Application of Data Assimilation to Transport Simulations of Tokamak Plasmas

Ryu E. ICHIKAWA[1)]*, Yuya MORISHITA[1)], Emi NARITA[1)], Mitsuru HONDA[2)], Hajime URANO[3)], Sadayoshi MURAKAMI[1)]

[1)] *Department of Nuclear Engineering, Kyoto University, Nishikyo, Kyoto 615-8540, Japan*
[2)] *Graduate School of Engineering, Kyoto University, Nishikyo, Kyoto 615-8530, Japan*
[3)] *Naka Institute for Fusion Science and Technology, National Institutes for Quantum Science and Technology, Naka, Ibaraki 311-0193, Japan*

(Received 15 January 2025 / Accepted 16 May 2025)

Data assimilation, a technique that uses actual measurements to optimize simulation models, is a powerful approach for achieving fast and accurate predictions of fusion plasma behavior. In this study, we validate the effectiveness of the data assimilation technique in the integrated simulation of tokamak plasmas. We use the data assimilation system ASTI, which has been successfully applied to real-time prediction and control of helical plasmas. We extend ASTI for transport simulation of tokamak plasmas and introduce a new data assimilation method that incorporates measurement error information. In this paper, we present simulation results using measurements from JT-60U plasma heated by neutral beam injection. Comparisons of several turbulent transport models are also provided. The results demonstrate that the data assimilation method is effective in tokamak simulation as well and expected to be useful for real-time prediction and control in the future.

Keywords: integrated simulation, data assimilation, the ensemble Kalman filter, JT-60U, NBI heating

DOI: 10.1585/pfr.20.1403036

1. Introduction

Magnetic fusion plasmas are complex systems in which a variety of physical phenomena interact with one another. To predict and control the behavior of fusion plasmas, the nonlinear dependencies among physical quantities must be taken into account.

Integrated simulations have been developed to predict and analyze plasma behavior and to design operational scenarios for burning plasmas. Real-time and control-oriented predictions of tokamak plasmas have also been achieved using integrated simulation codes [1–3]. However, modules in integrated simulation codes have uncertainties that arise from various approximations and assumptions of the physical model. These uncertainties interact and degrade the accuracy of simulation results. Moreover, many physical phenomena are not accounted for in the codes. As a result, predictions from integrated simulations often diverge from the actual plasma behavior.

To address these issues, we have introduced a data assimilation (DA) technique to integrated simulation. DA is a method to improve the simulation accuracy by optimizing the simulation model using actual observation data. This technique has been studied mainly in the fields of meteorology and oceanography for accurate forecasts. We have developed a DA system ASTI (Assimilation System for Toroidal plasma Integrated simulation) for integrated simulation of fusion plasmas. In the previous works, the prediction of LHD plasma with improved accuracy using ASTI has been achieved [4, 5]. ASTI has also been expanded as an adaptive model predictive control system, and the control experiment conducted at LHD has proven the effectiveness of the DA-based control of helical plasmas [6, 7].

While ASTI was originally developed for helical plasmas, we are currently expanding it for transport control of tokamak plasmas. As the first step in this expansion, this study focuses on predictive simulations of JT-60U plasmas. We use measurements from neutral beam injection (NBI) discharges and perform simulations with the DA technique. We also improve the DA method implemented in ASTI to account for the measurement error information. Similar techniques using the extended Kalman filter have been applied for reconstructions of density profile [8, 9] and q-profile [10]. Although the extended Kalman filter is a powerful approach, it may perform poorly for strongly nonlinear systems due to its reliance on linear approximations. In contrast, the ensemble Kalman filter (EnKF) [11] which we employ in this study, avoids such linearization and is therefore more suitable for nonlinear systems such as burning plasmas. Additionally, we establish a quantitative evaluation method for turbulent transport models using DA and compare the performance of several models.

*Corresponding author's e-mail: ichikawa.ryu.84n@st.kyoto-u.ac.jp

In Sec. 2, the DA method, the simulation model and the simulation configuration in this study are presented. The results given by the simulations with ASTI for JT-60U plasma are shown in Sec. 3. The conclusions are given in Sec. 4.

2. Data Assimilation

Simulation models inherently differ from actual physical systems, as they are simplified representations based on assumptions and approximations. DA is a method for probabilistically quantifying the uncertainty in simulations and estimating the conditional probability distribution of a state based on actual observations, using Bayes' theorem. ASTI employs the EnKF as the DA scheme.

2.1 The ensemble Kalman filter

The EnKF is an algorithm based on sequential Bayesian filters. It treats the system state probabilistically, rather than deterministically as in typical numerical simulations. In the EnKF, the system state is represented by a state-space model, which consists of a nonlinear, non-Gaussian system model and a linear, Gaussian observation model:

$$\mathbf{x}_t = f_t(\mathbf{x}_{t-1}, \mathbf{v}_t), \qquad \mathbf{v}_t \sim p(\mathbf{v}_t), \tag{1}$$

$$\mathbf{y}_t = H_t\mathbf{x}_t + \mathbf{w}_t, \qquad \mathbf{w}_t \sim N(\mathbf{0}, R_t). \tag{2}$$

Here, $\mathbf{x}_t$ and $\mathbf{y}_t$ represent the state vector and the observation vector at time t, respectively. In Eq. (1), the state vector $\mathbf{x}_{t-1}$ is time-evolved by the system model f_t, with system noise $\mathbf{v}_t$ added to represent deviations from the true state. In this study, we assume Gaussian system noise: $p(\mathbf{v}_t) = N(\mathbf{0}, Q_t)$. In Eq. (2), the observed variables are extracted from the state vector $\mathbf{x}_t$ using the observation matrix H_t, and Gaussian observation noise $\mathbf{w}_t$ is added to express the gap between the estimated value and the observed one.

From the system model, a non-Gaussian distribution $p(\mathbf{x}_t)$ and a Gaussian distribution $p(\mathbf{y}_t)$ are obtained. The objective of the EnKF is to estimate the conditional probability distribution $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ of the state $\mathbf{x}_t$ given the observations $\mathbf{y}_{1:t} \equiv \{\mathbf{y}_1, ..., \mathbf{y}_t\}$. This distribution is refered to as "the filtered distribution", and it is estimated using "the predicted distribution" $p(\mathbf{x}_t|\mathbf{y}_{1:t-1})$ and the latest observation $\mathbf{y}_t$.

The EnKF represents the probability distribution $p(\mathbf{x})$ using an ensemble of N samples, $\{\mathbf{x}^{(n)}\}_{n=1}^{N}$ via the Monte Carlo approximation:

$$p(\mathbf{x}) \doteq \frac{1}{N}\sum_{n=1}^{N}\delta\left(\mathbf{x} - \mathbf{x}^{(n)}\right), \tag{3}$$

where $\doteq$ denotes equality with Monte Carlo approximation.

In the EnKF, the two steps of the prediction and the filtering are repeated alternately using the state-space model described in Eqs. (1) and (2). In the prediction steps, each filtered ensemble member is time-evolved using the system model:

$$\mathbf{x}_{t|t-1}^{(n)} = f_t\left(\mathbf{x}_{t-1|t-1}^{(n)}, \mathbf{v}_t^{(n)}\right). \tag{4}$$

$\mathbf{x}_{t|t'}^{(n)}(1 \le n \le N, t'$ is an arbitrary time) denotes the nth member of the ensemble which approximates the probability distribution $p(\mathbf{x}_t|\mathbf{y}_{1:t'})$. System noise $\mathbf{v}_t^{(n)}$ is generated from Gaussian distribution $N(\mathbf{0}, Q_t)$. After prediction, ensemble of the filtered distribution is calculated from the predicted ensemble and the observation model, Eq. (2):

$$\mathbf{x}_{t|t}^{(n)} = \mathbf{x}_{t|t-1}^{(n)} + \widehat{K}_t\left(\mathbf{y}_t + \widetilde{\mathbf{w}}_t^{(n)} - H_t\mathbf{x}_{t|t-1}^{(n)}\right), \tag{5}$$

where $\widetilde{\mathbf{w}}_t^{(n)} = \mathbf{w}_t^{(n)} - (1/N)\sum_{k=1}^{N}\mathbf{w}_t^{(k)}$, and the observation noise $\mathbf{w}_t^{(n)}$ is generated from $N(\mathbf{0}, R_t)$. $\widehat{K}_t$ is called Kalman gain, and determines the degree of the correction to the state vector. $\widehat{K}_t$ is given as

$$\widehat{K}_t = \widehat{V}_{t|t-1}H_t^{\mathrm{T}}\left(H_t\widehat{V}_{t|t-1}H_t^{\mathrm{T}} + R_t\right)^{-1}, \tag{6}$$

where $\widehat{V}_{t|t-1}$ is a sample covariance matrix of the ensemble for the predicted distribution, $\left\{\mathbf{x}_{t|t-1}^{(n)}\right\}_{n=1}^{N}$. The key symbols used in this paper are summarized in Table 1.

2.2 System model

We employ the integrated simulation code TASK [12] as the system model for ASTI. The core of TASK code is the one dimensional transport code TASK/TR, which solves the 1D diffusive transport equations for particle and energy:

$$\begin{aligned}\frac{\partial}{\partial t}(n_s\mathcal{V}') = &-\frac{\partial}{\partial\rho}\Big(\mathcal{V}'\langle|\nabla\rho|\rangle n_s V_s \\ &-\mathcal{V}'\langle|\nabla\rho|^2\rangle D\frac{\partial n_s}{\partial\rho}\Big) + S_s\mathcal{V}',\end{aligned} \tag{7}$$

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{3}{2}n_sT_s\mathcal{V}'^{5/3}\right) = &-\mathcal{V}'^{2/3}\frac{\partial}{\partial\rho}\Big\{\mathcal{V}'\langle|\nabla\rho|\rangle \\ &\left(V_{K_s} + \frac{3}{2}V_s\right)n_sT_s - \mathcal{V}'\langle|\nabla\rho|^2\rangle \\ &\left(\frac{3}{2}DT_s\frac{\partial n_s}{\partial\rho} + n_s\chi_s\frac{\partial T_s}{\partial\rho}\right)\Big\} \\ &+P_s\mathcal{V}'^{5/3}.\end{aligned} \tag{8}$$

Table 1. Key symbols in this paper.

$\mathbf{x}_t$	State vector at time t
$\mathbf{y}_t$	Observation vector at time t
f_t	System model at time t
H_t	Observation matrix at time t
$\mathbf{v}_t$	System noise at time t
$\mathbf{w}_t$	Observation noise at time t
Q_t	System-noise covariance matrix at time t
R_t	Observation-noise covariance matrix at time t
$\mathbf{y}_{1:t}$	Set of observations, $\{\mathbf{y}_1, ..., \mathbf{y}_t\}$
$p(\mathbf{x}_t\|\mathbf{y}_{1:t-1})$	Probability distribution of the state vector at time t, given the observations at time 1 to $t-1$
$p(\mathbf{x}_t\|\mathbf{y}_{1:t})$	Probability distribution of the state vector at time t, given the observations at time 1 to t
$\left\{\mathbf{x}_{t\|t-1}^{(i)}\right\}_{i=1}^{N}$	Ensemble approximating the predicted distribution $p(\mathbf{x}_t\|\mathbf{y}_{1:t-1})$
$\left\{\mathbf{x}_{t\|t}^{(i)}\right\}_{i=1}^{N}$	Ensemble approximating the filtered distribution $p(\mathbf{x}_t\|\mathbf{y}_{1:t})$

The subscript s indicates the particle species (i for ions and e for electrons). T_s and n_s denote the temperature and density of species s, respectively. ρ is the normalized minor radius. $\langle\rangle$ denotes the magnetic surface average. $\mathcal{V}' = \partial\mathcal{V}/\partial\rho$ represents the plasma volume differentiated by ρ. V_{K_s} and V_s are the pinch velocities of energy and particles, respectively. In this study, we do not apply any physical model for the pinch velocities. Both V_{K_s} and V_s are set to zero in TASK, and only V_s is later modified in DA. S_s and P_s represent the particle and heat sources and sinks, respectively. In this study, S_s includes the NBI particle source and charge exchange, while P_s includes only NBI heating power. Heat and particle sources from NBI are pre-calculated using the OFMC code [13] for several time points, and loaded into TASK/TR. The particle sources from the transport and charge exchange of the peripheral neutral particles are evaluated using the AURORA module [14].

In this study, we assume ion and electron densities are approximately equal: $n_e \simeq n_i$. The thermal and particle diffusivities χ_s and D are assumed to be the sum of neoclassical element and turbulent components:

$$\chi_s = \chi_s^{\mathrm{TB}} + \chi_s^{\mathrm{NC}}, \tag{9}$$

$$D = D^{\mathrm{TB}} + D^{\mathrm{NC}}. \tag{10}$$

For the neoclassical components, we use Chang & Hinton's model [15, 16], which provides the ion thermal diffusivity as follows:

$$\chi_i^{\mathrm{NC}} = \frac{\varepsilon^{-3/2} q^2 \rho_i^2}{\tau_i} \left\{ \left[\frac{0.66 + 1.88\varepsilon^{1/2} - 1.54\varepsilon}{1 + 1.03\nu_{*i}^{1/2} + 0.31\nu_{*i}} \right] \left(1 + \frac{3}{2}\varepsilon^2\right) + \frac{0.59\nu_{*i}\varepsilon}{1 + 0.74\nu_{*i}\varepsilon^{3/2}} \left(1 + \frac{3}{2}\varepsilon^2 - \sqrt{1-\varepsilon^2}\right) \right\}, \tag{11}$$

where ε is the inverse aspect ratio, ρ_i the ion Larmor radius, τ_i the ion collision time, and ν_{*i} the ion collisionality. We assume the electron thermal diffusivity and the particle diffusivity using the ion thermal diffusivity:

$$\chi_e^{\mathrm{NC}} = \sqrt{\frac{m_e}{m_i}}\chi_i^{\mathrm{NC}}, \tag{12}$$

$$D^{\mathrm{NC}} = \frac{\chi_i^{\mathrm{NC}}}{10}. \tag{13}$$

For the turbulent components, we use following several turbulent transport models:

2.2.1 The CDBM model

Thermal diffusivity of the current diffusive ballooning mode (CDBM) model [17] is given as

$$\chi_s^{\mathrm{TB}} = C_s F(\hat{s}, \alpha)\alpha^{3/2}\frac{c^2}{\omega_{\mathrm{pe}}^2}\frac{v_A}{qR}, \tag{14}$$

where c is the speed of light, $\omega_{\mathrm{pe}} = \sqrt{n_e e^2/m_e\varepsilon_0}$ the electron plasma frequency, v_A the toroidal Alfvén velocity, q the safety factor, and R the major radius. The factor F is given by

$$F = \begin{cases} \dfrac{1}{\sqrt{2(1-2s')(1-2s'+3s'^2)}}, & \text{for } s' = \hat{s} - \alpha < 0, \\ \dfrac{(1+9\sqrt{2}s'^{5/2})}{\sqrt{2}(1-2s'+3s'^2+2s'^3)}, & \text{for } s' = \hat{s} - \alpha > 0, \end{cases} \tag{15}$$

where $\hat{s}$ is the magnetic shear and α is the normalized MHD pressure gradient. $\hat{s}$ and α are defined by

$$\hat{s} \equiv \frac{r}{q}\frac{\mathrm{d}q}{\mathrm{d}r}, \tag{16}$$

$$\alpha \equiv -q^2 R\frac{\mathrm{d}\beta}{\mathrm{d}r}, \tag{17}$$

where r is the minor radius and β is the ratio of the plasma pressure and the magnetic pressure, $(p_e + p_i)/(B_0^2/2\mu_0)$. The coefficient C_s is set to 12 following [18]. The effect of the magnetic curvature is neglected due to the large aspect ratio approximation.

2.2.2 The gyro-Bohm model

The gyro-Bohm model defines the thermal and particle diffusivities as

$$\chi_s^{\mathrm{TB}} = C_s\left(\frac{\rho_i}{L}\right)\left(\frac{T_s}{eB}\right), \tag{18}$$

$$D^{\mathrm{TB}} = C_d\left(\frac{\rho_i}{L}\right)\left(\frac{T_i}{eB}\right), \tag{19}$$

where the coefficients C_s and C_d are set to match the magnitudes of the diffusivities with those from the CDBM model. In this study, we use $C_e = 1.0$, $C_i = 1.2$, and $C_d = 2.0$.

2.2.3 The Alcator model

The Alcator model gives thermal and particle diffusivities as

$$\chi_s^{\mathrm{TB}} = C_s\frac{1}{n}, \tag{20}$$

$$D^{\mathrm{TB}} = C_d\frac{1}{n}. \tag{21}$$

The coefficients C_s and C_d are set using the same method as in the gyro-Bohm model, $C_s = 1.6 \times 10^{-19}$ [m^5/s] and $C_d = 1.0 \times 10^{-19}$ [m^5/s].

2.2.4 The constant model

We use the constant model only for the particle diffusivity. This model assumes the particle diffusivity as

$$D^{\mathrm{TB}} = C_d. \tag{22}$$

We employ $C_d = 0.8$ [m^2/s].

In the current system, MHD equilibrium is pre-calculated at discrete time points using the TOPICS code [19, 20], and the metric data, $\langle|\nabla\rho|\rangle$, $\langle|\nabla\rho|^2\rangle$ and $\mathcal{V}'$, are loaded to TASK/TR dynamically, but the time evolution is not solved in the simulation code. Safety factor q, which significantly affects Chang & Hinton's neoclassical diffusivity and the CDBM model turbulent diffusivity, is also pre-calculated in this study.

2.3 State vector

In this study, the state vector is defined as

$$\mathbf{x}_t = (\mathbf{T}_{\mathrm{e},t}, \mathbf{T}_{\mathrm{i},t}, \mathbf{n}_{\mathrm{e},t}, \mathbf{n}_{\mathrm{i},t}, \mathbf{c}_{\mathrm{e},t}, \mathbf{c}_{\mathrm{i},t}, \mathbf{d}_t, \mathbf{u}_t, \xi_{\mathrm{nn},t}, \xi_{\mathrm{Tn},t}, \mathbf{k}_{\mathrm{e},t}, \mathbf{k}_{\mathrm{i},t}). \quad (23)$$

Descriptions of these state variables are provided in Table 2. The 11-dimensional vectors represent radial profiles at ρ = 0.0, 0.1, …, 1.0. The 10-dimensional vector $\mathbf{u}_t$ represents the radial profile at ρ = 0.1, 0.2, …, 1.0 as the convection velocity at ρ = 0.0 is zero. The diffusivities, neutral density and temperature, and NBI heat deposition are optimized via factors $\mathbf{c}_{\mathrm{e},t}, \mathbf{c}_{\mathrm{i},t}, \mathbf{d}_t, \xi_{\mathrm{nn},t}, \xi_{\mathrm{Tn},t}, \mathbf{k}_{\mathrm{e},t}, \mathbf{k}_{\mathrm{i},t}$. These factors are initialized to one plus Gaussian noise and subsequently updated in the filtering procedure.

The magnitudes of the initial background error and the system noise for each variable are also listed in Table 2. The error and noise values with "%" represent the ratio of the standard deviation to the predicted value, while those without "%" denote the absolute standard deviation. These hyperparameters are determined by maximizing the marginal likelihood, which will be detailed in Sec. 2.5.

2.4 Observation models

In this study, we use the measurements from an NBI-heated plasma in JT-60U (shot: 44180) [21]. Figure 1 shows the time evolution of the NBI power. We define the time at which the first observation is obtained as 0.00 s. As shown in the figure, the injected heating power changes in steps. Radial profiles of the electron temperature T_e, electron density n_e, and ion temperature T_i are measured at six time points. We use the first measurement as the initial state and remaining five for DA.

An example of the observed profile data is shown in Fig. 2. As seen in the figure, locations of measured data are not evenly distributed along the radial direction. Moreover, the spatial distribution of observation points varies by time and for each variable. To assimilate these non-uniformly distributed measurements into the simulation model, the following two methods are used. Approach 1 follows the method proposed in the previous study [4], while Approach 2 introduces an improved method incorporates measurement error information.

Table 2. Variables in the state vector.

Variable	Description	Dimension	Initial background error	System noise
$\mathbf{T}_{\mathrm{e},t}$	Electron temperature at t	11	5.0%	5.0%
$\mathbf{T}_{\mathrm{i},t}$	Ion temperature at t	11	5.0%	5.0%
$\mathbf{n}_{\mathrm{e},t}$	Electron density at t	11	5.0%	5.0%
$\mathbf{n}_{\mathrm{i},t}$	Ion density at t	11	5.0%	5.0%
$\mathbf{c}_{\mathrm{e},t}$	Factor for electron thermal diffusivity at t	11	0.4	0.5
$\mathbf{c}_{\mathrm{i},t}$	Factor for ion thermal diffusivity at t	11	0.4	0.5
$\mathbf{d}_t$	Factor for particle diffusivity at t	11	0.4	0.5
$\mathbf{u}_t$	Correction term for particle convection velocity at t	10	0.5	0.2
$\xi_{\mathrm{nn},t}$	Factor for neutral density at the edge at t	1	0.1	0.1
$\xi_{\mathrm{Tn},t}$	Factor for neutral temperature at the edge at t	1	0.1	0.1
$\mathbf{k}_{\mathrm{e},t}$	Factor for electron NBI heat deposition at t	11	0.05	0.03
$\mathbf{k}_{\mathrm{i},t}$	Factor for ion NBI heat deposition at t	11	0.05	0.03

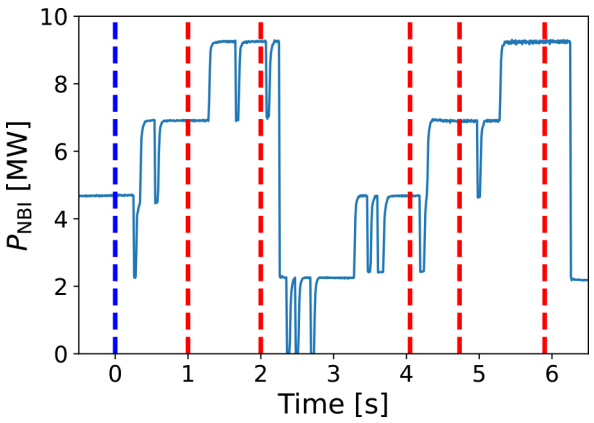

Fig. 1. Injected neutral beam power for shot 44180. Dashed lines indicate the time points at which temperature and density profiles were measured. We use the observation data at 0.00 s (blue line) as the initial state and the others (obtained at the times indicated by red lines) for the DA.

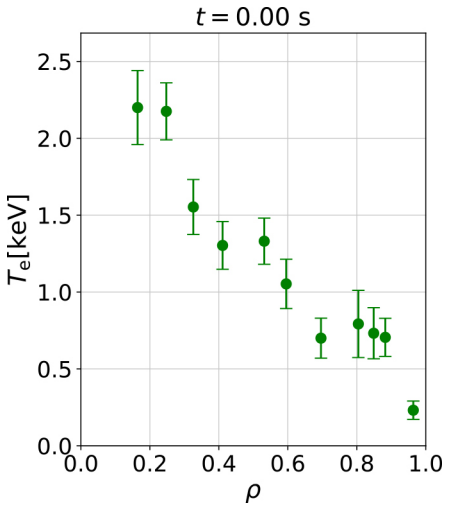

Fig. 2. Electron temperature profile at 0.00 s. Measurement points are unevenly distributed along the radial direction.

Approach 1: Profile fitting of the measurements

Firstly, we perform the profile fitting of the raw measurement data using Gaussian process regression (GPR) to produce uniformly spaced observations. Radial profiles of T_e, T_i and n_e are generated at ρ = 0.0, 0.1, …, 1.0, from which the observation vector is constructed as:

$$\mathbf{y}_t = \left(\widehat{\mathbf{T}}^{\mathrm{fit}}_{\mathrm{e},t}, \widehat{\mathbf{T}}^{\mathrm{fit}}_{\mathrm{i},t}, \widehat{\mathbf{n}}^{\mathrm{fit}}_{\mathrm{e},t}, \widehat{\mathbf{n}}^{\mathrm{fit}}_{\mathrm{i},t}\right). \tag{24}$$

Here, the hats denote the observation data, and the subscripts "fit" refer to the fitted profiles. Observation of the ion density $\widehat{\mathbf{n}}^{\mathrm{fit}}_{\mathrm{i},t}$ is decided according to the quasi-neutrality $\widehat{n}_\mathrm{i} \simeq \widehat{n}_\mathrm{e}$, as described in Sec. 2.2. The observation vector $\mathbf{y}_t$ is a 44-dimensional vector, composed of four 11-dimensional subvectors.

The observation matrix H_t is constructed simply by concatenating a 44 × 44 identity matrix with a 44 × 67 zero matrix:

$$H_t = (I_{44} \quad O_{44,67}), \tag{25}$$

where the number 67 represents the number of the non-observed parameters. Multiplying H_t with the system vector $\mathbf{x}_t$ extracts the observed components of $\mathbf{x}_t$.

The standard deviation of the Gaussian observation noise $\mathbf{w}_t$ is set proportional to the prediction error (i.e., the difference between the predicted and the observed values). The proportionality constant is set to 0.8 by maximizing the likelihood [4].

Approach 2: Linear interpolation using the observation matrix

The first approach is effective when measurement error information is unavailable. However, when the observed data contains measurement error information, as in this case, we can consider the credibility of each observation during the filtering step by reflecting the error information to the observation noise. In this section we propose a method which uses an improved observation model to consider the measurement error information.

To reflect the measurement error to the observation noise, we use the observed profiles without fitting them to uniform radial profiles. The observation vector is defined as:

$$\mathbf{y}_t = \left(\widehat{\mathbf{T}}^{\mathrm{raw}}_{\mathrm{e},t}, \widehat{\mathbf{T}}^{\mathrm{raw}}_{\mathrm{i},t}, \widehat{\mathbf{n}}^{\mathrm{raw}}_{\mathrm{e},t}, \widehat{\mathbf{n}}^{\mathrm{raw}}_{\mathrm{i},t}\right), \tag{26}$$

where the subscripts "raw" denote unprocessed data. Similar to the previous approach, the ion density profile $\widehat{\mathbf{n}}^{\mathrm{raw}}_{\mathrm{i},t}$ is decided by assuming $\widehat{n}_\mathrm{i} \simeq \widehat{n}_\mathrm{e}$. Since this observation vector $\mathbf{y}_t$ is a concatenation of the observed radial profile of each variable, the length of the vector is not constant and changes with times, depending on the shape of the actual measurements at each time step. While the observation matrix H_t is time constant in approach 1, in this approach we must set different H_t for each filtering step to assimilate $\mathbf{y}_t$ with time varying shape.

In this approach, the observation matrix H_t performs two functions:

1. Extracts the temperature and density components from the state vector $\mathbf{x}_t$.
2. Linear interpolates the extracted values to match the measurement locations in $\mathbf{y}_t$.

The observation matrix H_t has the following block form:

$$H_t = \begin{pmatrix} M_{T_\mathrm{e}} & & & \\ & M_{T_\mathrm{i}} & O & \\ & O & M_{n_\mathrm{e}} & O \\ & & & M_{n_\mathrm{i}} \end{pmatrix}, \tag{27}$$

where the zero matrix on the right has the shape of (dimension of the observation vector) × (number of the non-observed variables in the state vector). $M_x (x = T_\mathrm{e}, T_\mathrm{i}, n_\mathrm{e}, n_\mathrm{i})$ are linear interpolation matrices that maps state variables to measurement points. The observation model is thus written as:

$$\begin{pmatrix} M_{T_\mathrm{e}} & & & \\ & M_{T_\mathrm{i}} & O & \\ & O & M_{n_\mathrm{e}} & O \\ & & & M_{n_\mathrm{i}} \end{pmatrix} \begin{pmatrix} \mathbf{T}_{\mathrm{e},t} \\ \mathbf{T}_{\mathrm{i},t} \\ \mathbf{n}_{\mathrm{e},t} \\ \mathbf{n}_{\mathrm{i},t} \\ \mathbf{c}_{\mathrm{e},t} \\ \mathbf{c}_{\mathrm{i},t} \\ \mathbf{d}_t \\ \mathbf{u}_t \\ \xi_{\mathrm{nn},t} \\ \xi_{\mathrm{Tn},t} \\ \mathbf{k}_{\mathrm{e},t} \\ \mathbf{k}_{\mathrm{i},t} \end{pmatrix} + \mathbf{w}_t = \begin{pmatrix} \widehat{\mathbf{T}}^{\mathrm{raw}}_{\mathrm{e},t} \\ \widehat{\mathbf{T}}^{\mathrm{raw}}_{\mathrm{i},t} \\ \widehat{\mathbf{n}}^{\mathrm{raw}}_{\mathrm{e},t} \\ \widehat{\mathbf{n}}^{\mathrm{raw}}_{\mathrm{i},t} \end{pmatrix}. \tag{28}$$

Here, the observation noise is given by

$$\mathbf{w}_t = \begin{pmatrix} \mathbf{w}_{T_\mathrm{e},t} \\ \mathbf{w}_{T_\mathrm{i},t} \\ \mathbf{w}_{n_\mathrm{e},t} \\ \mathbf{w}_{n_\mathrm{i},t} \end{pmatrix} \sim N(\mathbf{0}, R_t), \tag{29}$$

where R_t is the diagonal covariance matrix of observation noise, constructed from the known measurement errors.

2.5 Simulation configuration

Before estimation with the EnKF, hyperparameters such as the initial background error, system noise levels, and ensemble size must be determined. We determine these hyperparameters by maximizing the log-likelihood. The log-likelihood for a parameter θ is estimated by [22]:

$$\begin{aligned} l(\theta) &= \log p(\mathbf{y}_{1:T}|\theta) \\ &= \sum_{t=1}^{T} \log p(\mathbf{y}_t|\mathbf{y}_{1:t-1}, \theta) \\ &\simeq \sum_{t_k=t_1}^{t_M} \log\left(\sum_{n=1}^{N} \alpha_{t_k}^{(n)}\right) - M\log N, \end{aligned} \tag{30}$$

where

$$\alpha_{t_k}^{(n)} = \frac{1}{\sqrt{(2\pi)^l \left| R_{t_k|} \right|}} \times \exp\left\{ -\frac{1}{2} \left(\mathbf{y}_{t_k} - H\mathbf{x}_{t_k|t_{k-1}}^{(n)} \right)^{\mathrm{T}} R_{t_k}^{-1} \left(\mathbf{y}_{t_k} - H\mathbf{x}_{t_k|t_{k-1}}^{(n)} \right) \right\}. \quad (31)$$

Since there are 12 system noise parameters, uniform search of the parameter space is impractical. To reduce the number of the evaluation cases, we restrict the search domain to fewer dimensions using the following three assumptions. Firstly, we fix the noise intensities of the observed variables at 5% of the predicted values. Secondly, assuming that the predictions of neutral density and temperature at the edge and NBI heating deposition are sufficiently accurate, we set the low noise intensities. Finally, we set the noise intensities of c_e, c_i, and d to be the same value. Under these simplifications, we conduct simulations with various noise intensities for c_e, c_i, d and u, and select the configuration that yields the highest log-likelihood. In this hyperparameter tuning step, the ensemble size is fixed at 520. We use the CDBM model for thermal turbulent transport and the constant model for particle turbulent transport. For the observation model, we use approach 2 presented in Sec. 2.4. The determined values for initial background error and system noise are shown in Table 2.

The ensemble size is then determined. A larger ensemble size improves estimation accuracy, but also increases computational cost. Figure 3 shows the log likelihood evaluated for ensemble size of 130, 260, 390, 520, 780, and 1,300. The results indicate that the log likelihood saturates around an ensemble size of 520; therefore we set the size to 520.

3. Results

3.1 Improvement of the observation model

We simulate particle and energy transport of bulk ions and electrons in an NBI discharge, using the DA method with two observation models introduced in the Sec. 2.4. Figure 4 shows the simulated ion temperature profiles at t = 2.00 s and t = 4.73 s. The black lines correspond to the observation model of approach 1 (which does not account for measurement error), and the blue lines correspond to approach 2 (which does account for measurement error). The orange lines represent results from simulations using TASK without DA. All simulations employ the CDBM model for thermal diffusivity and the constant model for particle diffusivity. The green dots with error bars represent observed values and their corresponding measurement errors.

It is evident from the figure that the reproducibility of the observations increases by applying the DA. While both DA

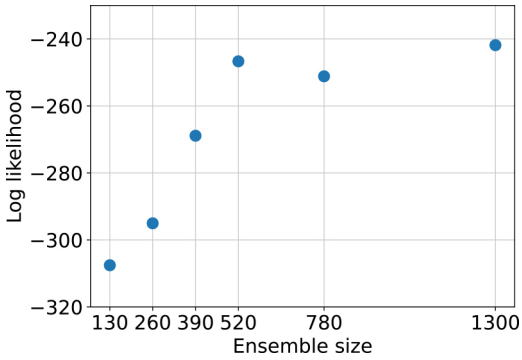

Fig. 3. Log likelihood for ensemble size of 130, 260, 390, 520, 780, and 1,300.

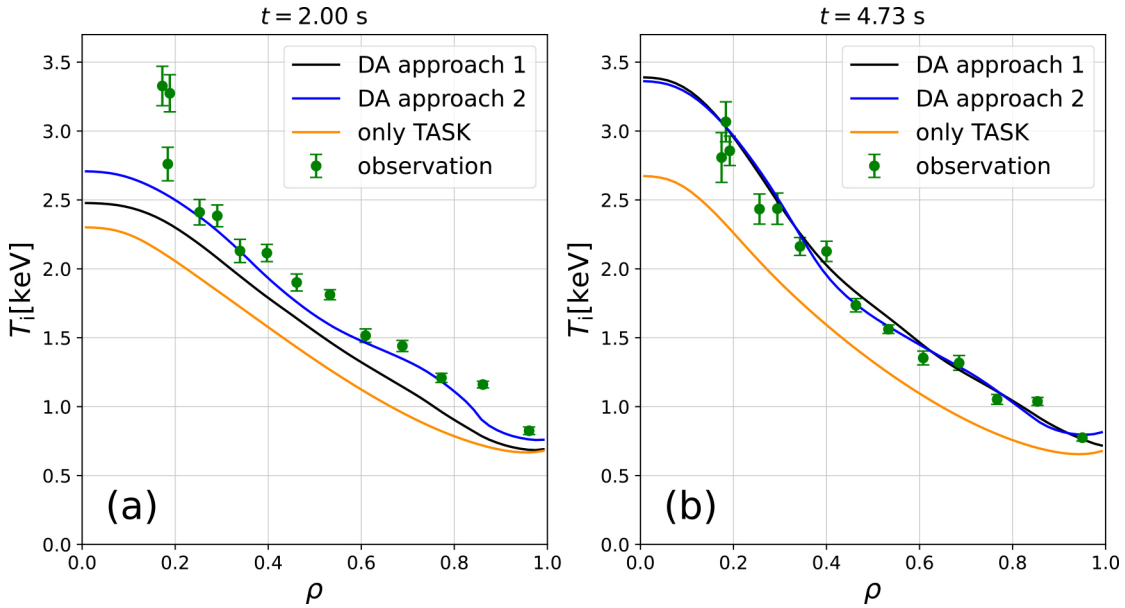

Fig. 4. A comparison of the results at (a) time 2.00 s and (b) time 4.73 s. The orange lines represent the results from the simulation using TASK without DA. The black and blue lines represent the results of the simulations with DA. The black ones correspond to the observation model of approach 1 (not reflecting the observation error), and the blue ones correspond to the observation model of approach 2 (reflecting the observation error). The green dots and error bars represent the observed values and the measurement errors.

approaches reproduce the observed ion temperature profile well at t = 4.73 s, differences in accuracy appear at t = 2.00 s. Also, the steep peak at the core region in the observation at 2.00 s is not reproduced in the results of either approaches. There are two likely reasons for these results: the number of DA performed and the off-axis heating at 2.00 s. In general, prediction accuracy improves with more assimilation steps. The prediction at t = 2.00 s reflects only a single assimilation at t = 1.00 s, whereas the prediction at t = 4.73 s is the result of three assimilations at t = 1.00, 2.00 and 4.05 s, which can be considered as the reason for the difference in accuracy.

Another factor is that the observed ion temperature profile at t = 2.00 s peaks at the core region despite off-axis heating. Figure 5 shows the NBI heating deposition profiles on ion at t = 2.00 s and 4.73 s. The heating peak at t = 2.00 s is located near ρ = 0.4, with relatively low power deposited in the core. Nevertheless, the observed ion temperature rises sharply at the core, indicating a possible significant reduction in heat diffusion near the magnetic axis. It has also been reported that during off-axis heating, heat transport is not completely diffusive, and some degree of inward heat flux is required to explain the observation [23]. However, such effects are not captured in the present simulation model.

Despite this limitation, approach 2 provides relatively accurate predictions at $\rho > 0.2$ at t = 2.00 s. This difference between two approaches stems from their respective DA procedures at t = 1.00 s. Figure 6 compares the optimizations of T_i and c_i at t = 1.00 s. The green dots in (a1, a2) indicate the assimilated observations. As mentioned in Sec. 2.4, approach 1 uses fitted profiles, while approach 2 uses raw measurements. Error bars indicate the standard deviations of observation noise. The difference in assimilated observation results in distinct filtered profiles of T_i (the red lines). In the filtered profile of T_i in approach 2, transport-barrier-like structures emerge around ρ = 0.4 and 0.8, whereas approach 1 yields smoother profiles. To reproduce these features, c_i is optimized accordingly in the filtering step. (b2) shows that the filtered profile of c_i is significantly suppressed near ρ = 0.8 in approach 2, reducing diffusivity in that region. The dip around $\rho = 0.2 \sim 0.3$ is also deeper in the filtered profile than that in approach 1.

In approach 2, observation noise is set based on actual measurement errors, while in approach 1 it is based on prediction error. In charge exchange recombination spectroscopy (CXRS), the core temperature measurements are generally less accurate than those at the edge [24]. This is considered to be another reason of the difference in the filtered profile of T_i between two approaches, since it leads to the focus of the DA shifting on improving the prediction accuracy of the periphery, rather than the core region.

Figure 7 shows the time evolution of electron temperature, hydrogen ion temperature and electron density. The improved observation model (approach 2) is employed in this result. We can see that the graph is discontinuous at each time point at which the observation is obtained. This shows that the temperature and density profiles are optimized by DA. The hatched areas show the standard deviations of the probability distribution. It is seen that the standard deviation is smaller after each assimilation. This is because the filtering procedure tend to narrow the width of probability distribution.

This study uses only the H-mode phase plasmas in the shot number 44180. ASTI is also capable of predicting L-mode plasmas, although the performance of each turbulent transport model requires further evaluation. The current system can follow L-H transitions by assimilating the observation after the transition, but the prediction of transition in advance requires the implementation of transition model to the system model. A detailed measurement that captures the structure of

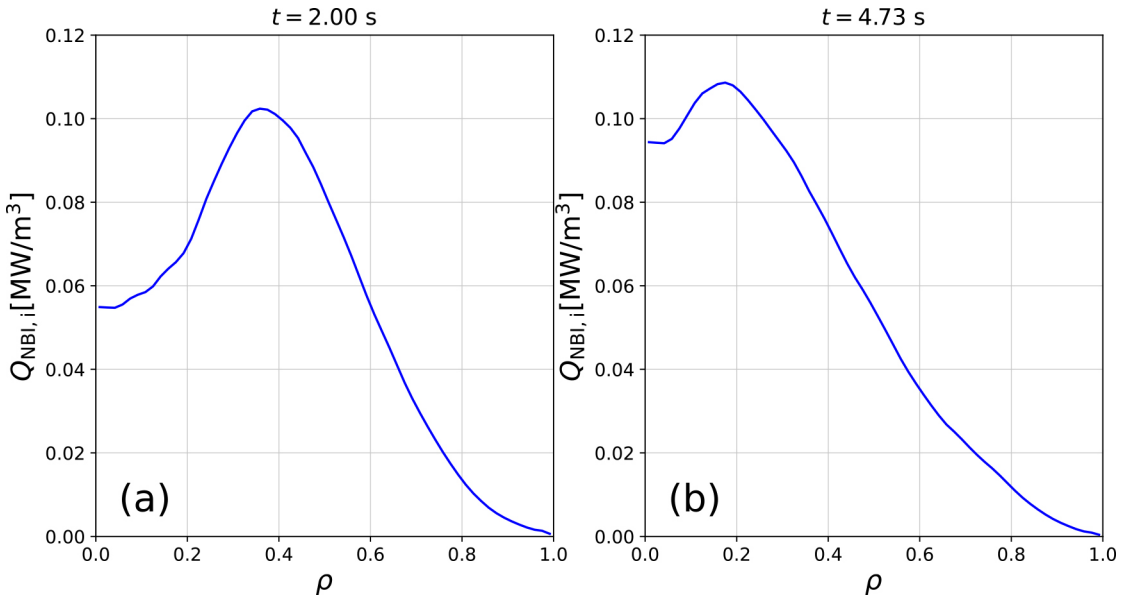

Fig. 5. Radial profile of the NBI heating power deposition on hydrogen ions at (a) 2.00 and (b) 4.73 s. The power deposition is pre-calculated by OFMC code. Off-axis heating is performed at 2.00 and on-axis heating is performed at 4.73 s.

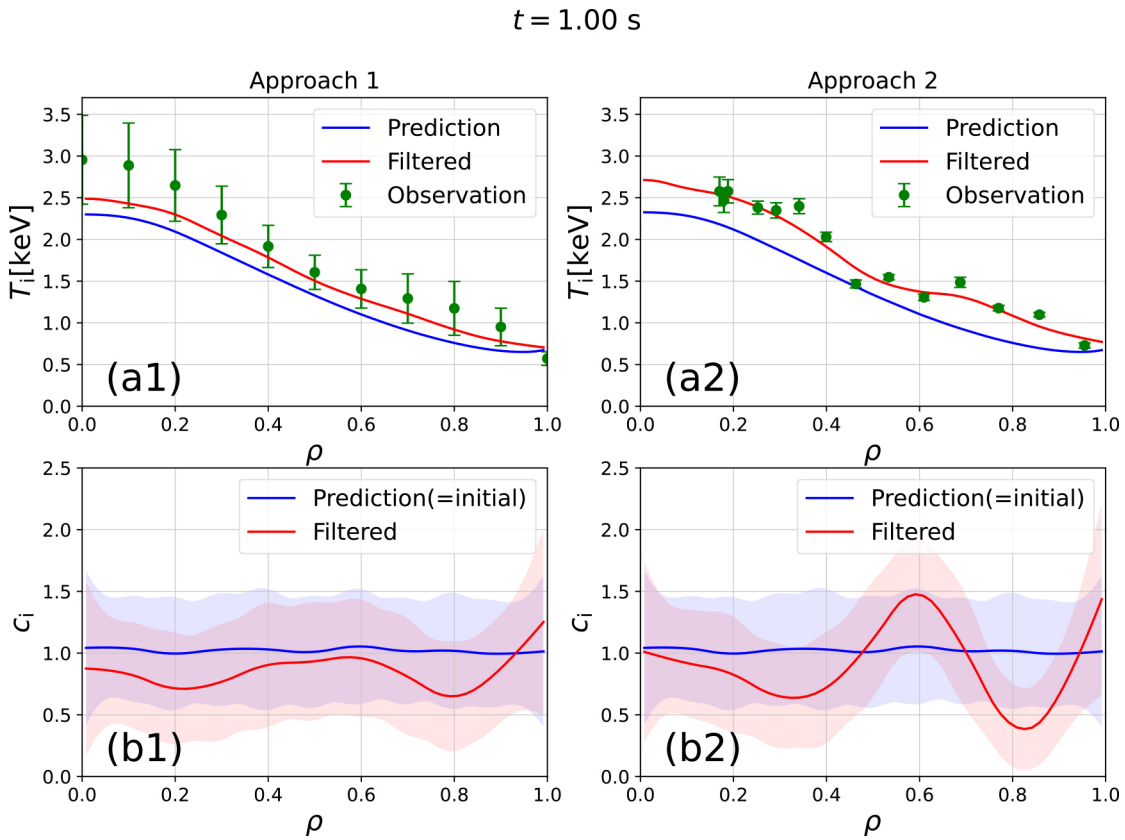

Fig. 6. (a1, a2) Radial profiles of T_i at time 1.00 s of approach 1 and 2, respectively. The green dots with error bars represent the observation and the observational errors, used for the assimilation in each approach. (b1, b2) Radial profiles of c_i at time 1.00 s of approach 1 and 2, respectively. The blue lines show the predicted profile and the red lines show the filtered profile. Since c_i changes only by the filtering, the predicted value of c_i at time 1.00 s is the same as the initial value, which is unity plus Gaussian noise. The hatched areas in (b1, b2) show the standard deviations of the probability distribution.

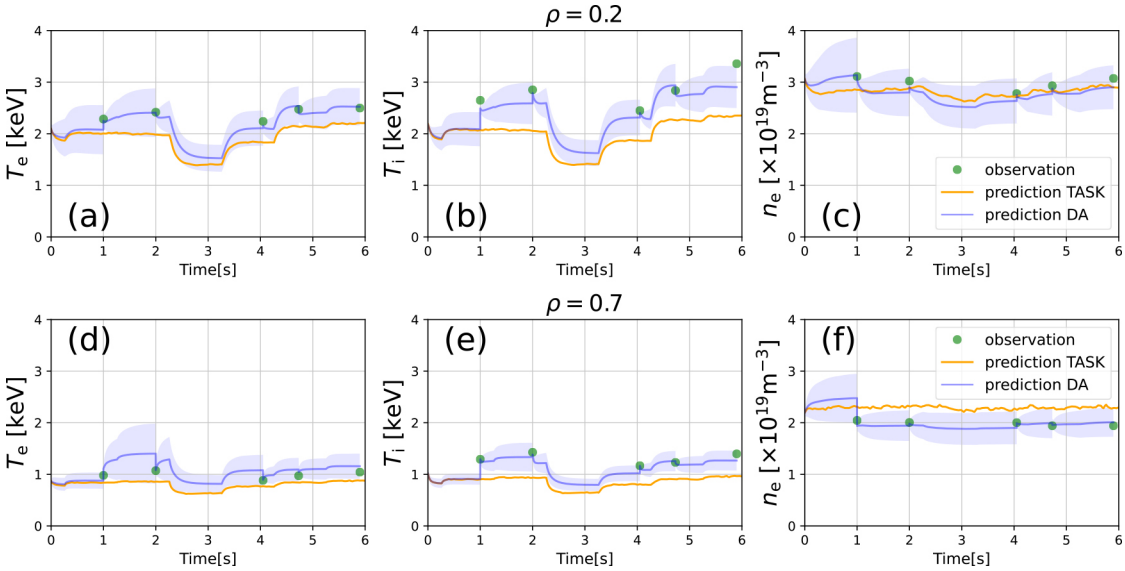

Fig. 7. Time evolution of electron temperature, hydrogen ion temperature and electron density at (a, b, c) ρ = 0.2 and (d, e, f) ρ = 0.7. The green dots indicate the observed values, the orange lines are the results of the simulation only by TASK, and the blue lines indicate the results from the simulation with DA. The hatched areas show the standard deviations of the probability distribution.

the H-mode pedestal will also be required to optimize the model. We believe that even a simplified model, when paired with detailed observations, can make good predictions of the transition phenomena.

3.2 Comparison of turbulent transport models

Estimating turbulent transport is essential for accurate prediction of burning plasmas. However, since this DA technique aims for real-time prediction and control, high-fidelity

and computationally expensive models are unsuitable. Instead, we optimize the diffusivities using DA to ensure accuracy with computationally inexpensive models. In this section, performances of turbulent transport models introduced in Sec. 2.2 are compared. We test the CDBM, gyro-Bohm, and Alcator models for thermal diffusivity. The same model is used for both ions and electrons. For particle diffusivity, we use the constant, gyro-Bohm, and Alcator models. We perform simulations with a total of nine combinations, three for each of thermal and particle diffusive models. The log-likelihood for each combination is summarized in Table 3.

The factors c_e, c_i, and d are used to optimize the diffusivities χ_e^{TB}, χ_i^{TB}, and D^{TB} respectively. The space-time distribution of these factors contain essential information on predictive performance of each turbulent transport model. Magnitude of variance in the space-time distributions of these factors indicate how strongly the diffusivity distribution calculated from the models have been corrected by the DA. Here, we allow dispersion for the spatial distribution, and only consider variance of the temporal distribution. Figure 8 shows the radial profile of time and ensemble averaged c_e,

$$\frac{1}{M}\sum_{t_k=t_1}^{t_M}\frac{1}{N}\sum_{n=1}^{N} c_e^{(n)}, \tag{32}$$

for each turbulent thermal transport model. Here, $t_1 \sim t_M$ indicate the times when DA takes place, and N is the size of the ensemble. Hatched areas on the graphs indicate the standard deviation of ensemble-averaged c_e over time:

$$\sqrt{\frac{1}{M}\sum_{t_k=t_1}^{t_M}\left(\frac{1}{N}\sum_{n=1}^{N} c_e^{(n)} - \frac{1}{M}\sum_{t_k=t_1}^{t_M}\frac{1}{N}\sum_{n=1}^{N} c_e^{(n)}\right)^2}. \tag{33}$$

Width of the hatched region indicates how strongly the correction of c_e is made at each time DA takes place. In the results shown in Fig. 8, the constant model is used as the turbulent particle transport model. Similar radial profile shapes are observed with the CDBM and Alcator models, though the Alcator model yields slightly smaller deviations. In contrast, the turbulent electron thermal diffusivity of the gyro-Bohm model at the edge is corrected by the factor of over 5 on average. Its standard deviation is also much larger than the case with other 2 models, indicating stronger corrections at each filtering procedure, suggesting poor baseline performance of the gyro-Bohm model. The "radial mean of std" at the top of each graph indicates the radial mean of the value of Eq. (33) normalized by the average magnitude. This parameter is considered to represent the predictive performance of each turbulent model. Table 4 shows the radial means of standard deviation for each combination of turbulent models.

From both the log-likelihood and the standard deviation of the optimization factors, the CDBM model is proven to offer high predictive accuracy. This aligns with previous findings [25] which demonstrated the effectiveness of the CDBM model for JT-60U discharges. Log-likelihood is the highest for the case in which the CDBM model and the constant model are used for thermal and particle transport, respectively. The strong performance of the constant model may be due to the relatively stable electron density observed in this discharge, which reduces the need for time-dependent corrections. Little time variation in observed density also leads to similar results of the constant and Alcator models.

Table 3. Log likelihood of the simulations with combinations of turbulent transport models.

		Thermal		
		CDBM	gyro-Bohm	Alcator
Particle	constant	−246.7	−337.3	−307.1
	gyro-Bohm	−337.0	−455.0	−528.0
	Alcator	−278.8	−356.4	−366.6

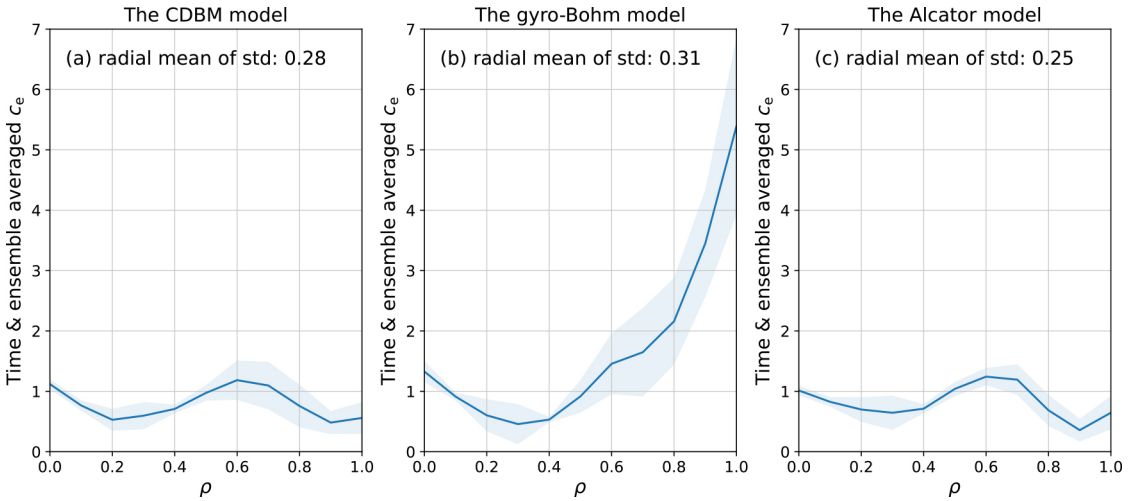

Fig. 8. Radial profile of time and ensemble averaged c_e obtained from simulations with (a) the CDBM model, (b) the gyro-Bohm model and (c) the Alcator model as the turbulent thermal transport model. The hatched parts represent the standard deviation of ensemble averaged c_e over time. "Radial mean of std" on the top of the graph shows the radial mean of the value of Eq. (33).

Table 4. Radial means of the standard deviation of the optimization factors c_e, c_i, and d over time, for each combination of turbulent models.

	c_e	Thermal		
		CDBM	gyro-Bohm	Alcator
Particle	constant	0.28	0.31	0.25
	gyro-Bohm	0.27	0.37	0.23
	Alcator	0.26	0.30	0.23

	c_i	Thermal		
		CDBM	gyro-Bohm	Alcator
Particle	constant	0.35	0.72	0.36
	gyro-Bohm	0.33	0.49	0.31
	Alcator	0.37	0.60	0.32

	d	Thermal		
		CDBM	gyro-Bohm	Alcator
Particle	constant	0.14	0.21	0.16
	gyro-Bohm	0.18	0.37	0.23
	Alcator	0.16	3.33	0.19

4. Conclusions

In this study, we have applied the DA system ASTI to predictive simulations of JT-60U plasmas. The results demonstrate that DA is effective in simulating particle and energy transport in tokamak plasmas. Simulations incorporating DA shows significantly better agreement with observations than those conducted using TASK code alone. We have proposed an improved observation model that incorporates measurement error information into filtering process. This observation model achieved better agreement with observations compared to the previous model. The new approach is effective in situations where the measurement errors are available from diagnostic data.

The simulations with DA also allows for quantitative comparison of different turbulent transport models. We have carried out the comparison and evaluation of several turbulent transport models. Our analysis have revealed that the combination of the CDBM model for thermal turbulent transport and the constant model for particle turbulent transport provides good predictions. Simulations using the CDBM model yields relatively accurate temperature predictions regardless of the particle transport model used, confirming the robustness of this model. The good performance of the constant model in this study is attributed to the limited temporal variation in the observed electron density profile. Future work will involve further evaluation of the particle turbulent transport model under conditions with larger density variations.

In the present simulations, time evolution of equilibrium is not solved in the code, and the pre-calculated data is loaded dynamically to TASK/TR code. In tokamak simulations, it is important to accurately assess time evolution of magnetic field and equilibrium. Our next step is to extend ASTI into a coupled transport-equilibrium prediction system for tokamak plasmas by implementing an equilibrium solver.

Acknowledgements

This work was supported by JSPS KAKENHI (Grant Numbers JP23K19033 and JP24K00609).

[1] F. Felici *et al.*, Plasma Phys. Control. Fusion **54**, 025002 (2012).
[2] F. Felici *et al.*, Nucl. Fusion **58**, 096006 (2018).
[3] J. Artaud *et al.*, Nucl. Fusion **58**, 105001 (2018).
[4] Y. Morishita *et al.*, Nucl. Fusion **60**, (2020).
[5] Y. Morishita *et al.*, Comput. Phys. Commun. **274**, 108287 (2022).
[6] Y. Morishita *et al.*, J. Comput. Sci. **72**, 102079 (2023).
[7] Y. Morishita *et al.*, Sci. Rep. **14**, 137 (2024).
[8] T. Blanken *et al.*, Fusion Eng. Des. **126**, 87 (2018).
[9] T. Bosman *et al.*, Fusion Eng. Des. **170**, 112510 (2021).
[10] M.C.C. Messmer *et al.*, Plasma Phys. Control. Fusion **61**, 035011 (2019).
[11] G. Evensen, Ocean Dyn. **53**, 343 (2003).
[12] A. Fukuyama, *Proc. of 20th Fusion Energy Conf.* (Vilamoura, Portugal, 2004).
[13] K. Tani *et al.*, J. Phys. Soc. Jpn. **50**, 1726 (1981).
[14] M. Hughes *et al.*, J. Comput. Phys. **28**, 43 (1978).
[15] C.S. Chang *et al.*, Phys. Fluids **25**, 1493 (1982).
[16] C.S. Chang *et al.*, Phys. Fluids **29**, 3314 (1986).
[17] A. Fukuyama *et al.*, Plasma Phys. Control. Fusion **37**, 611 (1995).
[18] M. Honda *et al.*, Nucl. Fusion **46**, 580 (2006).
[19] M. Kikuchi *et al.*, Plasma Phys. Control. Fusion **37**, 1215 (1995).
[20] H. Shirai *et al.*, Plasma Phys. Control. Fusion **42**, 1193 (2000).
[21] H. Urano *et al.*, Nucl. Fusion **53**, 083003 (2013).
[22] G. Ueno *et al.*, Q. J. R. Meteorol. Soc. **142**, 2055 (2016).
[23] P. Gohil *et al.*, Nucl. Fusion **38**, 425 (1998).
[24] Y. Koide *et al.*, Rev. Sci. Instrum. **72**, 119 (2001).
[25] N. Hayashi *et al.*, Nucl. Fusion **57**, 126037 (2017).