

# Application of Neural Ordinary Differential Equations for ITER Burning Plasma Dynamics

Zefang Liu<sup>1</sup>, Weston M. Stacey<sup>1</sup>

<sup>1</sup>Georgia Institute of Technology  
Atlanta, GA 30332, USA

liuzefang@gatech.edu, weston.stacey@nre.gatech.edu

## Abstract

The dynamics of burning plasmas in tokamaks are crucial for advancing controlled thermonuclear fusion. This study applies the NeuralPlasmaODE, a multi-region multi-timescale transport model, to simulate the complex energy transfer processes in ITER deuterium-tritium (D-T) plasmas. Our model captures the interactions between energetic alpha particles, electrons, and ions, which are vital for understanding phenomena such as thermal runaway instability. We employ neural ordinary differential equations (Neural ODEs) for the numerical derivation of diffusivity parameters, enabling precise modeling of energy interactions between different plasma regions. By leveraging transfer learning, we utilize model parameters derived from DIII-D experimental data, enhancing the efficiency and accuracy of our simulations without training from scratch. Applying this model to ITER’s inductive and non-inductive operational scenarios, our results demonstrate that radiation and transport processes effectively remove excess heat from the core plasma, preventing thermal runaway instability. This study underscores the potential of machine learning in advancing our understanding and control of burning plasma dynamics in fusion reactors.

## 1 Introduction

The dynamics of burning plasmas in tokamaks are critical for advancing controlled thermonuclear fusion. In the International Thermonuclear Experimental Reactor (ITER) (IAEA 2002), deuterium-tritium (D-T) fusion reactions generate 14.1 MeV neutrons and 3.5 MeV fusion alpha particles (Green et al. 2003). These alpha particles, confined by magnetic fields, transfer their energy primarily to core electrons before transferring it to ions. The heated core electrons emit various types of radiation, including electron cyclotron radiation (ECR), bremsstrahlung, and impurity radiation, which rapidly dissipate energy compared to the slower transport of energy to the edge. However, the energized electrons and remaining alpha particles also heat the core ions through collisional processes, potentially increasing fusion reactivity and leading to more fusion alpha particles. This positive feedback loop poses a risk of thermal runaway instability in ITER. Consequently, radiation and transport processes between different plasma regions with varying timescales are crucial for effective burning plasma operation.

In previous research by Liu and Stacey (Stacey 2021; Liu and Stacey 2021; Liu 2022; Liu and Stacey 2024),

a multi-region multi-timescale burning plasma dynamics model has been developed to simulate these complex interactions in tokamaks. Different regions, including the core, edge, scrape-off layer (SOL), and divertor, are modeled as separate nodes. This model incorporates essential mechanisms such as auxiliary heating, fusion alpha heating, radiations, collisional energy transfer, transport, and ion orbit loss (IOL). Neural ordinary differential equations (Neural ODEs) (Chen et al. 2018) have been employed to optimize the parametric diffusivity formula, and this multinodal model has been validated for deuterium plasmas from DIII-D (Liu and Stacey 2024).

In this study, we extend this multinodal burning plasma dynamics model, NeuralPlasmaODE<sup>1</sup>, to analyze ITER D-T plasmas. We make several assumptions suitable for ITER burning plasmas, modeling deuterons, tritons, alpha particles, and electrons in the core and edge nodes. This specialized multinodal model is applied to simulate inductive and non-inductive operational scenarios, examining the dynamics and energy flows among species through various mechanisms with multiple timescales. By leveraging Neural ODEs for the numerical derivation of diffusivity parameters, we enhance the model’s precision and efficiency. Additionally, we employ transfer learning to utilize model parameters derived from fitting DIII-D data and avoid the need for training from scratch. The simulation results indicate that radiation and transport processes efficiently remove excess heat from fusion alpha particles, preventing thermal runaway instability. This study highlights the potential of machine learning in improving our understanding and control of burning plasma dynamics in fusion reactors.

## 2 Related Work

Research on fusion-relevant plasmas, particularly focusing on burning plasma dynamics, has been extensive. Wang et al. (1997) used a 1.5D Tokamak Transport Simulation Code (TTSC) to simulate ITER burning plasmas, highlighting potential issues with fusion power excursions due to slight confinement improvements. Green et al. (2003) categorized burning plasma physics in ITER into energetic particle effects, self-heating phenomena, and reactor-scale physics, assessing various drive scenarios. Cordey et al. (2005) im-

---

<sup>1</sup>Repository: <https://github.com/zefang-liu/NeuralPlasmaODE>

proved the ELMy H-mode scaling law for energy confinement time using principal component regression, enhancing ITER performance predictions. Stacey (2007) reviewed abrupt transition phenomena in plasmas and studied theoretical thermal instabilities. Hill and Stacey (2017) developed a confinement tuning model specific to DIII-D experiments, showing improved temperature simulations over the ITER-98 scaling law. Hill (2019) further investigated control mechanisms for plasma power excursion in ITER, identifying electron cyclotron radiation (ECR) as a critical passive control mechanism, and presented a framework for multinodal dynamics modeling. Stacey (2021) introduced a spatially coarse nodal space-time dynamics model for burning plasmas, outlining core and edge particle and energy balance equations. Liu and Stacey (2024) proposed a multi-region multi-timescale transport model employing neural ordinary differential equations (Neural ODEs) to simulate intricate energy transfer processes in tokamaks, validated against DIII-D experimental data. These studies lay the foundation for our work, which builds a practical multinodal model using Neural ODEs and machine learning to simulate ITER plasmas.

### 3 Burning Plasma Dynamics Model

To effectively and efficiently simulate the complex energy transfer processes in ITER burning deuterium-tritium (D-T) plasmas, we implement a multi-region multi-timescale burning plasma dynamics model.

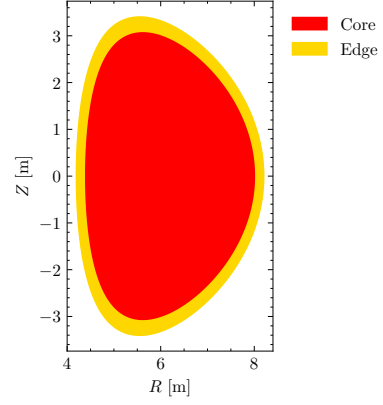
#### 3.1 ITER Plasma Geometry

The geometry of the burning plasma dynamics model is fundamental for simulating ITER plasmas. A conventional tokamak is viewed as a torus with a circular cross-section. The torus is divided into three regions: the core, edge, and scrape-off layer (SOL), following the flux surfaces from the inner to the outer side, while the divertor region is ignored in this geometry. Each region is represented as a separate node in this model, as illustrated in Figure 1. In this model, each node is a toroidal shell with interfaces represented as torus surfaces. The minor radii for the surfaces  $r_{\text{core}}$ ,  $r_{\text{edge}}$ , and  $r_{\text{sol}}$  correspond to the core, edge, and SOL surfaces respectively. The radial distances between these nodes are defined as  $\Delta r_{\text{core-edge}}$  for the core to edge,  $\Delta r_{\text{edge-sol}}$  for the edge to SOL, and  $\Delta r_{\text{sol-div}}$  for the distance from the SOL node center to its outer surface. These radial distances are calculated as  $\Delta r_{\text{core-edge}} = r_{\text{edge}}/2$ ,  $\Delta r_{\text{edge-sol}} = (r_{\text{sol}} - r_{\text{core}})/2$ , and  $\Delta r_{\text{sol-div}} = (r_{\text{sol}} - r_{\text{edge}})/2$ . Additionally, the normalized minor radius is defined as  $\rho = r/a$ , where  $a$  is the minor radius of the plasma.

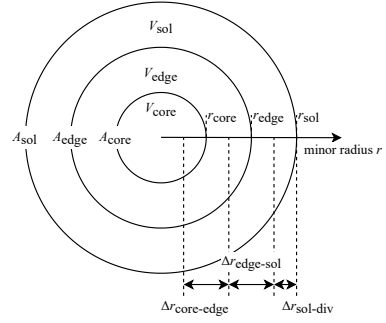
#### 3.2 Assumptions for ITER Plasmas

To simulate ITER D-T plasmas, the multinodal burning plasma model (Liu and Stacey 2024) requires several necessary modifications to balance accuracy and computational efficiency:

- Only the core and edge regions are modeled as separate nodes.



(a) Cross section of an ITER plasma



(b) Geometry of the multinodal model

Figure 1: Multinodal model geometry of tokamak plasmas, where the first figure shows the cross section of an ITER plasma, and the second figure is the simplified geometry in the multinodal model.

- The model computes the behavior of deuterons, tritons, alpha particles, and electrons, while including helium, beryllium, and argon as impurity particles.
- The simulation focuses solely on the D-T fusion reaction, while all atomic and molecular processes, as well as neutral and recycling particles, are neglected.
- The triton particle and thermal diffusivities are assumed to be equal to those of deuterons.
- A delay mechanism for fusion alpha heating is implemented in the core node.
- Predetermined energy deposition profiles for neutral beam injection (NBI) and radiofrequency (RF) heating (Kessel et al. 2007) are utilized.
- Electron cyclotron radiation (ECR) (Albajar, Johner, and Granata 2001; Albajar, Bornatici, and Engelmann 2009) parameters, fitted for typical ITER profiles, are applied.
- Ion orbit loss (IOL) (Stacey 2011, 2013; Wilks and Stacey 2016) is considered only for the edge node, with approximated loss timescales equated to transport times.

### 3.3 Particle and Energy Balance Equations

Based on the aforementioned assumptions, the multinodal burning plasma model is presented into the following particle and energy balance equations.

**Particle Balance Equations** Particle balance equations for deuterons (D), tritons (T), and alpha particles ( $\alpha$ ) in the core and edge nodes are

$$\frac{dn_{\sigma}^{\text{core}}}{dt} = S_{\sigma,\text{ext}}^{\text{core}} + S_{\sigma,\text{fus}}^{\text{core}} + S_{\sigma,\text{tran}}^{\text{core}}, \quad (1)$$

$$\frac{dn_{\sigma}^{\text{edge}}}{dt} = S_{\sigma,\text{ext}}^{\text{edge}} + S_{\sigma,\text{fus}}^{\text{edge}} + S_{\sigma,\text{tran}}^{\text{edge}} + S_{\sigma,\text{IOL}}^{\text{edge}}, \quad (2)$$

where  $\sigma \in \{D, T, \alpha\}$ . The electron densities are computed from the charge neutrality:

$$n_e^{\text{node}} = z_D n_D^{\text{node}} + z_T n_T^{\text{node}} + z_{\alpha} n_{\alpha}^{\text{node}} + \sum_z z_z n_z^{\text{node}}, \quad (3)$$

where  $\text{node} \in \{\text{core}, \text{edge}\}$  and  $z$  is for impurity particles. The particle source and sink terms are as follows.

The **external particle sources** are computed by summing the neutral beam injection (NBI), gas puffing (GAS), shatter pellet injection (SPI), and other external particle sources:

$$S_{\sigma,\text{ext}}^{\text{node}} = S_{\sigma,\text{NBI}}^{\text{node}} + S_{\sigma,\text{GAS}}^{\text{node}} + S_{\sigma,\text{SPI}}^{\text{node}} + \dots, \quad (4)$$

where  $\sigma \in \{D, T, \alpha\}$ . In this research, only particle sources from the NBI (Wesson and Campbell 2011) are considered. The **fusion terms** are computed from the D-T fusion reactions (Stacey 2012) by

$$S_{D,\text{fus}}^{\text{node}} = S_{T,\text{fus}}^{\text{node}} = -n_D^{\text{node}} n_T^{\text{node}} \langle \sigma v \rangle_{\text{fus}}, \quad (5)$$

$$S_{\alpha,\text{fus}}^{\text{node}} = n_D^{\text{node}} n_T^{\text{node}} \langle \sigma v \rangle_{\text{fus}}, \quad (6)$$

where  $\langle \sigma v \rangle_{\text{fus}}$  is the fusion reactivity (Bosch and Hale 1992). The **particle transport terms** in the core and edge nodes are computed by

$$S_{\sigma,\text{tran}}^{\text{core}} = -\frac{n_{\sigma}^{\text{core}} - n_{\sigma}^{\text{edge}}}{\tau_{P,\sigma}^{\text{core} \rightarrow \text{edge}}}, \quad (7)$$

$$S_{\sigma,\text{tran}}^{\text{edge}} = \frac{V_{\text{core}}}{V_{\text{edge}}} \frac{n_{\sigma}^{\text{core}} - n_{\sigma}^{\text{edge}}}{\tau_{P,\sigma}^{\text{core} \rightarrow \text{edge}}} - \frac{n_{\sigma}^{\text{edge}}}{\tau_{P,\sigma}^{\text{edge} \rightarrow \text{sol}}}, \quad (8)$$

where  $\tau_{P,\sigma}^{\text{core} \rightarrow \text{edge}}$  (or  $\tau_{P,\sigma}^{\text{edge} \rightarrow \text{sol}}$ ) is the particle transport time from the core (or edge) node to the edge (or SOL) node. These particle transport times (Liu 2022) are calculated by

$$\tau_{P,\sigma}^{\text{core} \rightarrow \text{edge}} = \frac{r_{\text{core}}^2}{2r_{\text{core}}} \frac{\Delta r_{\text{core-edge}}}{D_{\sigma}^{\text{core}}}, \quad (9)$$

$$\tau_{P,\sigma}^{\text{edge} \rightarrow \text{sol}} = \frac{r_{\text{edge}}^2 - r_{\text{core}}^2}{2r_{\text{edge}}} \frac{\Delta r_{\text{edge-sol}}}{D_{\sigma}^{\text{edge}}}, \quad (10)$$

where  $D_{\sigma}^{\text{core}}$  and  $D_{\sigma}^{\text{edge}}$  are the core and edge particle diffusivities respectively. The **ion orbit loss (IOL) terms** are computed by  $S_{\sigma,\text{IOL}}^{\text{edge}} = -F_{\sigma,\text{orb}}^{\text{edge}} / \tau_{P,\sigma,\text{IOL}}^{\text{edge}} \cdot n_{\sigma}^{\text{edge}}$ , where  $F_{\sigma,\text{orb}}$  is the particle loss fraction (Stacey 2011; Stacey and Schumann 2015), and  $\tau_{P,\sigma,\text{IOL}}^{\text{edge}}$  is the particle IOL timescale.

**Energy Balance Equations** Energy balance equations for deuterons, tritons, alpha particles, and electrons in the core and edge nodes are

$$\frac{dU_{\sigma}^{\text{core}}}{dt} = P_{\sigma,\text{aux}}^{\text{core}} + P_{\sigma,\text{fus}}^{\text{core}} + Q_{\sigma}^{\text{core}} + P_{\sigma,\text{tran}}^{\text{core}}, \quad (11)$$

$$\frac{dU_{\sigma}^{\text{edge}}}{dt} = P_{\sigma,\text{aux}}^{\text{edge}} + P_{\sigma,\text{fus}}^{\text{edge}} + Q_{\sigma}^{\text{edge}} + P_{\sigma,\text{tran}}^{\text{edge}} + P_{\sigma,\text{IOL}}^{\text{edge}}, \quad (12)$$

$$\frac{dU_e^{\text{core}}}{dt} = P_{\Omega}^{\text{core}} + P_{e,\text{aux}}^{\text{core}} + P_{e,\text{fus}}^{\text{core}} - P_R^{\text{core}} + Q_e^{\text{core}} + P_{e,\text{tran}}^{\text{core}}, \quad (13)$$

$$\frac{dU_e^{\text{edge}}}{dt} = P_{\Omega}^{\text{edge}} + P_{e,\text{aux}}^{\text{edge}} + P_{e,\text{fus}}^{\text{edge}} - P_R^{\text{core}} + Q_e^{\text{edge}} + P_{e,\text{tran}}^{\text{edge}}, \quad (14)$$

where  $\sigma \in \{D, T, \alpha\}$ , the nodal energy density is defined as  $U_{\sigma}^{\text{node}} = 3/2 \cdot n_{\sigma}^{\text{node}} T_{\sigma}^{\text{node}}$ . The energy source and sink terms are as follows.

The **ohmic heating power** (Stacey 2012) is computed from the plasma current  $I_P$  by

$$P_{\Omega}^{\text{node}} (\text{W/m}^3) = 2.8 \times 10^{-9} \frac{Z_{\text{eff}} I_P^2}{a^4 T_e^{3/2}}, \quad (15)$$

where  $Z_{\text{eff}}$  is the effective atomic number, the plasma current  $I_P$  is in A, the minor radius  $a$  in m, and the nodal electron temperature  $T_e = T_e^{\text{node}}$  in keV. The **auxiliary heating terms** contain the neutral beam injection (NBI) and radiofrequency (RF) heating (including ion cyclotron heating (ICH), electron cyclotron heating (ECH), and lower hybrid heating (LH)):

$$P_{i,\text{aux}}^{\text{node}} = P_{i,\text{NBI}}^{\text{node}} + P_{i,\text{ICH}}^{\text{node}}, \quad (16)$$

$$P_{e,\text{aux}}^{\text{node}} = P_{e,\text{NBI}}^{\text{node}} + P_{e,\text{ECH}}^{\text{node}} + P_{\sigma,\text{LH}}^{\text{node}}. \quad (17)$$

The **fusion power** is

$$P_{\sigma,\text{fus}}^{\text{node}} = n_D^{\text{node}} n_T^{\text{node}} \langle \sigma v \rangle_{\text{fus}} U_{f\sigma}, \quad (18)$$

where  $U_{f\sigma}$  is the fusion energy transferred to the species  $\sigma$  calculated by NBI heating formulas (Wesson and Campbell 2011). The delay effect between fusion alpha heating to electrons and ions is considered by a slowing-down timescale  $\tau_{se}^{\text{node}}$  as  $\tau_{se} = (3\sqrt{2\pi} T_e^{3/2}) / (m_e^{1/2} m_b A_D)$ . The fusion reaction rate  $S_{\text{fus}}^{\text{node}}$  for core ions is evaluated at

$$T_i^{\text{node}}(t - \tau_{se}) \approx T_i^{\text{node}}(t) - \left. \frac{dT_i^{\text{node}}}{dt} \right|_t \cdot \tau_{se}, \quad (19)$$

so the fusion heating to ions is postponed by this timescale compared with electrons. **Collisional energies** transferred between ions and electrons are

$$Q_{\alpha}^{\text{node}} = Q_{\alpha D}^{\text{node}} + Q_{\alpha T}^{\text{node}} + Q_{\alpha e}^{\text{node}}, \quad (20)$$

$$Q_e^{\text{node}} = -Q_{\alpha e}^{\text{node}} - Q_{De}^{\text{node}} - Q_{Te}^{\text{node}}, \quad (21)$$

$$Q_D^{\text{node}} = -Q_{\alpha D}^{\text{node}} + Q_{DT}^{\text{node}} + Q_{De}^{\text{node}}, \quad (22)$$

$$Q_T^{\text{node}} = -Q_{\alpha T}^{\text{node}} - Q_{DT}^{\text{node}} + Q_{Te}^{\text{node}}, \quad (23)$$

where  $Q_{\sigma\sigma'}^{\text{node}}$  is the energy transferred from the species  $\sigma'$  to  $\sigma$  (Stacey 2012). The **radiation terms** from electron

cyclotron radiation (ECR) (Albajar, Johner, and Granata 2001), bremsstrahlung (Stacey 2012), and impurity radiation (Stacey 2012; Roberts 1981; Morozov, Baronova, and Senichenkov 2007) are computed by

$$P_R^{\text{node}} = P_{\text{ECR}}^{\text{node}} + P_{\text{brem}}^{\text{node}} + P_{\text{imp}}^{\text{node}}. \quad (24)$$

The **energy transport terms** in the core and edge nodes are

$$P_{\sigma, \text{tran}}^{\text{core}} = -\frac{U_{\sigma}^{\text{core}} - U_{\sigma}^{\text{edge}}}{\tau_{E, \sigma}^{\text{core} \rightarrow \text{edge}}}, \quad (25)$$

$$P_{\sigma, \text{tran}}^{\text{edge}} = \frac{V_{\text{core}}}{V_{\text{edge}}} \frac{U_{\sigma}^{\text{core}} - U_{\sigma}^{\text{edge}}}{\tau_{E, \sigma}^{\text{core} \rightarrow \text{edge}}} - \frac{U_{\sigma}^{\text{edge}}}{\tau_{E, \sigma}^{\text{edge} \rightarrow \text{sol}}}, \quad (26)$$

where  $\tau_{E, \sigma}^{\text{core} \rightarrow \text{edge}}$  (or  $\tau_{E, \sigma}^{\text{edge} \rightarrow \text{sol}}$ ) is the energy transport time from the core (or edge) node to the edge (or SOL) node. These energy transport times (Liu 2022) are solved from

$$\tau_{E, \sigma}^{\text{core} \rightarrow \text{edge}} = \frac{r_{\text{core}}^2}{2r_{\text{core}}} \frac{\Delta r_{\text{core} \rightarrow \text{edge}}}{\chi_{\sigma}^{\text{core}}}, \quad (27)$$

$$\tau_{E, \sigma}^{\text{edge} \rightarrow \text{sol}} = \frac{r_{\text{edge}}^2 - r_{\text{core}}^2}{2r_{\text{edge}}} \frac{\Delta r_{\text{edge} \rightarrow \text{sol}}}{\chi_{\sigma}^{\text{edge}}}, \quad (28)$$

where  $\chi_{\sigma}^{\text{core}}$  and  $\chi_{\sigma}^{\text{edge}}$  are the core and edge thermal diffusivities respectively. The **IOL terms** for the edge node are computed by  $P_{D, \text{IOL}}^{\text{edge}} = -E_{D, \text{orb}}^{\text{edge}} / \tau_{E, D, \text{IOL}}^{\text{edge}} \cdot U_D^{\text{edge}}$ , where  $E_{D, \text{orb}}^{\text{edge}}$  is the energy loss fraction (Stacey 2011; Stacey and Schumann 2015), and  $\tau_{E, D, \text{IOL}}^{\text{edge}}$  is the energy IOL timescale.

### 3.4 Diffusivity Models

To compute internodal transport times, it is essential to have formulas for particle and thermal diffusivities. An empirical scaling for the effective thermal diffusivity in ELMy H-mode tokamak plasmas (Becker 2004) is given by:

$$\chi_{H98}(\rho) = \alpha_H B_T^{-3.5} n_e(\rho)^{0.9} T_e(\rho) |\nabla T_e(\rho)|^{1.2} \cdot q(\rho)^{3.0} \kappa(\rho)^{-2.9} M^{-0.6} R^{0.7} a^{-0.2} \text{ (m}^2/\text{s)}, \quad (29)$$

where the thermal diffusivity  $\chi_{H98}$  in  $\text{m}^2/\text{s}$ , normalized radius  $\rho = r/a$ , coefficient  $\alpha_H = 0.123$ , toroidal magnetic field  $B_T$  in T, electron density  $n_e$  in  $10^{19} \text{ m}^{-3}$ , electron temperature  $T_e$  in keV, electron temperature gradient  $\nabla T_e$  in keV/m, safety factor  $q = q_{\psi}$ , local elongation  $\kappa$ , hydrogenic atomic mass number  $M$  in 1 amu, major radius  $R$  in m, and minor radius  $a$  in m. The particle and thermal diffusivities for electrons and ions (Becker and Kardaun 2006) can be assumed as  $\chi_e(\rho) = \chi_i(\rho) = \chi_{H98}(\rho)$  and  $D_i(\rho) = 0.6\chi_{H98}(\rho)$ . This empirical scaling was used as the baseline by Liu and Stacey (2024).

For modeling ITER plasmas accurately, we apply a parametric diffusivity formula for diffusivities (Liu and Stacey 2024):

$$\begin{aligned} \frac{\chi(\rho)}{1 \text{ m}^2/\text{s}} &= \alpha_H \left( \frac{B_T}{1 \text{ T}} \right)^{\alpha_B} \left( \frac{n_e(\rho)}{10^{19} \text{ m}^{-3}} \right)^{\alpha_n} \\ &\cdot \left( \frac{T_e(\rho)}{1 \text{ keV}} \right)^{\alpha_T} \left( \frac{|\nabla T_e(\rho)|}{1 \text{ keV/m}} \right)^{\alpha_{\nabla T}} q(\rho)^{\alpha_q} \\ &\cdot \kappa(\rho)^{\alpha_{\kappa}} \left( \frac{M}{1 \text{ amu}} \right)^{\alpha_M} \left( \frac{R}{1 \text{ m}} \right)^{\alpha_R} \left( \frac{a}{1 \text{ m}} \right)^{\alpha_a}. \end{aligned} \quad (30)$$

where  $\alpha_H, \alpha_B, \dots, \alpha_a$  are parameters to be determined. This diffusivity formula can be expressed in vector form as  $\ln \chi_{\text{node}} = \mathbf{b}_{\text{node}} + \mathbf{W}_{\text{node}} \ln \mathbf{x}_{\text{node}}$ , where  $\chi_{\text{node}}$  represents the internodal diffusivities vector,  $\mathbf{b}_{\text{node}}$  is the bias vector,  $\mathbf{W}_{\text{node}}$  is the weight matrix, and  $\mathbf{x}_{\text{node}}$  is the vector of corresponding physical values.

## 4 Computational Framework

In this study, we enhance the computational framework of NeuralPlasmaODE (Liu and Stacey 2024) to simulate burning plasma dynamics in ITER. The framework consists of several modules designed for processing experimental data, simulating plasma behavior, and optimizing model parameters. Figure 2 shows a workflow diagram of this framework.

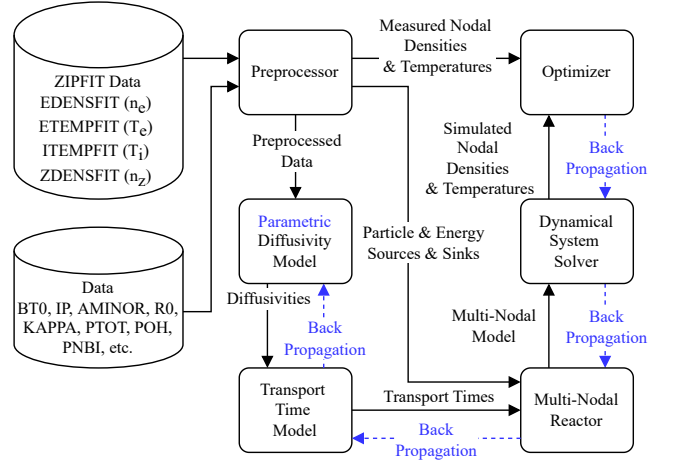


Figure 2: Computational framework of NeuralPlasmaODE, including cylinders as datasets, squares as modules, solid lines as forward flows, and dashed lines as back propagation processes.

Initially, a data module reads inputs such as two-dimensional plasma profiles and one-dimensional global parameters. A preprocessing module then standardizes these inputs into uniform time sequences and performs volume averaging on two-dimensional signals to derive nodal particle densities and temperatures. Central to the framework is a diffusivity model that calculates particle and thermal diffusivities based on experimental conditions, along with a transport time model that determines the timescales for particle and energy transport between nodes. These models feed into a reactor simulation module that integrates sources, sinks, and transport terms into a dynamical system. This system is solved using the neural ordinary differential equation (Neural ODE) (Chen et al. 2018) solver, which outputs estimated particle densities and temperatures. These estimates are then refined through an optimization module, which calculates the mean square error (MSE) between model predictions and experimental data. Using back-propagation for gradient computation and gradient descent for parameter updates, this module ensures the model parameters are optimized for accurate plasma behavior representation.

## 5 Simulations for ITER Plasmas

In this section, we present our simulations for ITER plasmas, including introducing the various simulation scenarios, discussing the simulation methodology, and analyzing the simulation results.

### 5.1 Simulation Scenarios

The multinodal burning plasma model (NeuralPlasmaODE) is applied to simulate the ITER deuterium-tritium (D-T) plasma. The ITER heating and current drive (H&CD) system (IAEA 2002; Shimada et al. 2007; Henderson et al. 2015) includes 33 MW neutral beam injection (NB), 20 MW ion cyclotron heating (IC), and 20 MW electron cyclotron heating (EC) in the initial campaign, with potential upgrades to 50 MW NB, 40 MW IC, 40 MW EC, and 40 MW lower hybrid heating (LH) in the future. The ITER tokamak can operate in both inductive and non-inductive modes. In the inductive mode, the ohmic current is the primary contributor to the total toroidal current, while in the non-inductive mode, highly energetic neutral atom injection and powerful radiofrequency radiation drive most of the toroidal current. For this study, we select inductive scenario 2, hybrid scenario 3, and non-inductive scenario 4 from the ITER design (IAEA 2002). Their typical operating conditions are shown in Table 1. The important parameters are contained, including geometries, electromagnetic values, auxiliary heating powers, densities, temperatures, and impurity fractions. Additionally, the typical radial profiles of electron and ion temperatures, and electron and helium densities are depicted in Figure 3. These profiles are integrated over the core and edge nodes to obtain nodal densities and temperatures, which are used to adjust the diffusivity parameters.

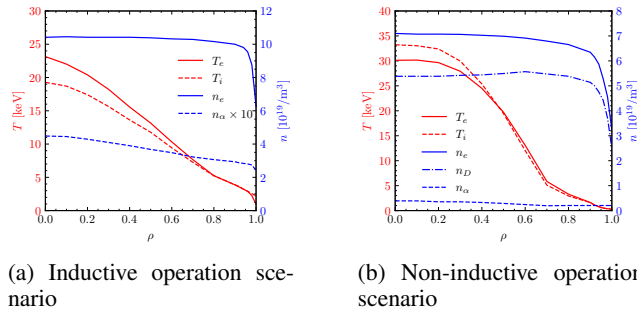


Figure 3: Typical radial profiles of plasma temperatures and densities in the ITER inductive and non-inductive operation scenarios (reproduced with permission from IAEA (2002)).

For simulating the ITER design scenarios, we model several essential particle and energy sources. For external particle sources, neutral beam injection introduces negative deuteron particles into the tokamak plasma, with an equal amount of tritons assumed to be supplied to maintain a balance between deuteron and triton particles. For external energy sources, neutral beam and radiofrequency (RF) heating systems provide auxiliary heating. In the inductively driven scenario, Wagner et al. (2010) concluded that the exact proportions of NB, IC, and EC heating are not critical, but IC

heating should be utilized to heat ions directly. Therefore, we assume all RF heating power is deposited into ions in this scenario.

### 5.2 Simulation Methodology

To accurately model the ITER burning plasma, the parameters in the diffusivity model need to be properly tuned using machine learning. In the previous study on simulating DIII-D plasmas with NeuralPlasmaODE, Liu and Stacey (2024) split experimental data into training and testing datasets. The training dataset was used to tune parameters in the diffusivity model, while the testing dataset evaluated the optimized model. However, since the ITER is still under construction and its experimental data are unavailable, we adopt transfer learning (Weiss, Khoshgoftaar, and Wang 2016), where the diffusivity parameters learned from DIII-D deuterium plasmas are transferred to ITER plasmas. We then apply a fine-tuning method to these diffusivity parameters using burning simulation results from prior research (IAEA 2002; Green et al. 2003) as optimization targets. The parameters in the diffusivity model are adjusted to match the ITER design scenarios during the current flat-top phase. Such current flat-top phase is used as the training set, while the plasma start-up phase serves as the testing set.

The optimization objective is defined as a vector including densities and temperatures in the core and edge nodes:

$$\left[ \frac{n_D^{\text{node}}}{10^{19} \text{ m}^{-3}}, \frac{n_\alpha^{\text{node}}}{10^{18} \text{ m}^{-3}}, \frac{n_e^{\text{node}}}{10^{19} \text{ m}^{-3}}, \frac{T_D^{\text{node}}}{1 \text{ keV}}, \frac{T_e^{\text{node}}}{1 \text{ keV}} \right], \quad (31)$$

with the nodal diffusivity parameters initialized to those of the DIII-D plasma. Densities and temperatures are initialized at current flat-top values instead of a cold plasma, with a time step of 0.2 s and a total simulation time of 10 s. After 14 epochs, the mean squared error (MSE) loss for scenario 2 in the current flat-top phase drops from 6.7085 to 0.0016 with a learning rate of 0.02. The optimized multinodal model is then used to simulate both inductive and hybrid scenarios. For non-inductive scenarios, diffusivity parameters are transferred from inductive scenarios and trained on scenario 4 with 2 epochs and a learning rate of 0.02, reducing the MSE loss from 13.2741 to 0.6333. After the optimization, initial temperatures are reset to 2 keV in the core node and 1 keV in the edge node for all species, with particle densities from Table 1 and Figure 3, except alpha particle densities set to  $10^{17} \text{ m}^{-3}$ . The fine-tuned multinodal model is then simulated for ITER scenarios during the plasma start-up phase over a total time of 15 s for testing.

### 5.3 Simulation Results

In this subsection, we show the results of our simulations for the different ITER operation scenarios. The outcomes for inductive, hybrid, and non-inductive scenarios are analyzed and discussed in detail.

**Inductive Operation Scenario** The inductive scenario represents a mode of operation where the majority of the plasma current is driven by ohmic heating and auxiliary heating systems. The simulation results for inductive scenario 2 are presented in Figure 4, showing the densities and

Parameter	Symbol	Inductive Scenario 2	Hybrid Scenario 3	Non-Inductive Scenario 4
Major radius	$R_0$ (m)	6.2	6.2	6.35
Minor radius	$a$ (m)	2.0	2.0	1.85
Volume	$V$ (m <sup>3</sup> )	831	831	794
Surface	$S$ (m <sup>2</sup> )	683	683	-
Elongation at the 95% flux surface	$\kappa_{95}$	1.70	1.70	1.85
Triangularity at the 95% flux surface	$\delta_{95}$	0.33	0.33	0.40
Toroidal magnetic field at the magnetic axis	$B_T$ (T)	5.3	5.3	5.18
Plasma current	$I_P$ (MA)	15	13.8	9.0
Safety factor at the 95% flux surface	$q_{95}$	3.0	3.3	5.3
Volume-averaged electron density	$\langle n_e \rangle$ (10 <sup>19</sup> m <sup>-3</sup> )	10.1	9.3	6.7
Volume-averaged ion temperature	$\langle T_i \rangle$ (keV)	8.0	8.4	12.5
Volume-averaged electron temperature	$\langle T_e \rangle$ (keV)	8.8	9.6	12.3
Fusion power	$P_{\text{fus}}$ (MW)	400	400	356
Auxiliary heating power	$P_{\text{aux}}$ (MW)	40	73	59
Radiofrequency heating power	$P_{\text{RF}}$ (MW)	7	40	-
Lower hybrid heating power	$P_{\text{LH}}$ (MW)	-	-	29
Neutral beam heating power	$P_{\text{NBI}}$ (MW)	33	33	30
Fusion energy gain factor	$Q$	10	5.4	6.0
Energy confinement time	$\tau_E$ (s)	3.7	2.73	3.1
Burn time	$t$ (s)	400	1070	3000
Helium fraction	$f_{\text{He}}$ (%)	3.2	2.5	4.1
Beryllium fraction	$f_{\text{Be}}$ (%)	2.0	2.0	2.0
Argon fraction	$f_{\text{Ar}}$ (%)	0.12	0.19	0.26
Effective impurity charge	$Z_{\text{eff}}$	1.66	1.85	2.07
Radiation power	$P_{\text{rad}}$ (MW)	47	55	37.6

Table 1: Typical parameters of inductive, hybrid, and non-inductive ITER operation scenarios (reproduced with permission from IAEA (2002)).

temperatures of deuterons, alpha particles, and electrons. Triton densities and temperatures are omitted as they are approximately the same as those of deuterons. Additionally, the temperatures of alpha particles are shown only for the cold ones, excluding fusion alpha particles at 3.5 MeV until they transfer their energy to electrons and ions. The results indicate that core and edge temperatures reach a steady state at around 11 s, and no energy excursion due to fusion alpha heating is observed.

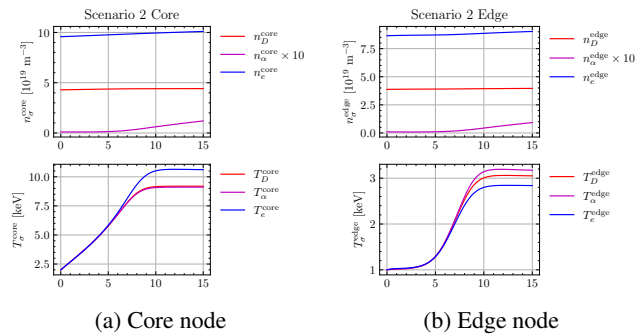


Figure 4: Densities and temperatures of ITER inductive operation scenario 2.

Figure 5 depicts the power changes over time, with posi-

tive values indicating energy gain and negative values indicating energy loss. At the start of the simulation, ohmic heating ( $P_{\text{oh}}^{\text{core}}$ ) and auxiliary heating ( $P_{\sigma,\text{aux}}^{\text{core}}$ ) provide most of the energy to core electrons and ions until the core ion temperature becomes high enough to initiate fusion reactions. Fusion alpha particles first heat the electrons ( $P_{e,\text{fus}}^{\text{core}}$ ), after which the heated electrons transfer their energy to ions through Coulomb collisions ( $Q_i^{\text{core}} = -Q_e^{\text{core}}$ ). Meanwhile, fusion heating is also deposited directly to core ions ( $P_{i,\text{fus}}^{\text{core}}$ ). The fusion heating is then removed by radiation power ( $P_{\text{rad}}^{\text{core}}$ ) and transport loss ( $P_{\sigma,\text{tran}}^{\text{core}}$ ).

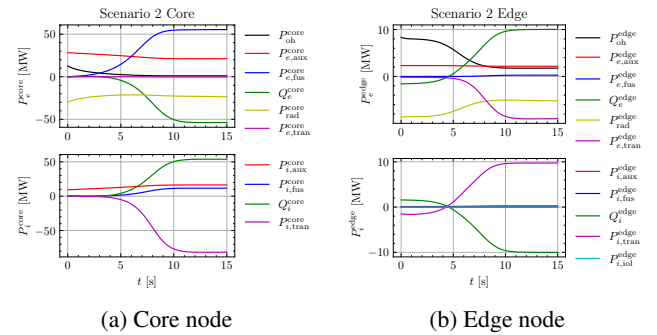


Figure 5: Powers of ITER inductive operation scenario 2.

**Hybrid Operation Scenario** A hybrid mode of operation, where a significant fraction of the plasma current is driven by non-inductive current drive power and the bootstrap current, is a promising route for establishing steady-state or non-inductive modes in ITER (IAEA 2002). To verify thermal stability with fusion alpha heating, we selected scenario 3 for simulation using the multinodal burning plasma dynamics model. The results for this scenario are shown in Figures 6 and 7, first displaying the densities and temperatures, followed by the power profiles. The ion and electron temperatures reach a steady state at around 6 s, which is shorter than in the inductive scenarios due to the higher auxiliary heating power. This increased auxiliary heating is offset by higher radiation and transport losses, preventing any power excursion in this hybrid scenario simulation.

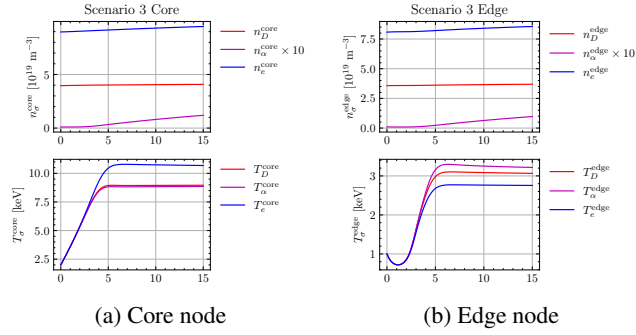


Figure 6: Densities and temperatures of ITER hybrid operation scenario 3.

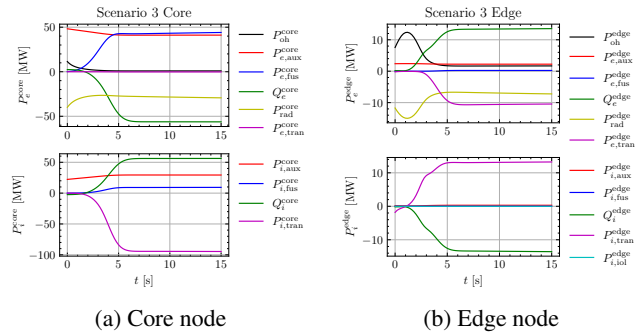


Figure 7: Powers of ITER hybrid operation scenario 3.

**Non-Inductive Operation Scenario** The non-inductive operation scenario selected for this study is weak negative shear (WNS) scenario 4. Compared to the inductive profiles, the typical radial profiles in Figure 3 show a higher core temperature with a steeper temperature gradient. The plasma current in the non-inductive scenario is lower than in the inductive one, but the safety factor at the 95% flux surface is higher. Additionally, the plasma current is roughly equally divided between current drive and bootstrap current. The simulation results for the non-inductive scenario 4 are presented in Figures 8 and 9. The ion and electron temperatures reach a steady state at around 12 s. Compared to inductive

scenario 2, non-inductive scenario 4 exhibits higher fusion and auxiliary heating. Additionally, the energy transported from core electrons to the edge is significantly greater. However, Coulomb collisional energy transfer occurs from ions to electrons since the core electron temperature is lower than the core ion temperature. No energy excursion due to fusion alpha heating is observed in this simulation.

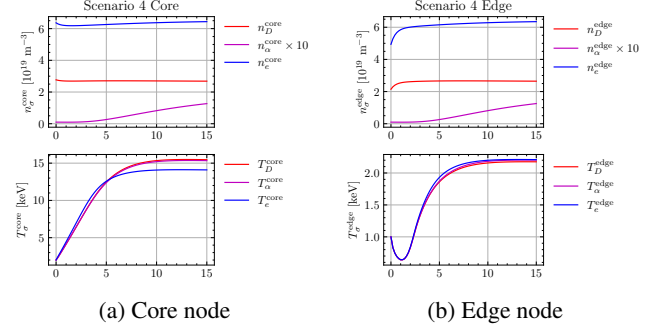


Figure 8: Densities and temperatures of ITER non-inductive operation scenario 4.

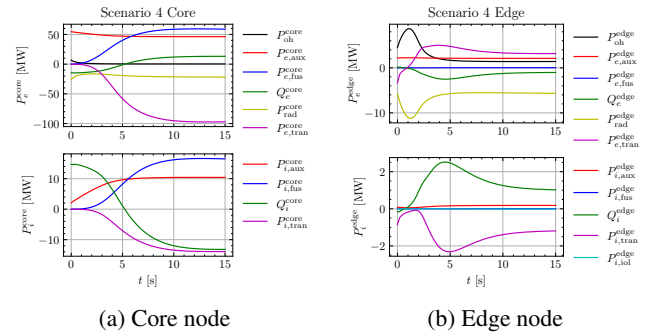


Figure 9: Powers of ITER non-inductive operation scenario 4.

## 6 Conclusion

In this research, we simulate ITER deuterium-tritium (D-T) plasmas using NeuralPlasmaODE, a multi-region multi-timescale transport model, to understand complex energy transfer processes. By employing neural ordinary differential equations (Neural ODEs) and leveraging transfer learning with parameters derived from DIII-D experimental data, we enhance simulation efficiency and accuracy. The model is applied to both inductive and non-inductive operational scenarios of ITER, showing that in inductive scenarios, core electrons dissipate energy through various radiation mechanisms, while in non-inductive scenarios, most plasma current is generated non-inductively, leading to higher core temperatures and significant energy transport to the edge. In both scenarios, radiation and transport processes effectively prevent thermal runaway instability, highlighting the potential of machine learning to advance our understanding and control of burning plasma dynamics in fusion reactors.

## References

- Albajar, F.; Bornatici, M.; and Engelmann, F. 2009. RAYTEC: a new code for electron cyclotron radiative transport modelling of fusion plasmas. *Nuclear Fusion*, 49(11): 115017.
- Albajar, F.; Johner, J.; and Granata, G. 2001. Improved calculation of synchrotron radiation losses in realistic tokamak plasmas. *Nuclear fusion*, 41(6): 665.
- Becker, G. 2004. Study of anomalous inward drift in tokamaks by transport analysis and simulations. *Nuclear fusion*, 44(9): 933.
- Becker, G.; and Kardaun, O. 2006. Anomalous particle pinch and scaling of  $\nu_{in}/D$  based on transport analysis and multiple regression. *Nuclear fusion*, 47(1): 33.
- Bosch, H.-S.; and Hale, G. M. 1992. Improved formulas for fusion cross-sections and thermal reactivities. *Nuclear fusion*, 32(4): 611.
- Chen, R. T.; Rubanova, Y.; Bettencourt, J.; and Duvenaud, D. K. 2018. Neural ordinary differential equations. *Advances in neural information processing systems*, 31.
- Cordey, J.; Thomsen, K.; Chudnovskiy, A.; Kardaun, O.; Takizuka, T.; Snipes, J.; Greenwald, M.; Sugiyama, L.; Rytter, F.; Kus, A.; et al. 2005. Scaling of the energy confinement time with  $\beta$  and collisionality approaching ITER conditions. *Nuclear fusion*, 45(9): 1078.
- Green, B.; et al. 2003. ITER: burning plasma physics experiment. *Plasma physics and controlled fusion*, 45(5): 687.
- Henderson, M.; Saibene, G.; Darbos, C.; Farina, D.; Figini, L.; Gagliardi, M.; Gandini, F.; Gassmann, T.; Hanson, G.; Loarte, A.; et al. 2015. The targeted heating and current drive applications for the ITER electron cyclotron system. *Physics of plasmas*, 22(2).
- Hill, M. D. 2019. *Burn Control Mechanisms in Tokamak Fusion Reactors*. Ph.D. thesis, Georgia Institute of Technology.
- Hill, M. D.; and Stacey, W. M. 2017. Confinement Tuning of a 0-D Plasma Dynamics Model. *Fusion Science and Technology*, 72(2): 162–175.
- IAEA. 2002. *ITER Technical Basis*. Number 24 in ITER EDA Documentation Series. Vienna: International Atomic Energy Agency.
- Kessel, C.; Giruzzi, G.; Sips, A.; Budny, R.; Artaud, J.; Basiuk, V.; Imbeaux, F.; Joffrin, E.; Schneider, M.; Murakami, M.; et al. 2007. Simulation of the hybrid and steady state advanced operating modes in ITER. *Nuclear Fusion*, 47(9): 1274.
- Liu, Z. 2022. *A Multi-Region Multi-Timescale Burning Plasma Dynamics Model for Tokamaks*. Ph.D. thesis, Georgia Institute of Technology.
- Liu, Z.; and Stacey, W. 2021. A Multi-Region Multi-Timescale Burning Plasma Dynamics Model for Tokamaks. In *APS Division of Plasma Physics Meeting Abstracts*, volume 2021, UP11–087.
- Liu, Z.; and Stacey, W. M. 2024. Application of Neural Ordinary Differential Equations for Tokamak Plasma Dynamics Analysis. *arXiv preprint arXiv:2403.01635*.
- Morozov, D. K.; Baronova, E.; and Senichenkov, I. Y. 2007. Impurity radiation from a tokamak plasma. *Plasma Physics Reports*, 33: 906–922.
- Roberts, D. 1981. Total impurity radiation power losses from steady-state tokamak plasmas. *Nuclear Fusion*, 21(2): 215.
- Shimada, M.; Campbell, D.; Mukhovatov, V.; Fujiwara, M.; Kirneva, N.; Lackner, K.; Nagami, M.; Pustovitov, V.; Uckan, N.; Wesley, J.; et al. 2007. Overview and summary. *Nuclear Fusion*, 47(6): S1.
- Stacey, W. M. 2007. A survey of thermal instabilities in tokamak plasmas: Theory, comparison with experiment, and predictions for future devices. *Fusion science and technology*, 52(1): 29–67.
- Stacey, W. M. 2011. The effect of ion orbit loss and X-loss on the interpretation of ion energy and particle transport in the DIII-D edge plasma. *Physics of Plasmas*, 18(10).
- Stacey, W. M. 2012. *Fusion plasma physics*. John Wiley & Sons.
- Stacey, W. M. 2013. Effect of ion orbit loss on distribution of particle, energy and momentum sources into the tokamak scrape-off layer. *Nuclear Fusion*, 53(6): 063011.
- Stacey, W. M. 2021. A Nodal Model for Tokamak Burning Plasma Space-Time Dynamics. *Fusion Science and Technology*, 77(2): 109–118.
- Stacey, W. M.; and Schumann, M. T. 2015. The distribution of ion orbit loss fluxes of ions and energy from the plasma edge across the last closed flux surface into the scrape-off layer. *Physics of Plasmas*, 22(4).
- Wagner, F.; Becoulet, A.; Budny, R.; Erckmann, V.; Farina, D.; Giruzzi, G.; Kamada, Y.; Kaye, A.; Koechl, F.; Lackner, K.; et al. 2010. On the heating mix of ITER. *Plasma Physics and Controlled Fusion*, 52(12): 124044.
- Wang, J.-F.; Amano, T.; Ogawa, Y.; and Inoue, N. 1997. Simulation of burning plasma dynamics in ITER. *Fusion technology*, 32(4): 590–600.
- Weiss, K.; Khoshgoftaar, T. M.; and Wang, D. 2016. A survey of transfer learning. *Journal of Big data*, 3: 1–40.
- Wesson, J.; and Campbell, D. J. 2011. *Tokamaks*, volume 149. Oxford university press.
- Wilks, T.; and Stacey, W. 2016. Improvements to an ion orbit loss calculation in the tokamak edge. *Physics of Plasmas*, 23(12).