scientific reports

OPEN

First application of data assimilation-based control to fusion plasma

Yuya Morishita[1✉], Sadayoshi Murakami[1], Naoki Kenmochi[2,5], Hisamichi Funaba[2], Ichihiro Yamada[2,5], Yoshinori Mizuno[2], Kazuki Nagahara[2], Hideo Nuga[2], Ryosuke Seki[2], Masayuki Yokoyama[2,5], Genta Ueno[3,4,6] & Masaki Osakabe[2,5]

Magnetic fusion plasmas, which are complex systems comprising numerous interacting elements, have large uncertainties. Therefore, future fusion reactors require prediction-based advanced control systems with an adaptive system model and control estimation robust to uncertainties in the model and observations. To address this challenge, we introduced a control approach based on data assimilation (DA), which describes the system model adaptation and control estimation based on the state probability distribution. The first implementation of a DA-based control system was achieved at the Large Helical Device to control the high temperature plasma. The experimental results indicate that the control system enhanced the predictive capability using real-time observations and adjusted the electron cyclotron heating power for a target temperature. The DA-based control system provides a flexible platform for advanced control in future fusion reactors.

Magnetic confinement fusion is a promising next-generation power source. To generate fusion-based power, the fusion plasma should be heated to attain and maintain a good state of confinement. However, fusion plasma is a typical complex system in which various physical quantities interact with each other to determine the overall behavior[1]. Therefore, a large number of variables should be considered for controlling the fusion plasma. An adaptive system (predictive) model is required to account for the latent variables that are difficult to observe or model (e.g., wall conditions[2]) along with uncertain elements related to stability, abrupt termination events, and energy and particle transport. These challenges exist in common with the control of complex systems.

Conventional controllers (e.g., proportional-integral-derivative (PID) controllers) exert simple control over systems with few variables. However, various controllers and observers should be combined to control nonlinear complex systems with large uncertainties and numerous variables. Consequently, the control system becomes extremely complicated and considerable effort is required to construct a well-coordinated control system. For such nonlinear systems, adaptive model predictive control is computationally expensive and difficult to achieve[3].

Aiming to overcome these challenges, we developed a control system based on data assimilation (DA) techniques[4,5]. In the DA framework, the state of the system is expressed as a probability distribution of the state vector (state distribution), which enables the calculation of the time evolution and conditional probability distribution (information imposition) of the system. Here, the state vector expresses the state of the target system and includes physical quantities and model parameters as components. In general, DA is employed to estimate unobservable variables and models[6–9] and optimize the model parameters for higher prediction accuracy[10–12]. This study used DA for state estimation (including the model parameters) and control estimation to achieve adaptive model predictive control. This approach facilitates the representation of multiple state variables, observations, and control objects within a single framework, thereby simplifying the control system. In addition, control constraints and domain knowledge can be incorporated into the system model and state variables. The proposed DA-based control system describes the system model adaptation and control estimation from a unified perspective based on state distribution and offers a flexible control platform for the advanced control of fusion reactors.

Recently, methods implementing data-driven approaches for predicting and controlling fusion plasmas, including the use of deep and reinforcement learning, have been investigated[13–15]. The data-driven approach

[1]Department of Nuclear Engineering, Kyoto University, Kyoto, Japan. [2]National Institute for Fusion Science, National Institutes of Natural Sciences, Toki, Japan. [3]The Institute of Statistical Mathematics, Research Organization of Information and Systems, Tachikawa, Japan. [4]The Joint Support-Center for Data Science Research, Research Organization of Information and Systems, Tachikawa, Japan. [5]Department of Fusion Science, The Graduate University for Advanced Studies, SOKENDAI, Toki, Japan. [6]Statistical Science Program, The Graduate University for Advanced Studies, SOKENDAI, Tachikawa, Japan. ✉email: morishita.yuya.7x@kyoto-u.ac.jp

is an inductive approach that constructs predictive and control models based on the given data. It facilitates flexible modeling and highly accurate approximations of target systems within the training data. Contrarily, it requires a large amount of training data and often has low prediction accuracy outside of the data. In addition, the training process is performed again for different scenarios. The physics-based approach, which is a deductive approach based on governing equations, is understandable and tractable by humans. However, it often fails to approximate real systems, particularly complex systems, with high accuracy. The DA-based control approach has the properties of both the data-driven and physics-based approaches. It can improve the prediction performance of the physical model using observation data and capture the internal state and control processes of the system. The proposed approach also enables the construction of a device-independent control system.

The proposed DA-based control system was implemented to control the plasma in the Large Helical Device (LHD), which is one of the largest superconducting plasma confinement devices in the world[16,17], located in Japan. The effectiveness of this system was verified by performing a simple control experiment. The helical device was desirable for the first demonstration experiment because of its stable magnetic field. In the experiment, the control system estimated the power of the electron cyclotron heating (ECH) required to deliver the target electron temperature. To bridge the gap between the system model and actual behavior, the sequentially observed electron temperature and density profiles were assimilated into the system model. This study is the first to demonstrate real-time DA and adaptive predictive control based on the DA of fusion plasmas. The proposed control scheme is applicable to fusion plasmas and other complex systems (e.g., traffic control, infectious disease control measures, and river level control), and is expected to advance complex system control.

Methods

Figure 1 shows a conceptual diagram of the DA-based control system. This system has a simple structure comprising a DA system, heating and fuel supply equipment, and measurement devices. This section describes the DA-based control system constructed for the LHD plasma control.

DA framework for adaptive predictive control

The DA system, which is the core component of the control system, computes the probability distribution of the state vector (state distribution), which is defined at time t as $\mathbf{x}_t = (\tilde{\mathbf{x}}_t^{\mathrm{T}}, \mathbf{u}_t^{\mathrm{T}})^{\mathrm{T}}$, where vector $\tilde{\mathbf{x}}_t$ is the part of $\mathbf{x}$ that includes the system state and model parameters, and vector $\mathbf{u}$ is the control input that determines the time evolution of the system (see Table 1 for an example). The superscript "T" denotes the transpose of a vector and Δt_z is the time interval for estimating the control inputs. The DA system optimizes $\tilde{\mathbf{x}}$ based on the observed information $\mathbf{y}$ (adaptation) and estimates $\mathbf{u}$ based on the predicted state distribution and target state $\mathbf{z}$ (predictive control).
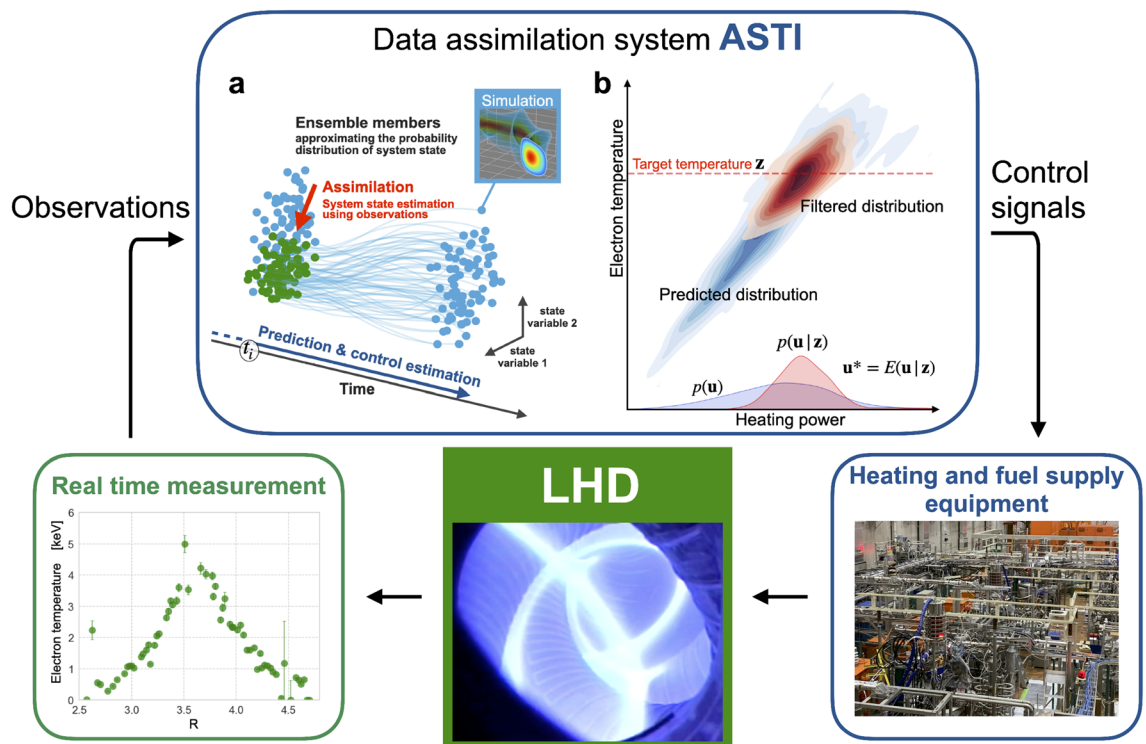

Figure 1. Overview of the DA-based control system for fusion plasmas. (**a**) The state distribution in the DA system (ASTI) is expressed as an ensemble. The parallel time evolution of each ensemble member approximates that of the state distribution. (**b**) The control estimation is performed by assimilating the target state into the predicted state distribution.

Consider a control problem in which the control input is adjusted at each time interval Δt_z and the system state is observed as the vector $\mathbf{y}$ at each time interval Δt_y. For simplicity, assume that $\Delta t_y = n\Delta t_z$ ($n \in \mathbf{N}$) and use the time notation $(i,j) = t_{i,j} = t_0 + i\Delta t_y + j\Delta t_z$ and $t_i = t_{i,0}$, where t_0 denotes the initial time. The DA framework for adaptive predictive control[18] is based on the following state-space model:

$$\mathbf{x}_{(i,j+1)} = f_{(i,j+1)}(\mathbf{x}_{(i,j)}, \mathbf{v}_{(i,j+1)}), \tag{1}$$

$$\mathbf{z}_{(i,j)} = H^{\mathbf{z}}\mathbf{x}_{(i,j)} + \mathbf{w}^{\mathbf{z}}_{(i,j)}, \tag{2}$$

$$\mathbf{u}^*_{(i,j)} = H^{\mathbf{u}}\mathbf{x}_{(i,j)} + \mathbf{w}^{\mathbf{u}}_{(i,j)}, \tag{3}$$

$$\mathbf{y}_i = H^{\mathbf{y}}\mathbf{x}_{(i,0)} + \mathbf{w}^{\mathbf{y}}_i. \tag{4}$$

Equation (1) represents the system model that describes the time evolution of the system $\mathbf{x}_{(i,j)} \rightarrow \mathbf{x}_{(i,j+1)}$, considering the effect of system noise $\mathbf{v}_{(i,j+1)}$. Assume that the value of $\mathbf{u}_t$ is constant in the prediction interval Δt_z, i.e.,

$$\mathbf{u}_{(i,j+1)} = \mathbf{u}_{(i,j)} + \mathbf{v}^{\mathbf{u}}_{(i,j+1)}, \tag{5}$$

$$\tilde{\mathbf{x}}_{(i,j+1)} = \tilde{f}_{(i,j+1)}(\tilde{\mathbf{x}}_{(i,j)}, \mathbf{u}_{(i,j+1)}, \tilde{\mathbf{v}}_{(i,j+1)}). \tag{6}$$

Here, the system noise for the control input $\mathbf{v}^{\mathbf{u}}_{(i,j+1)}$ is added to $\mathbf{u}_{(i,j)}$ before calculating the time-evolution. Equations (2)–(4) represent the relationship between the state vector $\mathbf{x}_{(i,j)}$ and vectors $\mathbf{z}_{(i,j)}$, $\mathbf{u}^*_{(i,j)}$, and $\mathbf{y}_i$ based on the noises $\mathbf{w}^{\mathbf{z}}_{(i,j)}$, $\mathbf{w}^{\mathbf{u}}_{(i,j)}$, and $\mathbf{w}^{\mathbf{y}}_i$, respectively. The matrices $H^{\mathbf{z}}$, $H^{\mathbf{u}}$, and $H^{\mathbf{y}}$ are linear operators for projecting the state vector in each corresponding space. The system noise $\mathbf{v}_{(i,j+1)}$ is assumed to follow a Gaussian distribution with zero mean and covariance matrix $Q_{(i,j+1)}$, i.e., $\mathbf{v}_{(i,j+1)} \sim N(\mathbf{0}, Q_{(i,j+1)})$. Similarly, $\mathbf{w}^{\mathbf{z}}_{(i,j)}$, $\mathbf{w}^{\mathbf{u}}_{(i,j)}$, and $\mathbf{w}^{\mathbf{y}}_i$ are assumed to follow the probability distributions $N(\mathbf{0}, R^{\mathbf{z}}_{(i,j)})$, $N(\mathbf{0}, R^{\mathbf{u}}_{(i,j)})$, and $N(\mathbf{0}, R^{\mathbf{y}}_i)$, respectively. The priority of the controlled and observed variables can be adjusted using $\mathbf{w}^{\mathbf{z}}_{(i,j)}$ and $\mathbf{w}^{\mathbf{y}}_i$.

In the DA system, the state distribution is approximated by an ensemble using parallel computing. Each ensemble member represents a simulation with slightly different conditions (e.g., initial conditions, model parameters, and control inputs). In this experiment, considering the time required for computation and communication, the prediction and control estimations were performed up to Δt_y ahead of the real time at $n = 3$ ($\Delta t_y = 3\Delta t_z$). The following are the computational steps for the adaptive predictive control:

- Prediction

$$p\left(\mathbf{x}_{(i,j)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)}\right) \rightarrow p\left(\mathbf{x}_{(i,j+1)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)}\right). \tag{7}$$

- z-filter

$$p\left(\mathbf{x}_{(i,j+1)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)}\right) \rightarrow p\left(\mathbf{u}_{(i,j+1)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)}, \mathbf{z}_{(i,j+1)}\right). \tag{8}$$

- u-filter

$$p\left(\mathbf{x}_{(i,j+1)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)}\right) \rightarrow p\left(\mathbf{x}_{(i,j+1)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j+1)}\right). \tag{9}$$

- y-filter

$$p\left(\mathbf{x}_{(i+1,0)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i+1,0)}\right) \rightarrow p\left(\mathbf{x}_{(i+1,0)} \mid \mathbf{y}_{0:i}, \mathbf{u}^*_{(0,1):(i+1,0)}\right). \tag{10}$$

here the subscript $t_1 : t_2$ denotes all the timings in $[t_1, t_2]$. Given the distribution $p(\mathbf{x}_{(i,j)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)})$, the prediction step computes $p(\mathbf{x}_{(i,j+1)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)})$, the distribution Δt_z ahead, based on the system model. This step can be performed by computing the time evolution of each ensemble member approximating the state distribution.

The z- and u-filter are the control estimation steps, and the y-filter is the adaptation step. These three steps are based on the Bayesian filter, which calculates the conditional probability distribution $p(\mathbf{x} \mid \xi)$ using the distribution $p(\mathbf{x})$ and the model representing the relationship between $\mathbf{x}$ and the imposed information ξ, $\xi = h(\mathbf{x}) + \mathbf{w}$. Here, h represents the relationship between ξ and $\mathbf{x}$ and $\mathbf{w}$ represents the associated noise. Generally, ξ denotes the observation data, and the Bayesian filter can assimilate the observed information into the state distribution. Kalman filter (EnKF)[19] and particle filter (PF)[20] are typical computational methods for the Bayesian filter that use an ensemble.

The z-filter step estimates the control input by assimilating the target $\mathbf{z}_{(i,j+1)}$ into the predicted distribution based on the Bayesian filter using Eq. (2),

$$p\left(\mathbf{x}_{(i,j+1)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)}\right) \rightarrow p\left(\mathbf{x}_{(i,j+1)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)}, \mathbf{z}_{(i,j+1)}\right), \tag{11}$$

and performing marginalization

$$p\left(\mathbf{u}_{(i,j+1)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)}, \mathbf{z}_{(i,j+1)}\right)$$

$$= \int p\left(\mathbf{x}_{(i,j+1)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)}, \mathbf{z}_{(i,j+1)}\right) d\tilde{\mathbf{x}}_{(i,j+1)}. \tag{12}$$

We used EnKF as the computational method for the Bayesian filter. This marginalization can be performed by removing $\tilde{\mathbf{x}}$ from the ensemble of the distribution $p(\mathbf{x}_{(i,j+1)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,j)}, \mathbf{z}_{(i,j+1)})$. The control input $\mathbf{u}^*_{(i,j+1)}$ is obtained as the expected value of the z-filtered distribution. The u-filter is applied by assimilating the estimated control input $\mathbf{u}^*_{(i,j+1)}$ into the predicted distribution using the Bayesian filter based on Eq. (3). The control estimation process can proceed by repeating the following three steps: prediction→z-filter→u-filter.

The y-filter corresponds to the adaptive process and suppresses the uncertainties in the system model by assimilating the observations into the state distribution. Because the observation times t_i differ from the latest u-filtered distribution $t_{i+1,0}$, the y-filter is executed by assimilating $\mathbf{y}_i$ into the joint distribution of the state vectors at two time points. The ensemble approximating the joint distribution $p(\mathbf{x}_{(i+1,0)}, \mathbf{x}_{(i,0)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i+1,0)})$ can be obtained by concatenating the u-filtered ensemble at $t_{i+1,0}$ with the stored ensemble at t_i. The filtered distribution $p(\mathbf{x}_{(i+1,0)}, \mathbf{x}_{(i,0)} \mid \mathbf{y}_{0:i}, \mathbf{u}^*_{(0,1):(i+1,0)})$ can be calculated by assimilating $\mathbf{y}_i$ with the concatenated ensemble using Eq. (4). The ensemble of $p(\mathbf{x}_{(i+1,0)} \mid \mathbf{y}_{0:i}, \mathbf{u}^*_{(0,1):(i+1,0)})$ is obtained by marginalizing $\mathbf{x}_{(i,0)}$.

This DA framework can be used to construct an adaptive predictive control algorithm without overlapping the prediction intervals. The control procedure for this experiment is summarized in Fig. 2. In $t_i < t < t_{i+1}$, while the real system evolves in time with the inputs $\mathbf{u}^*_{(i,1)}$, $\mathbf{u}^*_{(i,2)}$, and $\mathbf{u}^*_{(i+1,0)}$, the predictions and control estimates are performed from t_{i+1} to t_{i+2} in the DA system, as shown in Fig. 2a. At $t = t_{i+1}$, the system state is observed as $\mathbf{y}_{i+1}$, which is assimilated into the latest u-filtered distribution (adaptation), as shown in Fig. 2b. The loops of these processes illustrated in Fig. 2 provide an adaptive predictive control.

The proposed control system for LHD comprises a DA system "ASTI" (Assimilation System for Toroidal plasma Integrated simulation)[21], which uses an integrated transport simulation code TASK3D[22,23] as the system model for helical fusion plasmas and implements the EnKF and PF as DA techniques. TASK3D computes the heat and particle transport in a toroidal fusion plasma as a one-dimensional (1D) problem for the normalized minor radius. In this study, ASTI computed 256 ensemble members (TASK3D simulations) on a vector computer (NEC SX-Aurora TSUBASA, 16VE) connected to the LHD experimental system, and EnKF was used to perform the Bayesian filters. The computer had a maximum of 128 parallel processes, each of which was responsible for two ensemble members, thereby computing the 256 ensemble members. The control performance reached its maximum with 200 ensemble members in numerical experiments[18] to control the virtual LHD plasma generated by TASK3D using noise values similar to those in the LHD experiments. Although the behavior of the virtual plasma is different from that of the actual plasma, the system model and state vector are almost the same as in the experiment. Therefore, the effect of the number of ensemble members on the control performance is considered to be sufficiently small in the LHD experiment.

System model for LHD plasmas

We employed the TASK3D code[22,24] as the time evolution model f in Eq. (1). TASK3D is an integrated simulation code for helical fusion plasmas that solves the 1D diffusive transport problem in the radial ρ-direction. The parameter ρ is given by the magnetic flux surface, where 0 and 1 correspond to the center and edge of the plasma,
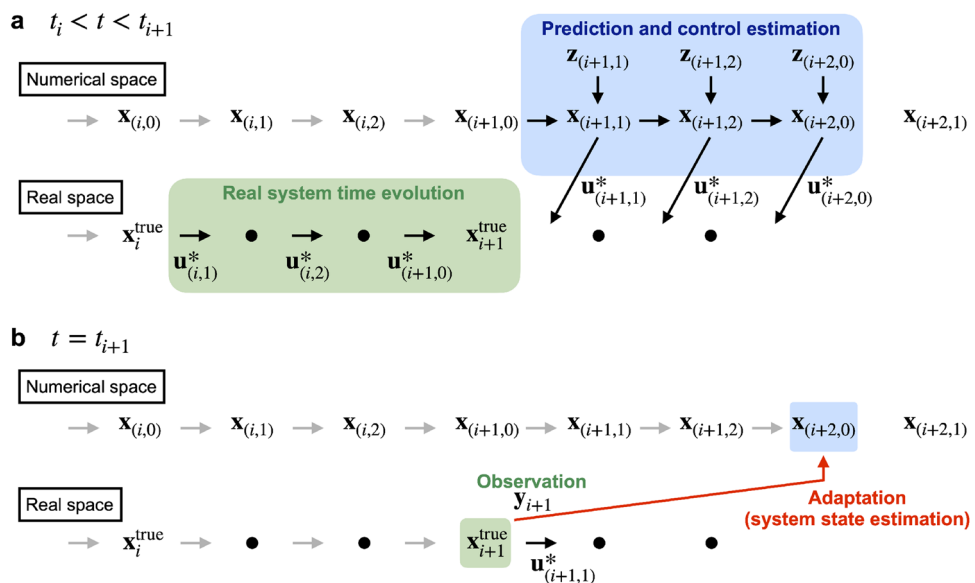

Figure 2. Control procedure of the LHD experiment ($\Delta t_y = 3\Delta t_z$). The control is performed by repeating the control estimation (**a**) and DA of the observation into the latest u-filtered distribution (**b**).

Variable			M_i
$\tilde{\mathbf{x}}$	n	Density	11
	T_e	Electron temperature	11
	T_i	Ion temperature	11
	c_e	Factor for electron thermal diffusivity	11
	c_i	Factor for ion thermal diffusivity	11
$\mathbf{u}$	P_{ECH}	ECH input power	1
$\mathbf{z}$	$T_{e,\rho=0}$	Electron temperature at plasma center	1
$\mathbf{y}$	n	Density	11
	T_e	Electron temperature	11

Table 1. State, target, and observation variables with their dimensions in the vectors (M_i).

respectively. We solved the heat transport equations for each electron and ion species in the LHD experiment. Assume that the electron and ion density profiles are similar, i.e., $n = n_e = n_i$. The radial profile in TASK3D was defined at 60 grid points. The geometric parameters required in the 1D transport simulation were evaluated by an equilibrium magnetic field precomputed using the VMEC code[25] for the typical LHD magnetic configuration. The major radius of the magnetic axis in vacuum was 3.6 m and the magnetic-field strength at the plasma center was 2.75 T. In TASK3D, changes in the geometric parameters were not considered in the time evolution calculation due to the computational cost.

For the electron and ion thermal diffusivities, the following gyro-Bohm models were employed: $\chi_e = C_e^{gB}(T_e/eB)(\rho_i/a)$ and $\chi_i = C_i^{gB}(T_i/eB)(\rho_i/a)$. where B, ρ_i, and a are the magnetic field strength, ion Larmor radius, and plasma minor radius, respectively. We set $C_e^{gB} = 1.5$ and $C_i^{gB} = 0.1$ as reasonable values based on previous studies[24,26]. The heating power source comprised the externally applied ECH, power exchange between species, and loss term by interaction with neutrals. ECH only contributes to the heating term of the electron and the following ECH model was employed in the experiments:

$$p_{ECH}(\rho) = A \exp\left(-\frac{1}{2}\frac{(\mu_{ECH} - \rho)^2}{\sigma_{ECH}^2}\right), \tag{13}$$

where $\mu_{ECH} = 0.15$ and $\sigma_{ECH} = 0.04$ are obtained from the ray-tracing calculations[27]. The coefficient A is determined from the total ECH input power, which is given by

$$P_{ECH} = \int_0^1 p_{ECH}(\rho)\mathscr{V}'(\rho)d\rho, \tag{14}$$

where $\mathscr{V}$ is the plasma volume and $\mathscr{V}' = d\mathscr{V}/d\rho$.

ECH control experiment

To demonstrate real-time adaptation and control estimation using the developed system, we considered a simple LHD experiment to control the central electron temperature by adjusting the ECH power[28,29]. Table 1 lists the state ($\tilde{\mathbf{x}}$ and $\mathbf{u}$), target state ($\mathbf{z}$), and observation variables ($\mathbf{y}$) used in the experiment. The radial profiles of the state variables were defined on 11 grid points ($\rho = 0, 0.1, 0.2, \ldots, 1$) in the state vector, and B-spline interpolation[30] was used to transform the radial profiles to those on the TASK3D grid (60 grid points)[21]. The factors for the thermal diffusivities c_e and c_i were introduced to optimize χ_e and χ_i with large uncertainties. In the prediction step, $c_e\chi_e$ and $c_i\chi_i$ were used instead of χ_e and χ_i in the TASK3D simulation. The ECH used the four gyrotrons: 330, 400, 600, and 600 kW for 5.5 s; thus, ASTI controlled the injection power by selecting a subset of these gyrotrons by switching the individual gyrotrons on or off. ASTI sends the estimated control signals to the gyrotrons every $\Delta t_z = 0.1$ s and receives the observed radial profiles of the electron temperature and density every $\Delta t_y = 0.3$ s from the real-time Thomson scattering measurement system.

To ensure that the behavior of the system model resembles that of the real system at the beginning of control, observation assimilation was performed with a fixed heating power of 731 kW during the first phase $t < 2.1$ s. ASTI assimilates the observations up to 1.8 s, after which the control estimation begins. Subsequently, the ECH power control commences at 2.1 s to produce the target electron temperature (4 keV) at 3.9 s and maintain it. The electron density at the plasma center was maintained at 1.5×10^{19} m^{-3} using the PID control to focus on the temperature control.

The covariance matrices for the noises $Q_{(i,j+1)}$, $R_{(i,j)}^{\mathbf{z}}$, $R_{(i,j)}^{\mathbf{u}}$, and $R_i^{\mathbf{y}}$ are the key parameters affecting the overall control performance. In addition to them, the ensemble mean $\hat{\mathbf{x}}_{(0,0)}$ and covariance matrix $V_{(0,0)}$ are required to generate the initial ensemble members. Diagonal matrices were used for these covariance matrices, i.e., the covariance component of the matrices was not considered.

The covariance matrix for the system noise, $Q_{(i,j+1)}$, controls the uncertainty of the system state, including the model parameters and control inputs. System noise was added before each prediction step, and $\tilde{\mathbf{x}}$ was assigned a slightly larger value of noise after processing by the y-filter to prevent the distribution from shrinking. The standard deviation of the system noise was fixed at the values listed in Table 2. The values of the standard deviation for

Variable		σ_{init}	σ_Q
$\tilde{\mathbf{x}}$	n	20%	2% (10%)
	T_e	15%	2% (3%)
	T_i	15%	2% (3%)
	c_e	0.2	0 (0.1)
	c_i	0.2	0 (0.1)
$\mathbf{u}$	P_{ECH}	0	0.5 (0)

Table 2. Standard deviations of the initial state distribution (σ_{init}) and system noise (σ_Q). The values in parentheses represent the additional system noise added after the y-filter. The values with% as the unit represent the rate for determining the standard deviation in proportion to the state distribution mean.

T_e, T_i, c_e, and c_i were determined based on a previous study on data assimilation for the LHD plasmas[18,21]. The noises for n_e were set to large values because ASTI did not compute the density transport. In control estimation, system noise for $\mathbf{u}$ is an important parameter determining the range of control inputs considered in a single control estimation and the change rate of $\mathbf{u}$. In this experiment, the standard deviation of the noise for P_{ECH} was set to a sufficiently large value to track the rate of change of the target temperature.

The covariance matrix $R^{\mathbf{z}}_{(i,j)}$ affects the performance of the z-filter and determines the proximity of the system state to the target state after the z-filter step. The diagonal components of $R^{\mathbf{z}}_{(i,j)}$ were determined at each z-filtering step as follows:

$$\left(R^{\mathbf{z}}_{(i,j)} \right)_{ll} = r_z^2 \left(H^{\mathbf{z}} V_{(i,j)} (H^{\mathbf{z}})^{\mathrm{T}} \right)_{ll}, \tag{15}$$

where $V_{(i,j)}$ denotes the covariance matrix of the ensemble that approximates the predicted distribution at $t_{i,j}$ and $r_z = 0.5$[18]. The subscript $(\ )_{ll}$ and superscript $^{\mathrm{T}}$ denote the l-th diagonal component and matrix transposition, respectively.

The covariance matrix $R^{\mathbf{u}}_{(i,j)}$ considers the uncertainty in the control input. In the experiment, the standard deviations of the control input noise were set to a sufficiently small value of 0.05 MW for P_{ECH}.

The covariance matrix $R^{\mathbf{y}}_i$ affects the performance of the y-filter and determines the effect of the observed information on the state distribution. The standard deviation of the observation noise is assumed to be proportional to the difference between the observation and mean of the state distribution[18],

$$\left(R^{\mathbf{y}}_i \right)_{ll} = r_y^2 \left(\mathbf{y}_i - H^{\mathbf{y}} \hat{\mathbf{x}}_{(i,0)} \right)_l^2, \tag{16}$$

where $r_y = 0.8$, $\hat{\mathbf{x}}_{(i,0)}$ is the mean of the ensemble approximating $p(\mathbf{x}_{(i,0)} \mid \mathbf{y}_{0:i-1}, \mathbf{u}^*_{(0,1):(i,0)})$, and $(\ )_l$ represents the l-th element of the vector. This assumption prevents the control instability caused by overfitting of the system model to noisy observations and maintains the variance of the y-filtered ensemble at an adequate magnitude.

The initial ensemble means of T_e and T_i are set to the steady-state radial profiles calculated by the TASK3D simulation for the initial ECH. The initial mean profile of n is set to $n(\rho) = 1.0 - 0.8\rho^8$ [$\times 10^{19} \text{m}^{-3}$] and those of c_e and c_i are set to 1. The initial ensemble variances are assigned the values listed in Table 2.

Real-time observation and control system

The DA system ASTI was connected to the real-time measurement system and ECH system[28,29] via the real-time communication system in LHD[31]. The temperature and density profiles along with their measurement errors obtained from the Thomson scattering measurement system[32,33] were sent from the Thomson data analysis PC to the vector engine server (SX-Aurora TSUBASA, NEC Inc.) via socket communication. The digitizers of the high-repetition-rate Thomson scattering system[34] were used for real-time measurements (10 Hz). In the vector engine server, ASTI considers the time delay relative to real time, which is tens of milliseconds. This is attributed to the temperature computation, density analysis, and communication delay. The electron density and temperature were observed at 144 radial points, of which 66 data points were available for real-time measurement. After removing the obvious outliers based on measurement errors, a mapping model was used to transform the observed data in the real space coordinate (major radius R) into a form suitable for DA (profiles on 11 points of the ρ-coordinate).

The mapping model acted as coordinate transformation, outlier removal, data smoothing, and data point extraction for DA, providing the observation vector $\mathbf{y}$ to ASTI. The mapping model was based on a multilayer perceptron with five hidden layers containing 500 units each using ReLu as the activation function. The mapping model was trained using data on the LHD experimental database (profiles observed by the Thomson scattering system and the mapped data obtained from the equilibrium calculations). The training data comprised 58 discharges (7769 time points) of the ECH plasma, of which a quarter were used as the test data. For all discharges in the training data, equilibrium magnetic fields were calculated and used to evaluate the density and temperature profiles in the ρ-coordinate for the teacher data. Before training, the mapped data were fitted with even-order polynomials up to the 8th order to smooth the profiles and extract the data points required for DA. An even function was used for fitting because the Neumann boundary condition $\partial/\partial\rho = 0$ was imposed on $\rho = 0$ in TASK3D. The variations of the magnetic surfaces were considered in this mapping model, and the observations in the ρ-coordinate were provided to ASTI. Since TASK3D did not solve for the magnetic equilibrium, errors occurred in the geometric parameters required for the 1D transport simulation (e.g., $d\rho/dR$). In this experiment,

these errors were considered small for helical devices and allowed. However, the real-time computation of the magnetic field, which is essential for tokamak control, will be addressed in future studies.

After the prediction and control estimation in ASTI were completed, the generated ECH control signals were sent to the computer (Jetson AGX Orin, NVIDIA Inc.) via socket communication. The digital signal for ECH control, which was converted to an analog signal using a D/A converter (EVAL-AD5686RSDZ, Analog Device Inc.) connected to a computer, was input to the ECH system to control the ON/OFF state of the ECH injection.

Results

Figure 3 shows the control results of the LHD experiment and Fig. 4 shows the prediction error (the absolute difference between "Prediction" and "Observation" in Fig. 3). The electron temperature increases to that of the target (4 keV) and is maintained beyond 3.9 s. The prediction error, shown in Fig. 4, increases to 2 keV in the transient section ($2.1\,\text{s} \leq t < 3.9\,\text{s}$) and decreases to approximately 0.2 keV in the steady-state section ($3.9\,\text{s} \leq t$). These results demonstrate the effectiveness of real-time adaptation and control estimation using the DA-based
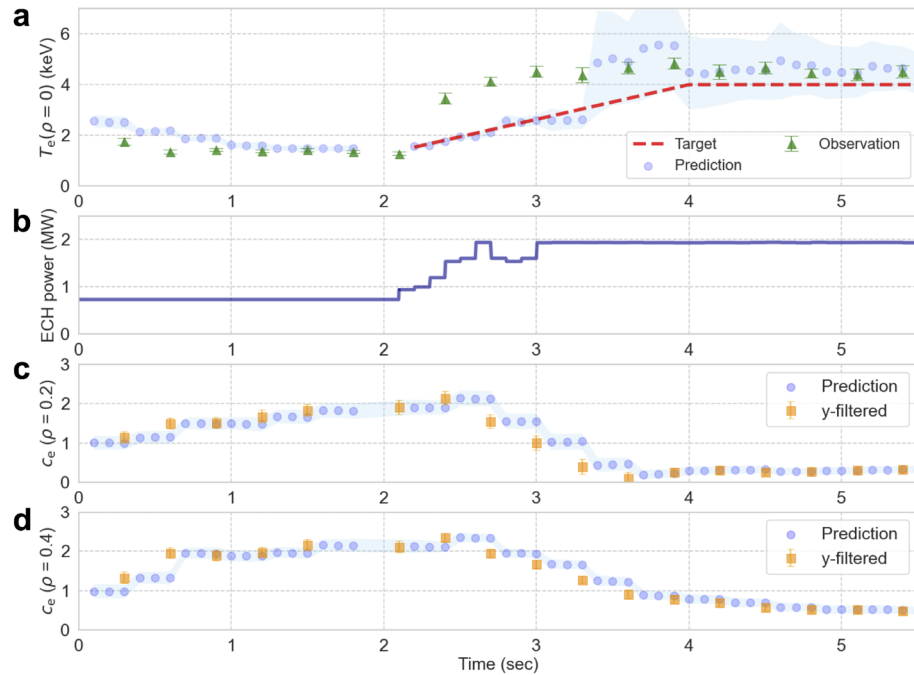
Figure 3. Results of a control experiment (#186500). (**a**) Control result of T_e at the plasma center. The plotted values labeled "Prediction" correspond to the expected values of the predicted distribution for $t \leq 2.1\,\text{s}$ and those of the u-filtered distributions for $t > 2.1\,\text{s}$. The shaded areas represent a single standard deviation of the distributions. The plotted values labeled "Observation" are those obtained by the mapping model from the Thomson scattering measurements. (**b**) ECH power adjusted using ASTI. (**c**) and (**d**) Distribution (expected value and one standard deviation) of (c_e) at $\rho = 0.2$ and 0.4 used in the prediction step ("Prediction") and y-filtered distribution ("y-filtered").
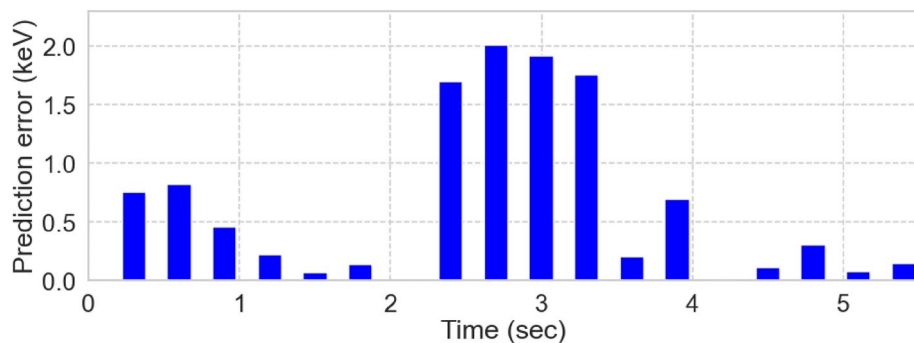

Figure 4. Prediction error defined as the absolute difference between "Prediction" and "Observation" in Fig. 3.

control system. The prediction error in the transient section occurred mainly because the gyro-Bohm model overestimated the thermal diffusivities as the electron temperature increased. At the beginning of the control (23 s), the c_e optimization lagged behind the variation in the actual thermal diffusivities. However, before 4 s, the gap between the model and the actual system was bridged, which improved the prediction performance. The prediction error also depends on the Thomson measurement error of 0.2–0.4 keV and accuracy of the mapping model.

In addition, the discrete nature of the control input and expansion of the u-filtered distribution from 3.4 s onward also affect the control accuracy. This expansion is attributed to the increase in ensemble members with small $c_e (< 0.3)$ near the center as shown in Fig. 3c, which increases the range of electron temperature variation. Because the state distribution is approximated by finite ensemble members, a large uncertainty in the predicted distribution makes it difficult to identify the relationship between the control inputs and the system state and may reduce the control accuracy. The expansion of the u-filtered distribution at 3.3 s was triggered by adding system noise with the fixed standard deviation of 0.1 to c_e, despite the small mean of c_e near the center. This expansion can be avoided by introducing a method to dynamically adjust the system noise by considering the characteristics of the system model. The probability distribution of model parameters should be adjusted by considering the physical characteristics and control accuracy. This issue will be addressed in a subsequent control experiment.

Change of state distribution in the control process

The ensemble members of the state distributions at 3.0 s are shown in Fig. 5. Figure 5a shows the z-filter process in which the target temperature is assimilated into the predicted ensemble. The z-filtered ensemble approximates the distribution $p(\mathbf{x}_{3.0} \mid \mathbf{y}_{0.3:2.4}, \mathbf{u}^*_{2.1:2.9}, \mathbf{z}_{3.0})$, where the subscript number indicates the time (s) and $t_1 : t_2$ denotes the values in $[t_1, t_2]$. As shown in Fig. 5b, the u-filtered ensemble is computed after the control input $\mathbf{u}^*_{3.0}$ is obtained as the expected value of this distribution. The distribution approximated by the u-filtered ensemble corresponds to the predicted distribution with zero or small uncertainty in the control input. The remaining uncertainties originate from the uncertainties in the model parameters and system state before model evolution calculation. The u-filtered ensemble is modified by assimilating the observation $\mathbf{y}_{2.7}$ (y-filter), as shown in Fig. 5c. This filter optimizes the state variables, including the model parameters, to adapt the system model to the real system. The control process proceeds to the subsequent prediction from the y-filtered ensemble.

The filtering steps were performed using EnKF in the simple control experiment. Strongly nonlinear relationships among the state variables and a significantly non-Gaussian state distribution (e.g., a multimodal distribution) necessitate the use of other filters (e.g., PF) and another definition of $\mathbf{u}^*$ (e.g., mode).

Real-time adaptation to the real system

The upper panels of Fig. 6 show the radial profiles of the predicted electron temperature (u-filtered distribution) and observed profiles at three different points in time. The radial profiles of the electron thermal diffusivity are shown in the lower panels. The thermal diffusivity can be estimated from the observed density and temperature profiles as

$$\chi(\rho) \sim \frac{-\frac{1}{\mathscr{V}'} \int_0^\rho p_{\text{ECH}} \mathscr{V}' \mathrm{d}\rho}{\langle |\nabla \rho|^2 \rangle n \frac{\partial T}{\partial \rho}}, \tag{17}$$

where $\langle \ \rangle$ represents the magnetic flux surface average. This estimate is less valid near the center where the temperature gradient is smaller.

After the first phase (observation assimilation), the temperature profile was predicted with high accuracy at the beginning of control estimation (t = 1.8 s), as shown in Fig. 6a,d. In the transient phase (2.1 s $\leq t <$ 3.9 s), insufficient modeling of the electron thermal diffusivity (Fig. 6e) underestimated the profile around the plasma center, as shown in Fig. 6b. These prediction errors contribute to the control error in the transient phase, as
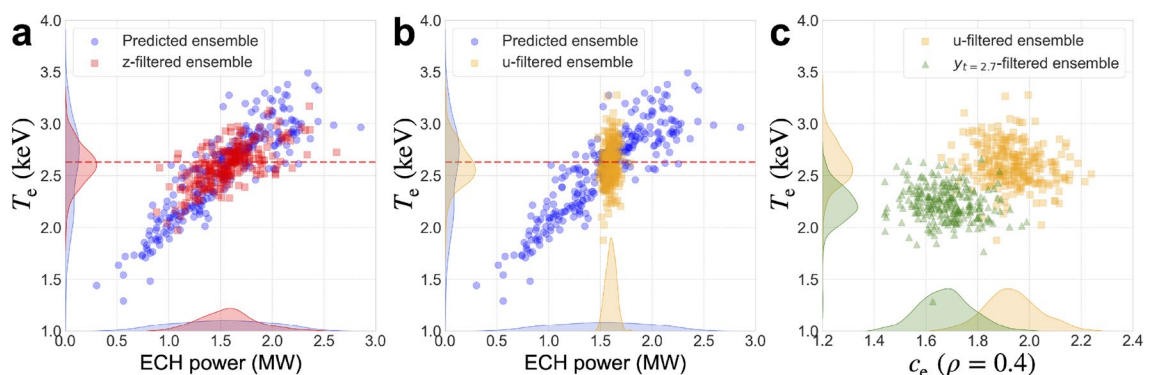

Figure 5. Variations in the ensembles that approximate the state distributions at $t = 3.0$ s in the control experiment. (**a**) Estimation of the control input (z-filter). (**b**) Predicted distribution modified by the estimated control input (u-filter). (**c**) Adaptation to the actual LHD plasma by assimilating the observation at $t = 2.7$ s (y-filter).
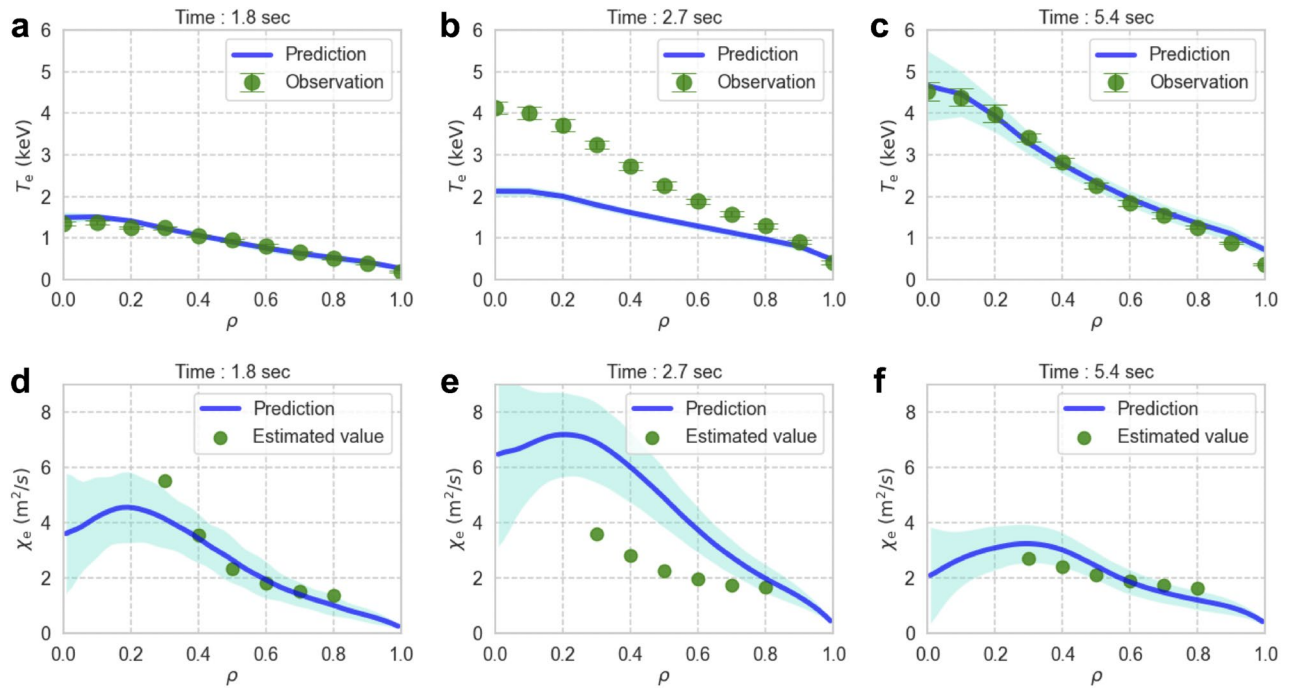
Figure 6. Time evolution of the radial profiles of electron temperature and thermal diffusivity. (**a**–**c**) Radial profiles of the predicted electron temperature corresponding to Fig. 3a and the observations (mapped to the ρ-coordinate) at three points in time. (**d**–**f**) Radial profiles of electron thermal diffusivity calculated in the prediction step ("Prediction") and the values estimated using Eq. (17) from the observed temperature profile at the corresponding time points ("Estimated value"). The shaded areas around the radial profiles represent a single standard deviation.

shown in Fig. 3. However, in the steady-state phase (3.9 s $\leq$ t), the temperature profile was predicted with high accuracy by optimizing the thermal diffusivity model, as shown in Fig. 6c,f.

These results demonstrate the real-time adaptation of the integrated transport simulation code to the actual LHD plasma. In the transient phase, changes in the circumstances and the discrepancy between the system model and real system degrade the prediction performance. Furthermore, in the control procedure used in this experiment, the observations obtained at time t_i were reflected in the control estimation after $t_i + \Delta t_y$, which caused a delay in adaptation, especially at the beginning of the control (2.1 s−). This delay can be a critical issue for some control problems. Therefore, to achieve a stable adaptive predictive control system, it is important to design the control procedure, Δt_y, and the observation noise to minimize this delay.

Reduction of the control error in the transient phase makes it necessary to improve the assumed transport model, adjust the adaptive capacity by controlling the uncertainty in model parameters, and minimize the adaptation delay. These limitations will be addressed in the future studies aimed at building a stable control system. The DA framework can be applied to construct an advanced control system that can optimize the timing, physical quantities, and measurement positions. For instance, it would be possible to develop a system that allows measurement devices to observe the system state depending on the uncertainty of the model parameters and system state.

Discussion

A control system based on DA was developed to control LHD plasmas. A simple experiment to control the electron temperature demonstrated adaptive predictive control using a nonlinear system model (integrated transport model) with large uncertainties. The proposed control framework simplifies the structure of the control system to accommodate complex control problems characterized by a large number of variables.

The adaptive capacity of a control system is limited by the assumed models and frequency of observations[18]. Therefore, the sophistication of the model, computational cost, and amount of information contained in the observations must be adjusted depending on the target control problem. Similar to other control systems, observation errors significantly affect the performance of the proposed DA-based control system. However, this system can explicitly account for the uncertainties in observations based on the observation noise. Robust and stable control can be achieved by exploiting this uncertainty; however, methods to adjust the uncertainties in observations and model parameters according to the situation should be developed.

Advanced control systems constructed for future fusion reactors would require models based on data-driven approaches, such as surrogate models[35–37], stability indicators[15,38,39], and tomography methods[40,41]. To this end, the DA-based control system, which expresses the system state as a probability distribution, is highly compatible

with other data-driven approaches and provides a flexible control platform that links the physics-based and data-driven models.

The results of the first application of DA-based control system in LHD do not exceed the control performance of conventional controllers (e.g., PID controller). However, the proposed approach has high potential in the control of future fusion reactors under large uncertainties because it can estimate and control the internal state from limited observed information and naturally handle various observed and control variables (e.g., temperature and density profiles) and plasma responses with different time scales in a single framework. Thus, the success of this demonstration paves the way for addressing challenging control problems in fusion plasma, such as multivariate control, radial profile control, electric field bifurcation control, and avoidance of terminating events using relevant alarm rates. DA-based control may become a key technology for solving challenging control problems in future fusion reactors and other complex systems by augmenting numerical models with real systems. Beginning in 2024, our roadmap includes conducting more sophisticated control experiments with this control system by incorporating additional measurement, heating, and fuel-supply devices into the system. Additionally, we are actively working on extending our control capabilities to include Tokamak control.

Data availability

The raw data were generated at the LHD facility. Automatic Integrated Data Analysis software and the analyzed data are available from the LHD data repository (https://doi.org/10.57451/lhd.analyzed-data). The calculated data depicted in the figures of this paper are available from the corresponding author upon reasonable request.

References

1. Fasoli, A. *et al.* Computational challenges in magnetic-confinement fusion physics. *Nat. Phys.* **12**(5), 411–423. https://doi.org/10.1038/nphys3744 (2016).
2. Winter, J. Wall conditioning in fusion devices and its influence on plasma performance. *Plasma Phys. Control. Fusion* **38**(9), 1503 (1996).
3. Mayne, D. Q. Model predictive control: Recent developments and future promise. *Automatica* **50**(12), 2967–2986. https://doi.org/10.1016/j.automatica.2014.10.128 (2014).
4. Gettelman, A. *et al.* The future of earth system prediction: Advances in model-data fusion. *Sci. Adv.* **8**(14), 3488. https://doi.org/10.1126/sciadv.abn3488 (2022).
5. Carrassi, A., Bocquet, M., Bertino, L. & Evensen, G. Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdiscip. Rev. Climate Change* **9**(5), 535. https://doi.org/10.1002/wcc.535 (2018).
6. King, J., Anchukaitis, K. J., Allen, K., Vance, T. & Hessl, A. Trends and variability in the southern annular mode over the common era. *Nat. Commun.* **14**(1), 2324. https://doi.org/10.1038/s41467-023-37643-1 (2023).
7. Tierney, J. E. *et al.* Spatial patterns of climate change across the paleocene-eocene thermal maximum. *Proc. Natl. Acad. Sci.* **119**(42), 2205326119. https://doi.org/10.1073/pnas.2205326119 (2022).
8. Hou, X. *et al.* Intracounty modeling of Covid-19 infection with human mobility: Assessing spatial heterogeneity with business traffic, age, and race. *Proc. Natl. Acad. Sci.* **118**(24), 2020524118. https://doi.org/10.1073/pnas.2020524118 (2021).
9. Kato, H., Yoshizawa, A., Ueno, G. & Obayashi, S. A data assimilation methodology for reconstructing turbulent flows around aircraft. *J. Comput. Phys.* **283**, 559–581. https://doi.org/10.1016/j.jcp.2014.12.013 (2015).
10. Tian, S., Van Dijk, A. I., Tregoning, P. & Renzullo, L. J. Forecasting dryland vegetation condition months in advance through satellite data assimilation. *Nat. Commun.* **10**(1), 469. https://doi.org/10.1038/s41467-019-08403-x (2019).
11. Miyoshi, T. *et al.* "Big data assimilation" revolutionizing severe weather prediction. *Bull. Am. Meteor. Soc.* **97**(8), 1347–1354. https://doi.org/10.1175/BAMS-D-15-00144.1 (2016).
12. Tian, S., Van Dijk, A. I., Tregoning, P. & Renzullo, L. J. Forecasting dryland vegetation condition months in advance through satellite data assimilation. *Nat. Commun.* **10**(1), 469. https://doi.org/10.1038/s41467-019-08403-x (2019).
13. Degrave, J. *et al.* Magnetic control of tokamak plasmas through deep reinforcement learning. *Nature* **602**(7897), 414–419. https://doi.org/10.1038/s41586-021-04301-9 (2022).
14. Wakatsuki, T., Suzuki, T., Oyama, N. & Hayashi, N. Ion temperature gradient control using reinforcement learning technique. *Nucl. Fusion* **61**(4), 046036. https://doi.org/10.1088/1741-4326/abe68d (2021).
15. Kates-Harbeck, J., Svyatkovskiy, A. & Tang, W. Predicting disruptive instabilities in controlled fusion plasmas through deep learning. *Nature* **568**(7753), 526–531. https://doi.org/10.1038/s41586-019-1116-4 (2019).
16. Takeiri, Y. *et al.* Extension of the operational regime of the LHD towards a deuterium experiment. *Nucl. Fusion* **57**(10), 102023. https://doi.org/10.1088/1741-4326/aa7fc2 (2017).
17. Osakabe, M. *et al.* Recent results from deuterium experiments on the large helical device and their contribution to fusion reactor development. *Nucl. Fusion* **62**(4), 042019. https://doi.org/10.1088/1741-4326/ac3cda (2022).
18. Morishita, Y., Murakami, S., Yokoyama, M. & Ueno, G. Data assimilation and control system for adaptive model predictive control. *J. Comput. Sci.* **72**, 102079. https://doi.org/10.1016/j.jocs.2023.102079 (2023).
19. Evensen, G. The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.* **53**(4), 343–367. https://doi.org/10.1007/s10236-003-0036-9 (2003).
20. Kitagawa, G. Monte Carlo filter and smoother for non-gaussian nonlinear state space models. *J. Comput. Graph. Stat.* **5**(1), 1–25. https://doi.org/10.1080/10618600.1996.10474692 (1996).
21. Morishita, Y., Murakami, S., Yokoyama, M. & Ueno, G. ASTI: Data assimilation system for particle and heat transport in toroidal plasmas. *Comput. Phys. Commun.* **274**, 108287. https://doi.org/10.1016/j.cpc.2022.108287 (2022).
22. Murakami, S. *et al.* Integrated transport simulations of high ion temperature plasmas of LHD. *Plasma Phys. Control. Fusion* **57**(5), 054009. https://doi.org/10.1088/0741-3335/57/5/054009 (2015).
23. Yokoyama, M. *et al.* Extended capability of the integrated transport analysis suite, task3d-a, for LHD experiment. *Nucl. Fusion* **57**(12), 126016. https://doi.org/10.1088/1741-4326/aa800a (2017).
24. Sakai, A. *et al.* Integrated particle transport simulation of NBI plasmas in LHD. *Plasma Fusion Res.* **10**, 3403048–3403048. https://doi.org/10.1585/pfr.10.3403048 (2015).
25. Hirshman, S. *et al.* Three-dimensional free boundary calculations using a spectral green's function method. *Comput. Phys. Commun.* **43**(1), 143–155. https://doi.org/10.1016/0010-4655(86)90058-5 (1986).

26. Wakasa, A., Fukuyama, A., Murakami, S., Miki, M., Yokoyama, M., Sato, M., Toda, S., Funaba, H., Tanaka, K., Ida, K., Yamada, H., Honda, M., & Nakajima, N. Integrated transport simulation of LHD plasmas using TASK3D, in *Proc. 23rd IAEA Fusion Energy Conf. (Daejon)*, P4.029 (2010).

27. Tsujimura, T. I. *et al.* Development and application of a ray-tracing code integrating with 3d equilibrium mapping in LHD ECH experiments. *Nucl. Fusion* **55**(12), 123019. https://doi.org/10.1088/0029-5515/55/12/123019 (2015).

28. Takahashi, H. *et al.* Extension of high t e regime with upgraded electron cyclotron resonance heating system in the large helical device. *Phys. Plasmas* **21**(6), 061506. https://doi.org/10.1063/1.4884365 (2014).

29. Shimozuma, T. *et al.* Ecrh-related technologies for high-power and steady-state operation in LHD. *Fusion Sci. Technol.* **58**(1), 530–538. https://doi.org/10.13182/FST58-530 (2010).

30. De Boor, C. *A Practical Guide to Splines* (Springer-Verlag, New York, 1978).

31. Nakanishi, H. *et al.* Real-time data streaming and storing structure for the LHD's fusion plasma experiments. *IEEE Trans. Nucl. Sci.* **63**(1), 222–227. https://doi.org/10.1109/TNS.2016.2515099 (2016).

32. Narihara, K., Yamada, I., Hayashi, H. & Yamauchi, K. Design and performance of the Thomson scattering diagnostic on LHD. *Rev. Sci. Instrum.* **72**(1), 1122–1125. https://doi.org/10.1063/1.1319368 (2001).

33. Yamada, I., Narihara, K., Funaba, H., Minami, T., Hayashi, H., Kohmoto, T., Group, L.E. Recent progress of the LHD Thomson scattering system. *Fusion Sci. Technol.* **58**(1), 345–351 (2010).

34. Funaba, H. *et al.* Electron temperature and density measurement by Thomson scattering with a high repetition rate laser of 20 khz on lhd. *Sci. Rep.* **12**(1), 15112. https://doi.org/10.1038/s41598-022-19328-9 (2022).

35. Narita, E., Honda, M., Nakata, M., Yoshida, M. & Hayashi, N. Quasilinear turbulent particle and heat transport modelling with a neural-network-based approach founded on gyrokinetic calculations and experimental data. *Nucl. Fusion* **61**(11), 116041. https://doi.org/10.1088/1741-4326/ac25be (2021).

36. Rodriguez-Fernandez, P., Howard, N. & Candy, J. Nonlinear gyrokinetic predictions of SPARC burning plasma profiles enabled by surrogate modeling. *Nucl. Fusion* **62**(7), 076036. https://doi.org/10.1088/1741-4326/ac64b2 (2022).

37. Dong, G. *et al.* Deep learning based surrogate models for first-principles global simulations of fusion plasmas. *Nucl. Fusion* **61**(12), 126061. https://doi.org/10.1088/1741-4326/ac32f1 (2021).

38. Murari, A. *et al.* Stacking of predictors for the automatic classification of disruption types to optimize the control logic. *Nucl. Fusion* **61**(3), 036027 (2021).

39. Yokoyama, T. *et al.* Prediction of radiative collapse in large helical device using feature extraction by exhaustive search. *J. Fusion Energy* **39**(6), 500–511 (2020).

40. Kenmochi, N., Nishiura, M., Nakamura, K. & Yoshida, Z. Tomographic reconstruction of imaging diagnostics with a generative adversarial network. *Plasma Fusion Res.* **14**, 1202117–1202117. https://doi.org/10.1585/pfr.14.1202117 (2019).

41. Carvalho, D. *et al.* Deep neural networks for plasma tomography with applications to jet and compass. *J. Instrum.* **14**(09), 09011. https://doi.org/10.1088/1748-0221/14/09/C09011 (2019).

Acknowledgements

Authors' contributions

Y.M. developed and ran the DA system and analyzed the data. S.M., N.K., and M.O. led the project and designed the control experiments. H.F. and I.Y. developed the measurement system. Y.Morishita, M.Y., and G.U. designed the DA system and control procedure. Y.M., N.K., H.F., Y.M., K.N., H.N., and R.S. integrated the DA, measurement, and heating systems. All the authors contributed to the preparation of this manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.