the 4.0-min isolation, scheduled after the observed trip time of 3.6 min, followed the no-isolation response and failed to avert the trip. Thus, 3.5 min was the latest evaluated isolation time predicted to avert the trip.

5. Conclusions

This study proposed a time-dependent action-conditioned prediction framework that forecasts the multivariate plant response from the observed plant state and a planned future action trajectory. On a letdown-line leak database collected with a full-scope 1400 MWe-class PWR simulator, the deep ensemble achieved a test nMAE of 0.00145 and an R^2 of 0.997, with a PICP of 0.925 at the nominal coverage of 0.95. By comparing delayed-isolation plans from the same observed state, the framework identified the latest evaluated isolation timing that averts the reactor trip, indicating that the formulation can support action-plan comparison under the modeled conditions.

ACKNOWLEDGEMENT

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (grant number: RS-2022-00144042) and Korea Hydro & Nuclear Power Co., Ltd. (No. 2025-AI-001-01).

REFERENCES

- [1] J. Yu, H. S. Park, and Y.-G. Lee, “Estimating allowable operator action time for mitigating multiple-failure accidents in Korean Advanced Power Reactor,” *International Journal of Energy Research*, vol. 2024, 1249111, 2024.
- [2] M. R. Endsley, “Toward a theory of situation awareness in dynamic systems,” *Human Factors*, vol. 37, no. 1, pp. 32–64, 1995.
- [3] H. Kim and J. Kim, “Long-term prediction of safety parameters with uncertainty estimation in emergency situations at nuclear power plants,” *Nuclear Engineering and Technology*, vol. 55, no. 5, pp. 1630–1643, 2023.
- [4] J. Bae, G. Kim, and S. J. Lee, “Real-time prediction of nuclear power plant parameter trends following operator actions,” *Expert Systems with Applications*, vol. 186, 115848, 2021.
- [5] H. Kim, S. Lee, and J. Kim, “Mitigation action-dependent prediction of nuclear plant parameters for severe accident management via conditional generative model,” *Reliability Engineering & System Safety*, vol. 276, 112892, 2026.