Impurity transport analysis using the EMC3-EIRENE code by taking into account plasma edge measurements on W7-X with an island divertor configuration

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Yu Luo
aus Nanchang

Düsseldorf, September 2025

aus dem Institut für
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Yunfeng Liang

2. Prof. Dr. Georg Pretzler

Tag der mündlichen Prüfung:

Zusammenfassung

Wendelstein 7-X (W7-X) soll den Dauerbetrieb eines Stellarators mit 10 MW für 30 Minuten im Inseldivertor-Konzept demonstrieren. Um das Risiko einer Beschädigung plasmaexponierter Komponenten zu verringern, ist eine strikte Kontrolle des Wärme- und Teilchenausstoßes erforderlich. Frühere Untersuchungen haben gezeigt, dass die Energiedissipation über Strahlung von Verunreinigungen eine wirksame Methode zur Minderung der Divertor-Wärmelasten darstellt. Allerdings erschweren die komplexe dreidimensionale Magnetfeldtopologie von W7-X und der begrenzte diagnostische Zugang die Analyse des Verunreinigungstransports erheblich. Als dreidimensionaler Randtransportcode ist daher EMC3-EIRENE zu einem wichtigen Werkzeug für quantitative Verunreinigungstransportstudien an W7-X geworden.

Um das Verhalten von Verunreinigungen im Inseldivertor besser zu analysieren, wurde in der Nachverarbeitung von EMC3-EIRENE ein synthetisches Diagnostikverfahren für das Divertor-Spektroskopiesystem entwickelt. Da die Spektrallinienstrahlung stark von den lokalen Plasmabedingungen abhängt, ermöglicht die synthetische Rekonstruktion eine engere Einschränkung der Eingabeparameter von EMC3-EIRENE, insbesondere der Koeffizienten des senkrechten Transports. Systematische Parameterscans deuten darauf hin, dass eine geeignete Abstimmung der Eingabeparameter zu EMC3-EIRENE-Simulationen führt, die besser mit den experimentellen Ergebnissen übereinstimmen. Im Vergleich zu räumlich homogenen Senkrechttransportkoeffizienten verbessern nicht-uniforme Koeffizientenverteilungen die Übereinstimmung zwischen Simulationen und Messungen zusätzlich. Durch die Kopplung des HINT-Gleichgewichtslösers an EMC3-EIRENE können der Einfluss von Gleichgewichtseffekten (Plasmabeta und toroidaler Plasmastrom) auf die Magnetfeldtopologie sowie die Auswirkungen der daraus resultierenden Topologieänderungen auf den Verunreinigungstransport bewertet werden. Mit toroidalem Strom wird die Restinsel nach innen gezogen, während die Verunreinigungsverteilung nahezu unverändert bleibt. Bei einem endlichen Plasmabeta von 2 % auf der Achse nimmt die Stochastisierung der Feldlinien zu und die Verbindungslänge in offenen Feldregionen verkürzt sich, wodurch die Impuritätsquelle über einen veränderten Divertor-Teilchenfluss verändert wird und die Transporteigenschaften modifiziert werden.

Die manuelle Anpassung der Eingaben von EMC3-EIRENE erfordert wiederholte Gesamtsimulationen mit variierenden Eingaben, was sowohl arbeitsintensiv als auch rechenaufwendig ist. Um dieses Problem zu lösen, wurde ein zweistufiges neuronales Netz als Surrogatmodell entwickelt, das diagnostische Messungen schnell auf EMC3-EIRENE-konsistente Randplasmavorhersagen abbildet. In der ersten Stufe reproduziert ein Feedforward-Neuronales Netz die synthetischen Diagnosesignale, die EMC3-EIRENE für einen gegebenen Parametersatz erzeugen würde, und liefert innerhalb eines Bayes'schen Rahmens posteriori Schätzungen der Eingabeparameter von EMC3-EIRENE, die mit den Messungen konsistent sind. In der zweiten Stufe werden diese abgeleiteten Eingaben auf vollständige

Plasmaparameterverteilungen abgebildet: Ein Autoencoder wird auf vollständigen Simulationsergebnissen trainiert, um einen kompakten latenten Raum zu definieren; anschließend wird ein Mapping-Netzwerk trainiert, das den latenten Raum aus den bayesianisch abgeleiteten Eingaben vorhersagt; zuletzt werden das Mapping-Netzwerk und der Decoder gemeinsam feinabgestimmt, um die vollständigen Randplasmadaten zu rekonstruieren. Zusammen ermöglichen diese Schritte EMC3-EIRENE-konsistente Vorhersagen direkt aus Diagnosen und eröffnen so großangelegte Parameterscans, eine schnelle Szenarioplanung sowie eine Inter-Shot-Analyse im Kontrollraum während des W7-X-Betriebs.

Abstract

Wendelstein 7-X (W7-X) aims to demonstrate steady-state stellarator operation at 10 MW for 30 minutes using an island divertor configuration. To reduce the risk of damaging plasma-facing components, rigorous control of heat and particle exhaust is required. Previous research has shown that energy dissipation via impurity radiation is an effective way to mitigate divertor heat loads. However, W7-X's complex three-dimensional magnetic topology and limited diagnostic access make impurity transport analysis particularly challenging. As a three-dimensional edge transport code, EMC3-EIRENE has therefore become an important tool for quantitative impurity transport studies on W7-X.

To better analyze impurity transport behavior in the island divertor, a synthetic diagnostic for the divertor spectroscopy system was developed in EMC3-EIRENE post-processing. Because line radiation is highly sensitive to local plasma conditions, the synthetic reconstruction allows tighter constraints on EMC3-EIRENE input parameters, particularly the cross-field transport coefficients. Systematic parameter scans indicate that appropriate tuning of the input parameters yields EMC3-EIRENE simulations that more closely match the experimental results. Compared with spatially uniform cross-field coefficients, non-uniform settings further improve agreement between simulations and measurements. By coupling the HINT equilibrium solver to EMC3-EIRENE, the influence of equilibrium effects (plasma beta and toroidal plasma current) on magnetic topology, and the impact of the resulting topology changes on impurity transport, could be evaluated. With toroidal current, the remnant island is pulled inward while the impurity distribution remains almost unchanged. With a finite on-axis plasma beta of 2 %, field-line stochastization increases and the connection length in open-field regions shortens, thereby altering the impurity source via a changed divertor particle flux and modifying transport property.

Manual tuning of EMC3-EIRENE inputs involves iterating full simulations over varying inputs, which is laborious and computationally expensive. To address this problem, a two-stage neural-network surrogate was built to rapidly map diagnostic measurements to EMC3-EIRENE-consistent edge-plasma predictions. In the first stage, a feedforward neural network reproduces the synthetic diagnostic signals that EMC3-EIRENE would generate for a given input set and, within a Bayesian framework, returns posterior estimates of EMC3-EIRENE input parameters consistent with the measurements. In the second stage, those inferred inputs are mapped to full plasma parameter distributions: an autoencoder is trained on full simulation outputs to define a compact latent space; a mapping network is then trained to predict the latent space from the Bayesian inferred inputs; lastly, the mapping network and decoder are fine-tuned jointly to reconstruct the full edge-plasma parameters. Together, these steps provide EMC3-EIRENE-consistent predictions directly from diagnostics, enabling large-scale parameter scans, rapid scenario planning, and inter-shot analysis in the control room during W7-X operation.

Eidesstattliche Versicherung

Ich versichere an Eides statt, dass die vorliegende Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Ort, Datum: ________________ Unterschrift: ________________

Publications

As first author

- Subsections 3.2.2, 5.2.2, 5.3.1, 5.3.2 and 5.3.3, as well as figures 5.4, 5.5, 5.7, 5.8, 5.10, 5.11, 5.12, 5.13, and 5.14, have been previously published in [1]. The author of this thesis carried out the data selection and analysis, as well as the implementation and evaluation of the Bayesian neural network presented in the publication.
- Section 6.1 as well as figures 6.1 to 6.19 have been submitted to Nuclear Fusion under the title "Neural network-based surrogate model for 3D edge-plasma transport in the standard configuration of W7-X".

As co-author

- The author of this thesis also contributed as a co-author to publications [2, 3, 4, 5, 6]. Specific contributions include processing portions of the diagnostic datasets for [2, 3, 5], contributing to the design of the neural-network architecture for [4], and participating in the spectroscopic diagnostics operation during OP2.1 [6].

[1] Y. Luo, et al., "A neural network-based method for input parameter optimization of edge transport modeling utilizing experimental diagnostics", Nuclear Fusion 65, 096016 (2025). DOI: 10.1088/1741-4326/adf75f.

[2] S. Xu, et al., "Modeling of plasma beta effects on the island divertor transport in the standard configuration of W7-X", Nuclear Fusion 63, 066005 (2023). DOI: 10.1088/1741-4326/acc7b8.

[3] J. Yang, et al., "Dynamic characteristics of edge plasma profiles during the opening of edge islands on the J-TEXT tokamak", Nuclear Fusion 64, 056030 (2024). DOI: 10.1088/1741-4326/ad37cc.

[4] J. Q. Cai, et al., "Improved training framework in a neural network model for disruption prediction and its application on EXL-50", Plasma Science & Technology 26, 055102 (2024). DOI: 10.1088/2058-6272/ad1571.

[5] J. Yang et al., "First observation of beta-induced Alfvén eigenmode inside the edge magnetic island on the J-TEXT tokamak", Nuclear Fusion 64, 024001 (2024). DOI: 10.1088/1741-4326/ad1788.

[6] O. Grulke et al., "Overview of the first Wendelstein 7-X long pulse campaign with fully water-cooled plasma facing components", Nuclear Fusion 64, 112002 (2024). DOI: 10.1088/1741-4326/ad2f4d.

List of Figures

List of Tables

Nomenclature

β_{ax} on-axis plasma beta

$\chi_{\perp}$ cross-field energy transport coefficient

$\nabla_{\parallel}$ gradient along and across the magnetic field

$\nabla_{\perp}$ gradient across the magnetic field

$D_{\perp}$ cross-field particle transport coefficient

ι rotational transform

λ_q energy decay length

L_c connection length

L_{loss} loss function

$L_{radiance}$ line-emission radiance

$\langle\cdot\rangle^n$ n-th moment

$\boldsymbol{\pi_a}$ viscous stress tensor

q safety factor

ASHA Asynchronous Successive Halving Algorithm

BOHB Bayesian Optimization with Hyperband

CNN convolutional neural network

FNN feed-forward neural network

LCFS last closed flux surface

LOS line of sight

MAP maximum a posteriori

MHD magnetohydrodynamics

MLP multilayer perceptron

Contents

1 Introduction

Humankind has long sought an inexhaustible energy source, and nuclear fusion may provide the solution. The most practical reaction on Earth is deuterium-tritium fusion, which releases 17.6 MeV per event. Deuterium can be extracted from seawater, while tritium is produced by bombarding lithium with neutrons. In recent years, fusion's promising potential has attracted the attention of governments and private investors worldwide, with hopes of achieving commercial fusion reactors within this century.

1.1 Global energy challenges and the promise of fusion

Energy and environmental issues are unavoidable in today's society. At present, fossil fuels [1] are still the primary source of energy for human production and daily life. However, due to the non-renewable nature of fossil fuels and the large amount of pollutants [2] they produce during use, this forces us to develop efficient, clean and renewable new energy sources, such as nuclear energy, wind energy, hydropower, tidal energy, geothermal energy, and hydrogen energy. Related studies [3] indicate that, in terms of power generation capacity, nuclear power performs the best, while solar photovoltaic energy performs the lowest. Current nuclear power stations utilize fission energy. During the fission reaction process, a large amount of highly radioactive and long-lived nuclear waste is generated, posing a challenging issue for their disposal. More importantly, nuclear power stations carry safety risks, and in the event of an accident, they can lead to catastrophic consequences for local populations and the ecological environment. The Chernobyl nuclear accident [4] in the former Soviet Union in 1986 and the Fukushima nuclear accident [5] in Japan in 2011 serve as stark reminders that the cautious use of fission nuclear power stations is imperative.

Another form of nuclear energy is nuclear fusion, where two or more atomic nuclei (lighter than iron Fe) combine to form one or more different atomic nuclei, releasing tremendous energy due to mass loss. As shown in Figure 1.1, the binding energy per nucleon of elements with lighter mass numbers is smaller than that of elements with intermediate mass numbers. Hence, nuclear fusion can release energy. Compared to nuclear fission, nuclear fusion has three main advantages. Firstly, fuel sources are abundant, as deuterium can be extracted from seawater and tritium can be obtained through lithium-neutron reactions. Secondly, fusion reaction conditions are stringent, making operation safer. Thirdly, fusion produces clean reaction products without generating difficult-to-handle waste.

The nuclear force between nucleons is short-range, exerting significant influence only when nucleons are within distances on the order of 10^{-15} m, where it is great than the Coulomb force.

Consequently, in order to achieve a fusion reaction, the reacting particles need sufficient kinetic energy to overcome the Coulomb potential barrier. Figure 1.2 shows some important fusion reactions, it indicate that the most feasible fusion reaction on earth is the *Deuterium-Tritium* reaction, with its reaction formula is: ${}_1^2\mathrm{D} + {}_1^3\mathrm{T} \rightarrow {}_2^4\mathrm{He} + \mathrm{n} + 17.6\ \mathrm{MeV}$. *D-T* reaction has highest cross-section at low temperature, compared to other reactions. The cross-section refers to the likelihood of a particular nuclear reaction occurring between two particles when they collide. As evident, the temperature for *D-T* reaction is at least 10 keV.

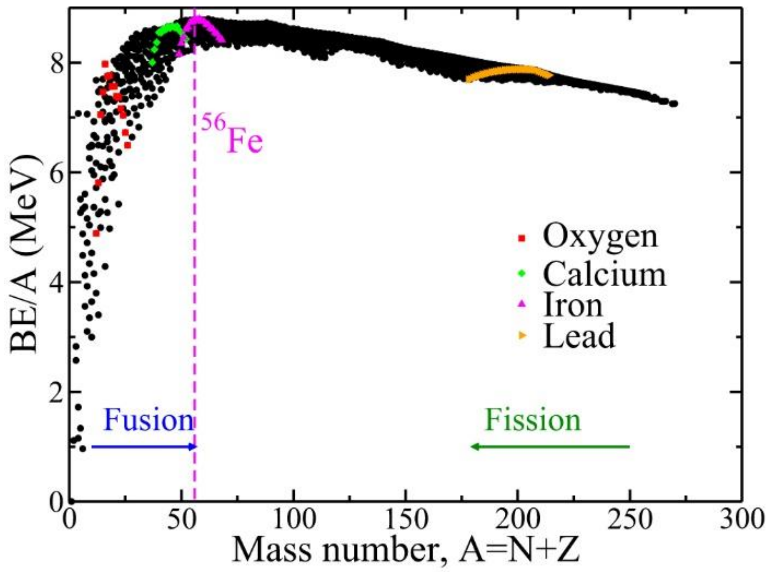

Figure 1.1: Binding energy per nucleon. Image from [6].

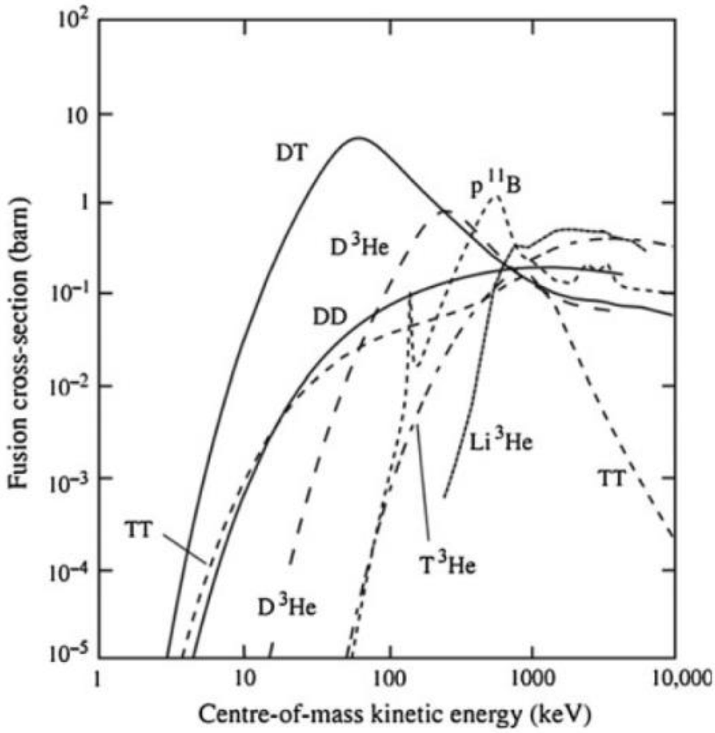

Figure 1.2: Cross-sections of different fusion reactions as a function of the reacting particle energy. Image from [7].

Another consideration for nuclear fusion is the energy balance. For a nuclear fusion reactor to be economically viable, it must produce net energy. In the 1997 paper [8],John D. Lawson presents calculations regarding the power balance in thermonuclear reactors operating under various idealized conditions. The terms involved in energy balance include thermonuclear power, energy loss and auxiliary heating. For the thermonuclear power in a *D-T* reaction is

$$P_{DT} = n_D n_T < \sigma v > E \qquad 1.1.1$$

where n_D and n_T are the densities of deuterium and tritium, $< \sigma v >$ is the rate coefficient of reaction and E is the fusion energy. If $n_D = n_T = \frac{1}{2}n$,

$$P_{DT} = \frac{1}{4}n^2 < \sigma v > E \qquad 1.1.2$$

For a hot plasma, energy loss includes radiation and conduction. Radiation consist of bremsstrahlung, impurities line emission, cyclotron radiation and so on. Conduction loss arises due to the huge gradient of edge plasma temperature. In a steady-state tokamak, energy loss is balanced by auxiliary heating, expressed as $P_L = P_H$. The factor energy confinement time is defined as:

$$\tau_E = \frac{W}{P_H} = \frac{3\bar{n}\bar{T}V}{P_H} \qquad 1.1.3$$

where W is total energy of plasma and V is the plasma volume.

Now, α-particle heating in *D-T* fusion is only considered, as neutrons readily escape the plasma without interaction. For energy balance, it is expressed as:

$$P_H + P_\alpha = P_L \qquad 1.1.4$$

Using equations 1.1.2 and 1.1.3, this balance can be written

$$P_H + \frac{1}{4}\overline{n^2 < \sigma v >} E_\alpha = \frac{3\bar{n}\bar{T}V}{\tau_E} \qquad 1.1.5$$

If the plasma can be self-sustaining without auxiliary heating,

$$\frac{1}{4}\overline{n^2 < \sigma v >} E_\alpha \geq \frac{3\bar{n}\bar{T}V}{\tau_E} \Longrightarrow n\tau_E \geq \frac{12\bar{T}}{<\sigma v> E_\alpha} \qquad 1.1.6$$

This is the ignition condition for fusion reaction. This product has the minimum at $T \approx 30$ keV and the requirement is

$$n\tau_E \geq 1.5 \times 10^{20}\ m^{-3}s \qquad 1.1.7$$

Another more useful factor for measuring fusion progress is triple product of density, temperature, and confinement time. Compared to $n\tau_E$, the triple product has its minimum at $T \approx 14$ keV. The minimum triple product for parabolic density and temperature profiles is

$$nT\tau_E \geq 5 \times 10^{21} m^{-3} keVs \qquad 1.1.8$$

Currently, human cannot reach ignition conditions, but the fusion reaction can be achieved with auxiliary heating power. A gain factor Q in magnetic confinement fusion is defined as:

$$Q = \frac{P_{fusion}}{P_{heating}} \qquad 1.1.9$$

where P_{fusion} is fusion power and $P_{heating}$ is auxiliary heating power. $Q = 1$ represents that the fusion power is equal to external heating power. For ignition condition, the $P_{heating}$ is zero, so Q is ∞. To date, the record plasma gain in magnetic confinement fusion is $Q = 0.67$ at JET tokamak [9]. For the future device ITER (International Thermonuclear Experimental Reactor) tokamak, it is designed to achieve the gain factor >10 [10], indicating that it can produce ten time fusion power than auxiliary heating power.

1.2 Magnetic confinement fusion

No material can withstand the high temperature required for fusion reactions, so it can't use simple containers to confine the fuel particles. For instance, the sun utilizes gravitational forces to confine plasma and sustain fusion reactions. At present, two main approaches are currently considered potentially achievable for controlled fusion: inertial confinement fusion and magnetic confinement fusion. Inertial confinement fusion is to use the shockwave compression from high-power lasers to rapidly compress pellets containing deuterium-tritium fuel to achieve extremely high temperatures and densities, allowing the fuel within the pellets to meet the Lawson criterion. For magnetic confinement fusion, it utilizes strong magnetic field to trap charged particles and make particles away from material walls. Compared to Inertial confinement fusion, magnetic confinement fusion typically has longer confinement time but lower particle densities.

In order to confine charged particles, a straightforward magnetic configuration is the toroidal closed magnetic field. However, this simple magnetic configuration can't trap particle effectively due to the non-uniform magnetic field [11], which can cause particles to drift perpendicular to magnetic field. There are given by

$$\boldsymbol{v}_{\nabla B} + \boldsymbol{v}_{cur} = m\left(v_{\parallel}^2 + \frac{1}{2}v_{\perp}^2\right)\frac{\boldsymbol{B}\times\nabla B}{B^3} \quad 1.2.1$$

The gradient and the curvature of magnetic field can induce the opposite drifts for ion and electron, leading to charge separation. This electric field, in turn, induces further drift motion, unrelated to the charge of the charged particles, causing the plasma to collectively drift outward rapidly and undergo significant loss. The electric drift velocity is written

$$\boldsymbol{v}_{\boldsymbol{E}\times\boldsymbol{B}} = \frac{\boldsymbol{E}\times\boldsymbol{B}}{B^2} \quad 1.2.2$$

To prevent plasma loss, it needs to avoid the charge separation. Introducing poloidal field components into the toroidal magnetic field enables the rapid neutralization of the space electric field by charged particles moving along magnetic field lines. In the past decades, scientists have tried many magnetic configurations, such as magnetic mirrors, reversed-field pinches, stellarators, tokamaks, and others. At present, the magnetic confinement devices that countries are interested in are mainly tokamak and stellarator. Both tokamaks and stellarators use helical magnetic fields to confine plasma and avoid drift losses. In tokamaks, the helical magnetic field consists of the toroidal field from external coils and the poloidal field generated by plasma current in the vacuum chamber. While for stellarators, the twisting magnetic field is entirely provided by external coils. For helical magnetic field, the degree of distortion can be described by *pitch angle* $\theta = arctan\left(\frac{B_{pol}}{B_{tor}}\right) \approx \frac{B_{pol}}{B_{tor}}$. It is more commonly characterized by the *safety factor* $q = \frac{d\phi}{d\varphi}$, where ϕ is the toroidal magnetic flux and φ is the poloidal magnetic flux. The physical meaning of q is the number of turns the magnetic field lines travel in the toroidal direction for each turn in the poloidal direction. It usually uses the *rotational transform* $\iota = q^{-1}$ in stellarators. Figure 1.3 visually illustrates the differences between tokamaks and stellarators. Although stellarators face several drawbacks [12], the three-dimensional magnetic configuration requires highly complex superconducting coils and complicates blanket integration and maintenance; the island-divertor geometry is sensitive to plasma current, making precise strike-line control more difficult; and unoptimized machines exhibited strong neoclassical cross-field transport with energy confinement often below that of advanced tokamaks. Nevertheless, their advantages are compelling: they operate without a sustained toroidal plasma current, thereby avoiding current-driven MHD instabilities and many associated operational limits; they can access densities set by the Sudo or radiation limits [13] rather than the Greenwald limit [14]; and their island or non-resonant divertors provide longer connection lengths, enhanced the cross-field transport, and larger wetted areas that improve power exhaust.

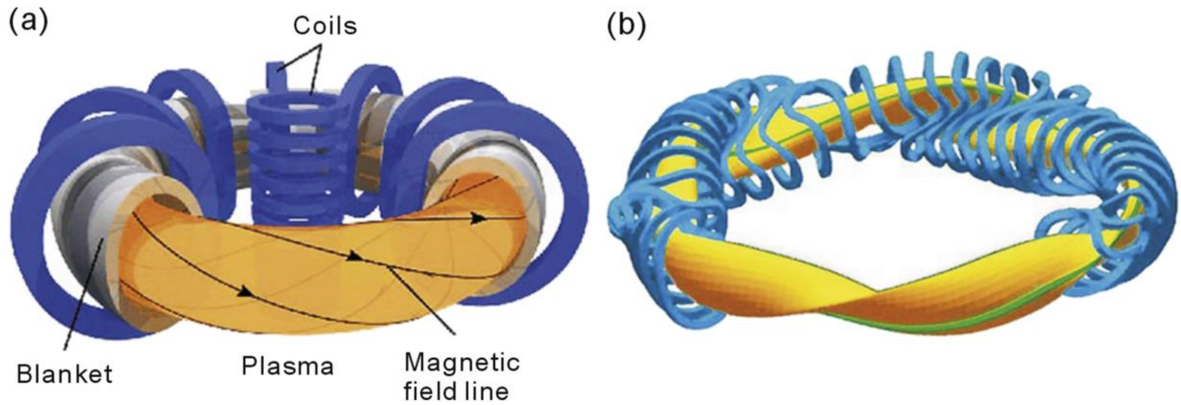

Figure 1.3: The magnetic configuration for (a) tokamak, (b) stellarator. Image from [15].

1.3 The Wendelstein 7-X stellarator

In 1951, American physicist Lyman Spitzer first proposed the stellarator concept with a helical magnetic field completely generated by the external coils at Princeton University [16]. In order to achieve magnetic rotational transform, he presented two types [17]. The first simplest configuration is known as the figure-8 device, as shown in Figure 1.4. Field coils encircle the figure-eight torus, producing rotational transform. Another method to induce rotational transform is by using the continuous helical coils wound along the magnetic axis. As depicted in Figure 1.5, the toroidal field is generated by toroidal field coils, and the poloidal field is provided by helical windings. However, classic stellarator faces many challenges. First, the helical windings are located in the strong magnetic field generated by the toroidal-field coils and are subject to strong force. Second, the complex coils arrangement makes installation and maintenance difficult.

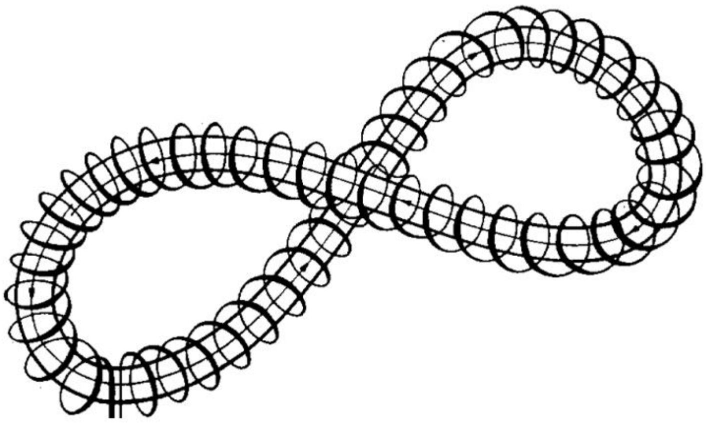

Figure 1.4: Figure-eight stellarator. Image from [18].

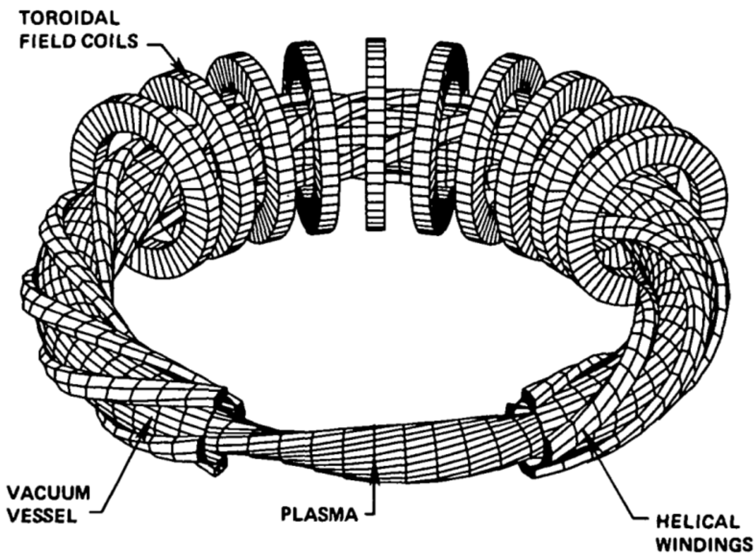

Figure 1.5: classical stellarator with continuous helical coils. Image from [19].

With the improvement of computer computing power and the development of processing technology, many new stellarator designs have emerged. According to the coil designs, there are three types: continuous helical coils, helical-axis stellarators and modular coils. For continuous helical coils, a typical device is the helitron Large Helical Device (LHD) [20]. Unlike classical stellarators, the Large Helical Device (LHD) does not utilize toroidal coils. Instead, it employs adjacent helical coils with currents flowing in the same direction. As for helical-axis stellarators, TJ-II is a representative device [21], which mainly consists of 32 toroidal coils, a central circular coil and a helical coil. Its rotational transform can be widely adjusted. For modular coils, Wendelstein 7-X, currently the world's largest stellarator, features five-fold symmetry modules. The advantages of modular coils are obvious in terms of disassembly and maintenance.

1.3.1 Magnetic configurations on W7-X

This thesis research focuses on W7-X, therefore, a detailed introduction to W7-X is provided. As shown in Figure 1.6, the main magnetic field of W7-X is produced by 50 non-planar coils and 20 planar coils. Additionally, there are 5 trim coils to reduce error fields and 10 control coils to adjust the position and size of magnetic islands. Figure 1.6 (b) gives the coils system in one period. By adjusting the currents in the planar and non-planar coils, different magnetic configurations can be achieved in W7-X. When equal currents are passed through only the 5 non-planar coils, the standard configuration in W7-X with the 5/5 magnetic island is obtained. The poloidal mode number of the magnetic islands can be adjusted via changing the edge iota value. Supplying equal current to two planar coils in the same direction as the non-planar coils effectively increases the toroidal field component, thereby reducing the iota value and achieving a low-iota configuration with 5/6 magnetic island. Conversely, supplying equal current to the planar coils in the opposite direction to the non-planar coils adds a toroidal field component in the opposite direction, increasing the iota value and resulting in a high-iota configuration with 5/4 magnetic island.

In W7-X, the magnetic field within one period is somewhat similar to a magnetic mirror, being weaker in the middle (φ = 36°) and stronger at the ends (φ = 0° or 72°). If the non-planar coils 1-5 are supplied with progressively increasing currents ($I_1 < I_2 < I_3 < I_4 < I_5$), a low magnetic mirror ratio configuration (low mirror configuration) can be achieved. Conversely, supplying the non-planar coils with progressively decreasing currents ($I_1 > I_2 > I_3 > I_4 > I_5$) will create a high magnetic mirror ratio configuration. Additionally, when the planar coils are supplied with reverse currents, the toroidal field components generated by the two planar coils will cancel each other out, leaving a vertical field component. If the current in planar coil A is in the same direction as the non-planar coils, it creates a vertical field that shifts all magnetic surfaces inward (inward shifted configuration). If the current in planar coil A is in the opposite direction, it creates a vertical field that shifts all magnetic surfaces outward (outward shifted configuration).

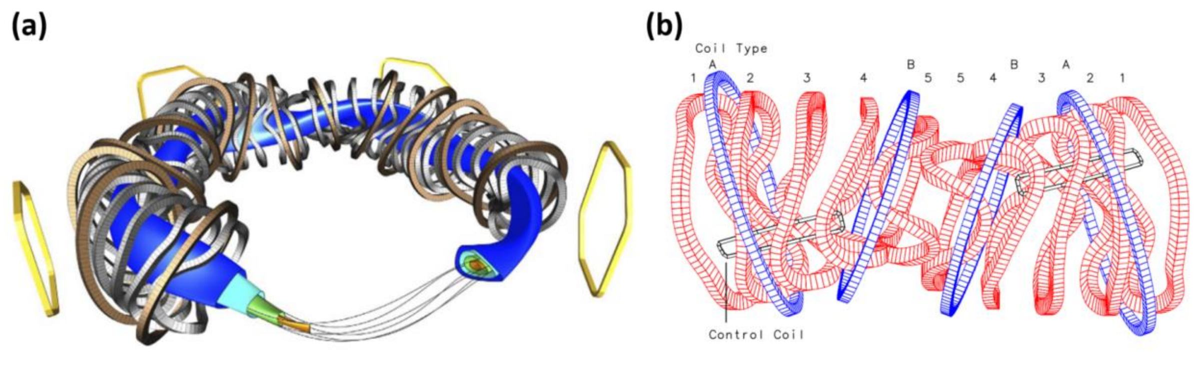

Figure 1.6: (a) Layout of W7-X, (b) coil system in one period, non-planar coils in red, planar coils in blue. Subfigure (a) taken from [22], subfigure (b) taken from [23].

1.3.2 Island divertor concept

Due to the cross-field transport of plasma from the core to the edge, the interaction of the plasma with plasma-facing component is inevitable. This brings about many challenges, such as particle and energy fluxes exhaust, material erosion and redeposition, and plasma contamination. To address these problems, magnetic confinement devices adopt limiter or divertor configurations. A limiter is a component that directly intersects with the closed magnetic field to prevent the plasma from making contact with the device walls. The advantages of limiter are its simple design and ease of implementation. However, since the limiter is directly inserted into the plasma, it is subjected to significant particle and heat fluxes, and it has limited capability for impurity screening. As for the divertor configuration, it typically uses external currents to produce a magnetic separatrix, which is known as the last closed flux surface (LCFS). Beyond the separatrix, charged particles in the open magnetic field lines are guided towards designated divertor plates. In Figure 1.7, basic configurations of a limiter and a divertor are illustrated. In both limiter and divertor configurations, there is a region with open magnetic field lines called the scrape-off layer (SOL) and the last closed flux surface (LCFS). For the divertor configuration, there is also a private flux region and an X-point where the poloidal magnetic field is zero. In the scrape-off layer, an important parameter that characterizes the target-to-target length along an open magnetic field lines is called the connection length L. Compared to limiter, a divertor configuration has many advantages despite of it's complex structure. These advantages include: (1) enhanced plasma confinement by reducing recycling, (2) reduced heat load on targets, (3) improved impurity control, (4) greater operational flexibility. Therefore, for high-parameter long-pulse steady-state operation, the divertor is a better choice.

As an advanced divertor concept, the island divertor was successfully implemented on the W7-AS stellarator [24-26] and further optimized on W7-X [27, 28]. More recently, an island divertor configuration has also been constructed and studied on the J-TEXT tokamak [29-31]. In W7-X, the presence of a natural radial magnetic field allows for the formation of a natural chain of magnetic islands at the plasma edge. The significant radial magnetic field component and the weak magnetic shear in W7-X enable the generation of magnetic islands with radial widths on the order of centimeters. These islands are intersected by plates, resulting in the formation of open field lines and residual islands within the scrape-off layer. As depicted in Figure 1.8, the divertor configuration of W7-X has the same elements as tokamak divertors, including the scrape-off layer, divertor plates and baffles, X-point, and private flux region. However, W7-X exhibits distinct features: its magnetic topology is not toroidally symmetric, and its divertors are discontinuous. These differences result in significant disparities in the edge magnetic topology between tokamak and stellarator divertors, which can be characterized by the connection length L_c. For a tokamak, the connection length L_c can be approximated by $2\pi R/(N\iota)$, where R is the major radius, N is the number of nulls and ι is the rotational transform. Typically, the connection length is on the order of tens of meters. While for an island divertor, the connection length does not directly depend on the rotational transform. Instead, it is inversely proportional to the magnetic shear of the field lines. Similar to tokamaks, the connection length in island divertors can be written as $L_c = 2\pi R/(N r_i \iota')$, where N is the number of islands, r_i is the radial size of island and ι' is the shear at the resonance [32]. Therefore, for a low-shear stellarator like W7-X, the connection length is usually longer than that of a tokamak, reaching the order of hundreds of meters.

The length of L_c determines the time it takes for charged particles to reach the materials, significantly affecting the particle and heat flux distributions on limiter or divertor targets. Now, let's consider a simple SOL model where all ionization processes occur within the closed flux surfaces. In other words, there is no particle source term in the SOL region, and all particles are from core by cross-field diffusion. The connection length of magnetic field lines in the SOL region, from one target plate to another, is given by $L_c=2\pi R/(N\iota)$. Assuming the parallel particle velocity is the ion sound speed c_s, the characteristic transport time in the parallel direction can be determined as $\tau_\parallel = L_c/0.5c_s$. The particle radial characteristic width in the SOL is λ_n, and the perpendicular velocity is $v_\perp = \Gamma/n$. By assuming that the characteristic transport time in the parallel direction equals the radial characteristic time, the relationship can be obtained:

$$\frac{\lambda_n}{v_\perp} = \frac{L_c}{0.5c_s} \qquad 1.3.1$$

Substituting $v_\perp = \Gamma/n = -D_\perp/\lambda_n$, the particle decay length on the target is given by:

$$\lambda_n = \sqrt{\frac{D_\perp L_c}{0.5c_s}} \qquad 1.3.2$$

Similarly, the energy decay length in the SOL region can be obtained as follows:

$$\lambda_q = L\sqrt{\frac{\chi_\perp}{\chi_\parallel}} \qquad 1.3.3$$

where $\chi_\perp$ and $\chi_\parallel$ are cross-field and parallel heat transport coefficients, respectively. If λ_q is larger, the radial spreading of the heat flux on the divertor target plates is greater, resulting in a lower heat load per unit area on the target plates. Therefore, for stellarators, which have longer connection lengths and consequently larger λ_q, this characteristic is advantageous for reducing the peak heat load on the target plates. Consequently, compared to a divertor, an island divertor is more favorable for long-pulse steady-state operation.

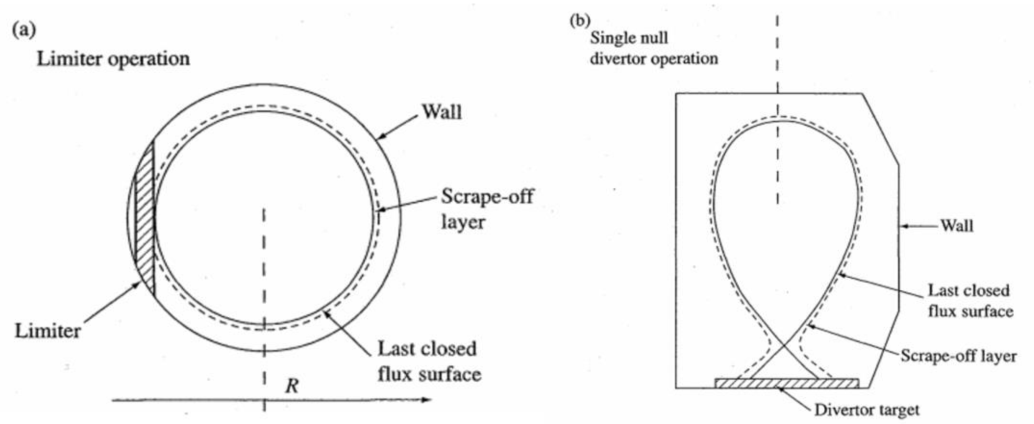

Figure 1.7: Cross-section in a tokamak (a) with a limiter configuration, (b) with a single-null divertor configuration. Image adapted from [11].

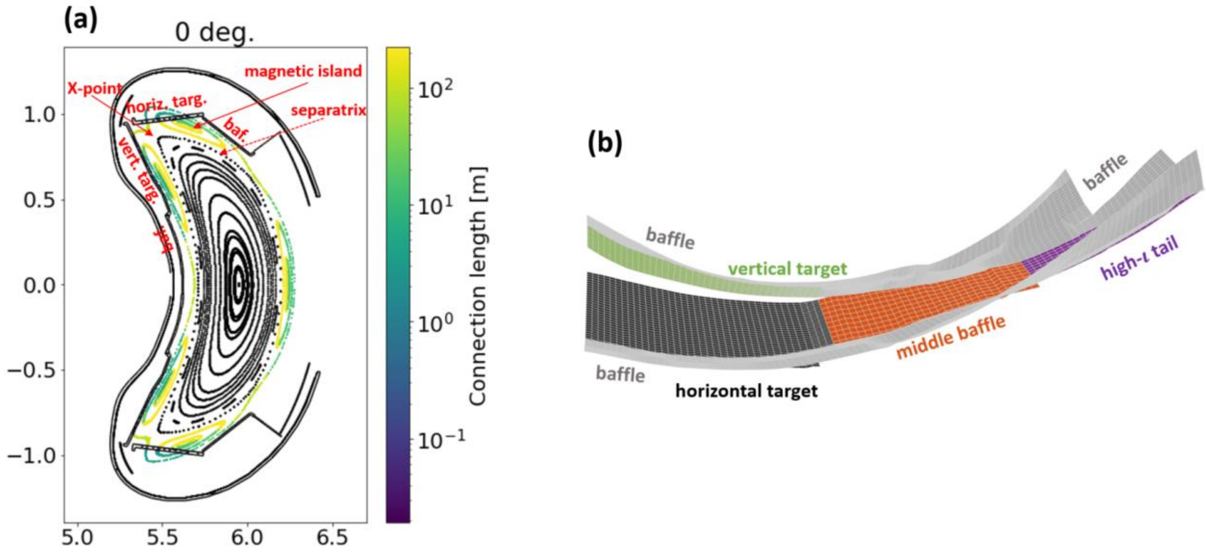

Figure 1.8: (a) Poincaré plot of the standard W7-X magnetic configuration on a bean-shaped cross-section, (b) divertor geometry in W7-X, with each color denoting a distinct divertor component.

1.4 The importance of impurity transport

For long-pulse, high-performance operation, the main obstacle is protecting the divertor targets from excessive heat loads. For widely used tungsten targets, the maximum tolerable load is about 10 MW m^{-2} [33], which strongly constrains the high-performance operation. When power leaves the core and enters the scrape-off layer (SOL), the energy decay length is short, so the heat spreads over only a narrow "wetted" area. The resulting peak flux on the targets is very large [34]. Heat damage to the divertor targets releases large amounts of high-Z impurities, but their concentration in the core must be kept very low (for tungsten, below 10^{-5}); otherwise, the resulting bremsstrahlung radiation will severely degrade plasma performance. The most effective mitigation strategy is to dissipate the power by radiation before it reaches the plates, that is, to operate in a detached regime.

As the line-averaged density is raised, the edge plasma evolves through three characteristic regimes [35].

1. Low-recycling regime: the divertor plasma remains hot and dilute; downstream density rises roughly linearly with the upstream value.
2. High-recycling regime: continued fuelling cools the SOL, strong parallel temperature gradients appear, and the downstream density increases more rapidly.
3. Detached regime: the electron temperature near the plate falls to only a few eV, the local density saturates and even rolls over, and the heat flux reaching the targets decreases sharply.

In W7-X, detachment is sustained by impurity radiation, supplied passively via sputtering from plasma-facing components or actively via puffing low-Z gases. Maintaining stable detachment requires a solid understanding of impurity transport. Too high a density or excessive impurity injection can trigger edge radiation instabilities such as MARFE [36-38] or even disruptions [39], whereas core accumulation dilutes the fuel and degrades performance.

Diagnostics alone cannot capture detail impurity behavior, so both qualitative and quantitative studies must rely on transport models. The classical two-point model [35] treats the SOL as only two "points" (upstream and downstream) and, by solving simplified fluid equations, links the downstream parameters to upstream conditions. The later "extended" two-point model includes energy losses, momentum losses and convective heat transport, but it still neglects cross-field transport, impurities and plasma-neutral interaction. As a one-dimensional formulation, it cannot represent complex 2D or 3D geometries and is not valid in high-recycling or detached regimes. For more accurate analysis, researchers turn to integrated transport codes that include many additional physics processes, such as the 2D code SOLPS-ITER [40] and the fully 3D code EMC3-EIRENE [41]. These sophisticated tools solve not only the plasma fluid equations but also incorporate plasma-wall interactions and neutral-particle transport. These codes have greatly deepened our understanding of divertor physics and have become indispensable tools for plasma-transport studies and the design of next-generation divertor concepts.

1.5 Research objectives and thesis outline

W7-X controls heat and particle exhaust with a fully three-dimensional island divertor. The complex edge geometry, together with limited diagnostic access, makes quantitative studies of edge-plasma transport, especially impurity transport, considerably more difficult. The three-dimensional transport code EMC3-EIRENE, which couples the EMC3 fluid model with the kinetic neutral solver EIRENE, can deal with such geometry and has already advanced work on edge transport, radiation dissipation, divertor heat load, particle exhaust [42-53]. Most published EMC3-EIRENE studies, however, focus on individual physics questions and draw on part of the W7-X boundary-diagnostic data. While such targeted analyses clarify specific phenomena, an overall understanding still demands that all available edge measurements be used to constrain the many input parameters in EMC3-EIRENE.

This motivation yields the main research objective: to investigate edge-plasma behaviour, especially impurity transport, by constraining EMC3-EIRENE with multiple W7-X edge diagnostics in the island divertor configuration. To make this objective practical, a neural-network surrogate is introduced that converts the multiple diagnostics directly into EMC3-EIRENE-consistent predictions, removing repeated manual tuning and the need for full runs. To support this goal, a synthetic construction of the FZJ divertor endoscope spectroscopic system is developed in EMC3-EIRENE so that its line-emission signals can be further used to refine the input parameters.

To meet the research objectives, the thesis needs to addresses five key questions:

1. How can synthetic diagnostics, particularly spectroscopic measurements, be implemented in EMC3-EIRENE? (Chapter 3)

2. Which strategies improve agreement between EMC3-EIRENE simulations and upstream/downstream diagnostics on W7-X? (Chapter 4)
3. What automated method can rapidly infer the EMC3-EIRENE input set most consistent with the diagnostics, avoiding manual trial-and-error? (Chapter 5)
4. How can a surrogate model achieve fast EMC3-EIRENE input-to-output predictions at much lower computational cost? (Chapter 6)
5. Can an EMC3-EIRENE-AI framework be developed that, in real time, converts diagnostic measurements with uncertainties into full 3D plasma profiles and key transport characteristics—namely the cross-field particle diffusivity and the thermal diffusivity? (Chapter 6)

The thesis is organized as follows.

Chapter 2 gives a concise overview of the EMC3-EIRENE code. It first outlines the derivation of the magnetohydrodynamic equations adopted in the model, then describes the impurity transport module added to EMC3, and explains how EMC3 is coupled to the neutral-particle code EIRENE. The chapter concludes with a discussion of the capabilities and limitations of EMC3-EIRENE. The purpose of this chapter is to clarify the code's range of applicability.

Chapter 3 presents the experimental diagnostics used for comparison and explains how their synthetic diagnostics are implemented in EMC3-EIRENE. The main focus is the calibration of the divertor endoscope spectrometer and edge and divertor spectroscopy developed by our group. Because spectroscopic signals cannot be obtained directly from the raw EMC3-EIRENE outputs, the chapter gives a step-by-step description of the synthetic construction. Photon-emissivity coefficients are first downloaded from Open-ADAS, then a Monte-Carlo algorithm traces each line of sight through the EMC3-EIRENE mesh and calculates the contribution of every intersected cell to the chosen line emission. Finally, the contributions are integrated to produce synthetic signals that can be compared with the measurements.

Chapter 4 refines the EMC3-EIRENE input set by comparing synthetic signals with the corresponding experimental diagnostics. Spatially non-uniform cross-field transport coefficients are introduced to improve the agreement. The chapter also examines how equilibrium effects (plasma beta and toroidal current) modify the edge magnetic topology and, in turn, the temperature, density and impurity profiles.

Chapter 5 replaces the manual iterative alignment of EMC3-EIRENE with edge diagnostics by using a Bayesian neural-network framework. From diagnostic measurements with their uncertainties, the network rapidly infers input parameters consistent with the observations. This greatly accelerates impurity transport analysis, reducing it to a single full EMC3-EIRENE run using the inferred inputs.

Chapter 6 develops EMC3-EIRENE-AI, a surrogate model for rapid edge transport analysis. It further removes the need for a full run after inference by combining the Chapter-5 Bayesian neural-network inference with a newly trained EMC3-EIRENE surrogate that maps the inferred inputs to full-code outputs. EMC3-EIRENE-AI provides near-real-time, code-consistent predictions directly from multiple diagnostic measurements.

Chapter 7 summarizes the thesis, highlights its limitations, and outlines directions for future work.

2 Edge plasma transport code EMC3-EIRENE

Edge plasma physics is a key topic in magnetic confinement fusion research, because it not only influences core confinement but also controls the heat and particle loads on the divertor. A complex magnetic topology, such as that of W7-X, makes the problem even more challenging. A comprehensive edge transport model must capture not only main-plasma transport but also neutrals and the production and motion of and impurities. Indeed, the edge plasma regime is strongly governed by impurity behaviour and by plasma-neutral interactions. When the plasma approaches detachment, the energy and momentum losses driven by impurities and neutrals become particularly critical.

This chapter starts with a brief derivation of the magnetohydrodynamic equations and explains the key approximations adopted to obtain a tractable fluid model. It then describes the EMC3 code in detail [54], covering the underlying assumptions, its impurity transport module, and the diffusion approach used to represent anomalous cross-field transport. The discussion moves on to the coupling between EMC3 and the kinetic neutral solver EIRENE [55], outlining the steps required for a full EMC3-EIRENE run. The chapter ends with a review of the capabilities and current limitations of the code package.

2.1 Fluid description of plasma transport

When plasma is treated with a fluid approach, it is regarded as an electromagnetic fluid whose parameters depend only on spatial coordinates and time. In a kinetic description, by contrast, particles are distinguished by their velocities, and the distribution function becomes a function of velocity space in addition to position and time. Assuming the plasma locally follows a Maxwell distribution, one can obtain fluid equations by taking moments of the kinetic equation and suitably approximating the higher-order moments. The derivation of the fluid conservation laws from the kinetic equation is outlined below.

We first introduce the distribution function $f_a(\boldsymbol{r}, \boldsymbol{v}, t)$, which gives the number of particles of species a contained in the phase-space element $d\boldsymbol{v}d\boldsymbol{r}$ around position $\boldsymbol{r}$ and velocity $\boldsymbol{v}$ at time t. The evolution of this function is governed by the Boltzmann equation

$$\frac{\partial f_a}{\partial t} + \boldsymbol{v}\frac{\partial f_a}{\partial \boldsymbol{r}} + \frac{\boldsymbol{F_a}}{m_a} \cdot \frac{\partial f_a}{\partial \boldsymbol{v}} = C_a \qquad 2.1.1$$

This equation describes the non-equilibrium evolution of species a in phase space $(\boldsymbol{r}, \boldsymbol{v})$. Here, f_a is the distribution function of species a, $\frac{\partial f_a}{\partial \boldsymbol{r}}$ is its spatial gradient, $\boldsymbol{F_a}$ is the external force on species a, which in a plasma is mainly the Lorentz force $\boldsymbol{F} = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B})$, $\frac{\partial f_a}{\partial \boldsymbol{v}}$ is the gradient in velocity space, and $C_a = \left(\frac{\partial f_a}{\partial \boldsymbol{t}}\right)_c$ represents the rate of change of the distribution function due to inter-particle collisions. If the plasma is sufficiently hot so that collisions can be neglected, the Boltzmann equation reduces to

$$\frac{\partial f_a}{\partial t} + \boldsymbol{v}\frac{\partial f_a}{\partial \boldsymbol{r}} + \frac{\mathrm{q}(\mathbf{E}+\mathbf{v}\times\mathbf{B})}{m_a}\frac{\partial f_a}{\partial \boldsymbol{v}} = 0 \tag{2.1.2}$$

known as the Vlasov equation, which expresses conservation of particle number in phase space under collisionless conditions. When collisions cannot be ignored but are dominated by binary Coulomb encounters, the Boltzmann equation can be approximated by the Fokker-Planck equation

$$\frac{\partial f_a}{\partial t} + \boldsymbol{v}\frac{\partial f_a}{\partial \boldsymbol{r}} + \frac{\mathrm{q}(\mathbf{E}+\mathbf{v}\times\mathbf{B})}{m_a}\frac{\partial f_a}{\partial \boldsymbol{v}} = -\frac{\partial}{\partial v_i}(f_a\langle\triangle\, v_i\rangle) + \frac{1}{2}\frac{\partial^2}{\partial v_i \partial v_j}\left(f_a\langle\triangle\, v_i \triangle\, v_j\rangle\right) \tag{2.1.3}$$

where $\triangle\, v_i$ is the change in the i-th velocity component during a collision.

Before taking moments of the Boltzmann equation, the statistical moments of the distribution function $f_a(\boldsymbol{v}, \boldsymbol{r}, t)$ for species $\boldsymbol{a}$ are introduced. The n-th moment is defined as

$$\langle f_a(\boldsymbol{v}, \boldsymbol{r}, t)\rangle^n = \int v^n f_a(\boldsymbol{v}, \boldsymbol{r}, t) d\boldsymbol{v} \tag{2.1.4}$$

Common low-order moments are: Zero-th moment (density), $n_a = \int f_a d\boldsymbol{v}$; First moment (fluid velocity), $\boldsymbol{u_a} = \frac{1}{n_a}\int \boldsymbol{v} f_a d\boldsymbol{v}$; Second moment (temperature), $T_a = \frac{m_a}{3n_a}\int(\boldsymbol{v} - \boldsymbol{u_a})^2 f_a d\boldsymbol{v}$. Multiplying both sides of Eq. (2.1.1) by an arbitrary function $\phi(\boldsymbol{v})$ and integrating over velocity space gives

$$\int \phi(\boldsymbol{v})\frac{\partial f_a}{\partial t}d\boldsymbol{v} + \int \phi(\boldsymbol{v})\boldsymbol{v}\cdot\frac{\partial f_a}{\partial \boldsymbol{r}}d\boldsymbol{v} + \int \phi(\boldsymbol{v})\frac{\boldsymbol{F_a}}{m_a}\cdot\frac{\partial f_a}{\partial \boldsymbol{v}}d\boldsymbol{v} = \int \phi(\boldsymbol{v})C_a d\boldsymbol{v} \tag{2.1.5}$$

Using the low-order moments, this becomes

$$\frac{\partial}{\partial t}(n_a\langle\phi(\boldsymbol{v})\rangle) + \frac{\partial}{\partial \boldsymbol{r}}\cdot(n_a\langle\phi(\boldsymbol{v})\boldsymbol{v}\rangle) - \frac{n_a q}{m_a}E\cdot\langle\frac{\partial\phi(\boldsymbol{v})}{\partial \boldsymbol{v}}\rangle - \frac{n_a q}{c m_a}\langle(\mathbf{v}\times\mathbf{B})\cdot\frac{\partial\phi(\boldsymbol{v})}{\partial \boldsymbol{v}}\rangle = \int \phi(\boldsymbol{v})C_a d\boldsymbol{v} \tag{2.1.6}$$

Choosing $\phi(\boldsymbol{v}) = 1$ yields the particle-balance equation

$$\frac{\partial n_a}{\partial \boldsymbol{v}} + \frac{\partial}{\partial \boldsymbol{r}}\cdot(n_a\boldsymbol{u}_a) = \int C_a d\boldsymbol{v} \tag{2.1.7}$$

With $\phi(\boldsymbol{v}) = m_a\boldsymbol{v}$ one obtains the momentum equation

$$\frac{\partial}{\partial t}(m_a n_a \boldsymbol{u}_a) + \frac{\partial}{\partial \boldsymbol{r}}\cdot(m_a n_a\langle\boldsymbol{v}\boldsymbol{v}\rangle) - n_a q_a\left(\boldsymbol{E} + \frac{1}{c}(\boldsymbol{u}\times\boldsymbol{B})\right) = \int m_a\boldsymbol{v}C_a d\boldsymbol{v} \tag{2.1.8}$$

Taking $\phi(\boldsymbol{v}) = \frac{1}{2}m_a v^2$ gives the energy equation

$$\frac{\partial}{\partial t}\left(\frac{3}{2}n_a T_a + \frac{1}{2}m_a n_a u_a^2\right) + \frac{\partial}{\partial \boldsymbol{r}}\cdot\left(\frac{1}{2}m_a n_a\langle\boldsymbol{v}v^{\boldsymbol{2}}\rangle\right) - n_a q_a\mathbf{E}\cdot\boldsymbol{u}_a = \int\frac{1}{2}m_a v^2 C_a d\boldsymbol{v} \tag{2.1.9}$$

These three conservation equations still contain higher-order moments $\langle \boldsymbol{vv} \rangle$ and $\langle \boldsymbol{v}v^{2} \rangle$. To close the set, approximations for these moments are required.

Introduce several quantities related to higher moments: pressure P_a, viscous stress tensor $\boldsymbol{\pi_a}$, and heat flux $\boldsymbol{h_a}$:

$$P_a = n_a T_a \tag{2.1.10}$$

$$\boldsymbol{\pi_a} = m_a n_a \left\langle (\boldsymbol{v} - \boldsymbol{u}_a)(\boldsymbol{v} - \boldsymbol{u}_a) - \frac{(\boldsymbol{v}-\boldsymbol{u}_a)^2}{3}\boldsymbol{I} \right\rangle \tag{2.1.11}$$

$$\boldsymbol{h_a} = \frac{m_a n_a}{2} \langle (\boldsymbol{v} - \boldsymbol{u}_a)(\boldsymbol{v} - \boldsymbol{u}_a)^2 \rangle \tag{2.1.12}$$

Here $\boldsymbol{I}$ is the unit tensor. For a magnetized plasma where collision times are long compared with gyro periods, Braginskii assumed the distribution is a Maxwellian plus a small perturbation and obtained simplified forms:

$$\boldsymbol{h_a} = -k_{a\parallel} \nabla_{\parallel} T_a - k_{a\perp} \nabla_{\perp} T_a - k_{a\wedge} \boldsymbol{b}_{\parallel} \times \nabla T_a \tag{2.1.13}$$

$$\boldsymbol{b}_{\parallel} \cdot \boldsymbol{\pi_a} \cdot \boldsymbol{b}_{\parallel} = -\eta_{a\parallel} \boldsymbol{b}_{\parallel} \cdot \nabla_{\parallel} u_{a\parallel} \tag{2.1.14}$$

with $k_{a\parallel}$, $k_{a\perp}$, $k_{a\wedge}$ the parallel, perpendicular, and cross heat-conductivities, $\nabla_{\parallel}$ and $\nabla_{\perp}$ the gradients along and across the magnetic field, and $\eta_{a\parallel}$ the parallel viscosity coefficient. Consequently,

$$\langle \boldsymbol{vv} \rangle = \frac{1}{m_a n_a} (\boldsymbol{\pi_a} + P_a \boldsymbol{I} + m_a n_a \boldsymbol{u}_a \boldsymbol{u}_a) \tag{2.1.15}$$

$$\langle \boldsymbol{v}v^{2} \rangle = \frac{2}{m_a n_a} \left(\boldsymbol{h_a} + \boldsymbol{u}_a \cdot \boldsymbol{\pi_a} + \frac{5}{2} n_a T_a \boldsymbol{u}_a + \frac{m_a n_a}{2} {u_a}^2 \boldsymbol{u}_a\right) \tag{2.1.16}$$

2.2 The EMC3 solver

The EMC3 code, developed by Prof. Y. Feng and co-workers, solves a reduced Braginskii fluid model on fully three-dimensional magnetic geometry [54]. The code was first applied to studies on the W7-AS stellarator. In toroidal devices, the plasma consists of a main ion species plus a small amount of impurities. Impurities arise from two sources: (1) intrinsic impurities released from plasma-facing components or residual gases such as water, and (2) impurities that are intentionally injected to reduce divertor heat load or to study impurity transports. For both main ions and impurities, anomalous transport across the magnetic field is described by a diffusion model. The cross-field particle flux is

$$\boldsymbol{\Gamma_{i\perp}} = -D_{\perp} \nabla_{\perp} n_i \tag{2.2.1}$$

where $D_{\perp}$ is a free diffusion coefficient. EMC3 allows different diffusion values for different species and permits spatially non-uniform profiles. The cross-field heat flux is

$$\boldsymbol{h_{i\perp}} = -\chi_{\perp} n_i \nabla_{\perp} T_i \tag{2.2.2}$$

with $\chi_{\perp}$ the free heat-diffusion coefficient. Because anomalous transport is not yet well understood, these coefficients are usually chosen empirically or adjusted to match diagnostics.

To simplify the fluid equations, EMC3 adopts several assumptions for the main plasma: Steady state, $\frac{\partial}{\partial t} = 0$; electron contributions from impurity ionization are neglected; quasi-neutrality, $n_e = n_i$; electrons and ions share the same mean velocity, $\boldsymbol{u_e} = \boldsymbol{u_i}$, giving $\boldsymbol{j} = 0$; bohm sheath condition at the boundary.

Under these assumptions, the particle balance for species a becomes

$$\nabla \cdot (n_a u_{a\parallel} \boldsymbol{b}_\parallel - D_\perp \boldsymbol{I}_\perp \cdot \nabla n_a) = S_p \qquad 2.2.3$$

where $\boldsymbol{I}_\perp = \boldsymbol{I} - \boldsymbol{b}_\parallel \boldsymbol{b}_\parallel$ and S_p is the particle source from electron-neutral collisions.

The momentum equation is

$$\nabla \cdot \boldsymbol{b}_\parallel \left(m_a n_a u_{a\parallel}^2 - \eta_{a\parallel} \boldsymbol{b}_\parallel \cdot \nabla_\parallel u_{a\parallel} \right) - \nabla \cdot \boldsymbol{I}_\perp \cdot D_\perp \nabla (m_a n_a u_{a\parallel}) + \boldsymbol{b}_\parallel \cdot \nabla P = S_m \qquad 2.2.4$$

with S_m the momentum source from plasma-neutral collisions.

The energy equations for electrons and ions are

$$\nabla \cdot \boldsymbol{b}_\parallel \left(\frac{5}{2} T_e n_e u_{e\parallel} - k_{e\parallel} \boldsymbol{b}_\parallel \cdot \nabla T_e \right) - \nabla \cdot \boldsymbol{I}_\perp \cdot \left(\chi_{e\perp} n_e \nabla T_e + \frac{5}{2} T_e D_\perp \nabla n_e \right)$$
$$= -\frac{3m_e}{m_i} \frac{n_e}{\tau_{ei}} (T_e - T_i) + S_{ee} - S_{e,cool} \qquad 2.2.5$$

$$\nabla \cdot \boldsymbol{b}_\parallel \left(\frac{5}{2} T_i n_i u_{i\parallel} - k_{i\parallel} \boldsymbol{b}_\parallel \cdot \nabla T_i \right) - \nabla \cdot \boldsymbol{I}_\perp \cdot \left(\chi_{i\perp} n_i \nabla T_i + \frac{5}{2} T_i D_\perp \nabla n_i \right)$$
$$= +\frac{3m_e}{m_i} \frac{n_e}{\tau_{ei}} (T_e - T_i) + S_{ei} \qquad 2.2.6$$

where S_{ee} and S_{ei} are energy sources from neutral collisions, and $S_{e,cool}$ is the radiative cooling sink due to impurities.

Because impurity concentrations are much lower than those of the main ions, impurity transport can be solved separately. Neglecting viscosity, the parallel force balance gives the impurity transport equation

$$\nabla \cdot (n_a u_{a\parallel} \boldsymbol{b}_\parallel - D_{a\parallel} \boldsymbol{b}_\parallel \boldsymbol{b}_\parallel \cdot \nabla n_a - D_{a\perp} \boldsymbol{I}_\perp \cdot \nabla n_a) = S_p^{imp} \qquad 2.2.7$$

where S_p^{imp} is the source term from ionization and recombination between adjacent charge states. Only nearest-neighbour transitions are retained, i.e.

$$S_p^{imp} = R_{Z-1\to Z}^{ion} + R_{Z+1\to Z}^{rec} - R_{Z\to Z+1}^{ion} - R_{Z\to Z-1}^{rec} \qquad 2.2.8$$

In equation 2.2.5 the electron-cooling term $S_{e,cool}$ represents the energy lost through ionization and radiation:

$$S_{e,cool} = \sum_{n=0}^{Z_{max}-1} E_n R_{n\to n+1}^{ion} + \sum_{n=0}^{Z_{max}-1} P_n^{rad} \qquad 2.2.9$$

where E_n is the ionization energy of charge state n and P_n^{rad} is its line-radiation power. These equations form the basis of the EMC3 fluid model.

2.3 Coupling EMC3 with the EIRENE

EIRENE, developed by D. Reiter and co-authors, is a kinetic solver for neutral-particle transport. By coupling EMC3 with EIRENE, neutral motion, atomic and molecular reactions, and plasma-neutral interactions can be treated self-consistently, greatly extending the code's capability. Figure 2.1 illustrates the coupling workflow. A three-dimensional mesh is first built from the magnetic topology and the plasma-facing components (PFCs). The PFCs define the simulation domain, and their material properties are supplied so that EIRENE can handle wall interactions correctly. The combined code uses a Monte Carlo technique and therefore follows a large number of particle trajectories. EMC3 solves the charged-particle transport equations and passes background data, electron temperature, ion temperature, electron density, flow velocity, and the particle and energy loads on divertor targets, to EIRENE. With these inputs, EIRENE solves the Boltzmann equation for neutrals and returns the resulting particle, momentum, and energy source terms to EMC3. The two codes iterate in this way until convergence. All coefficients required to determine the impurity charge-state distribution and the associated radiation are taken from the ADAS database [56], which is built on collisional-radiative modelling. In practice, convergence is declared when the relative change in electron temperature, ion temperature, and density between successive iterations falls below about 1 %.

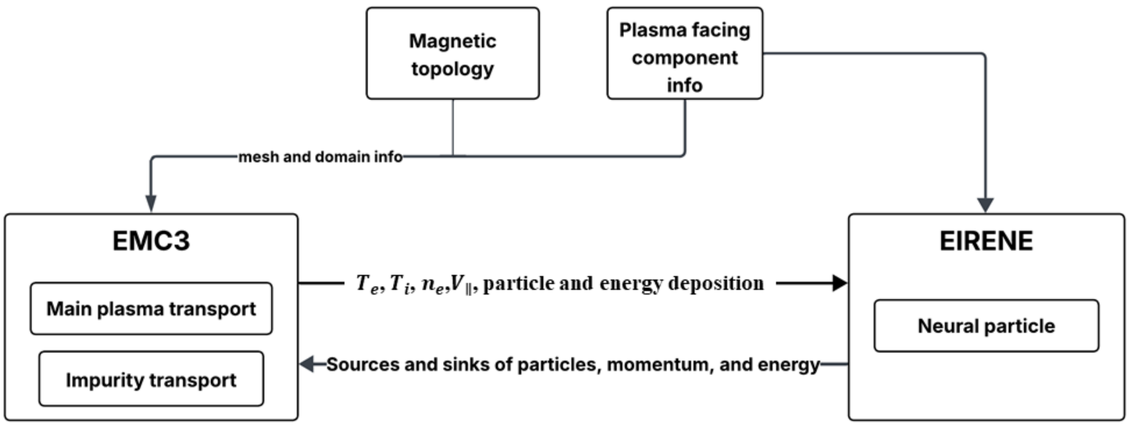

Figure 2.1: Schematic workflow of the coupled EMC3-EIRENE calculation.

2.4 Capabilities and limitations of EMC3-EIRENE

Typical input parameters for EMC3-EIRENE include the auxiliary heating power, the total radiated power, the separatrix density, the cross-field transport coefficients, the impurity species, the way impurities are introduced (e.g. gas puffing or sputtering), the sputtering yield, the initial energy of sputtered particles and their angular distribution, and so on. Radiated power and sputtering yield are not independent; they are coupled self-consistently through the main-plasma, impurity-transport and neutral-gas equations solved in EMC3-EIRENE together with the ADAS data base. ADAS, which is built on collisional-radiative (CR) modelling, supplies the CR coefficients that the code uses to determine impurity charge-state populations and the associated radiative losses. Users can switch individual atomic and molecular processes on or off within the code. The magnetic equilibrium is taken into account when the computational grid is generated; to study beta effects [11] on the edge, one first obtains a pressure-balanced field with the HINT code [57, 58], imports that equilibrium into EMC3-EIRENE, and then computes the resulting edge plasma profiles. With its flexible set of input parameters, EMC3-EIRENE can be readily adapted to a wide range of research scenarios.

The output is equally rich, providing electron temperature and density, ion temperature, target particle and heat fluxes, parallel Mach number, total radiation, impurity charge-state densities, and many other quantities. These features make EMC3-EIRENE a powerful tool for edge transport studies, and its outstanding strength is the ability to handle fully three-dimensional boundary magnetic structures. The code was first applied to the 3-D island divertor of W7-AS [32, 54, 59]. Since then it has been used extensively on W7-X, for example to analyze detachment [49], to study conditions and benefits of X-point radiation [48], to model impurity transport in different magnetic configurations [42, 45], to follow neutral particle behavior in attached and detached conditions [51], and to assess how plasma beta shifts the detachment threshold [47]. On LHD it has been employed to examine impurity transport in the stochastic layer [60, 61]. Tokamak applications include impurity studies during toroidally localized lithium injection on EAST [62, 63] and edge plasma transport under lower-hybrid-wave-induced magnetic perturbations on the same device [52]. Further examples are edge-transport simulations of the snowflake divertor on TCV [64, 65] and impurity transport during toroidally localized divertor gas injection experiments on Alcator C-Mod [66].

However, EMC3-EIRENE also has limitations. In the present version, it ignores $\boldsymbol{E}\times\boldsymbol{B}$ drifts. By assuming $n_e = n_i$ and $\boldsymbol{u_e} = \boldsymbol{u_i}$, the code sets the current $\boldsymbol{J}$ to zero and therefore gives $\boldsymbol{E} = \eta\boldsymbol{J} = 0$, which removes the $\boldsymbol{E}\times\boldsymbol{B}$ drift. On W7-X, drift effects are non-negligible at low density [67, 68]. Figure 2.2 shows heat-flux measurements from nine infrared cameras viewing different 9 divertor targets at attached plasma and all signals have been mapped to the same toroidal location as the divertor probe. Clear differences appear between the upper and lower divertors. The upper targets receive consistently higher heat fluxes. Such up-down and toroidal asymmetries, driven by drifts, cannot be reproduced with the present version of EMC3-EIRENE. At higher density, the drift-induced asymmetry becomes small and can be neglected. Accordingly, EMC3-EIRENE has limitations for problems where drift physics controls the observed asymmetry.

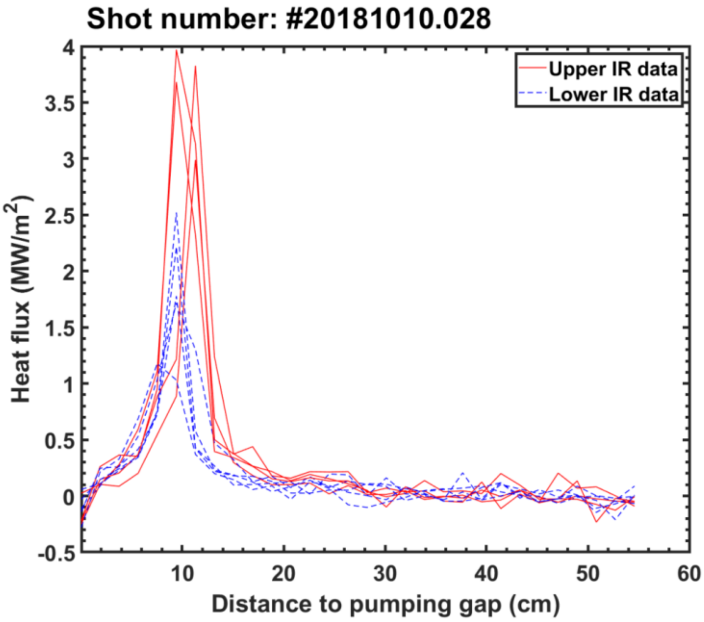

Figure 2.2: Heat flux profiles on W7-X divertor targets measured by nine IR cameras (upper targets: solid red; lower targets: dashed blue), mapped to the divertor probe toroidal position and plotted versus distance from the pumping gap.

Moreover, the three-dimensional nature of EMC3-EIRENE makes it computationally demanding. In W7-X studies, the device's five-fold symmetry is usually exploited to shrink the toroidal domain from 360° to 36°, reducing the mesh to roughly one-tenth of its full size. Even with this simplification, a single simulation at relatively fine resolution still requires on the order of several thousand CPU-core hours. Such heavy resource demands remain a primary barrier to the routine, widespread use of EMC3-EIRENE.

2.5 Summary

This chapter has explained how the three-dimensional edge-transport code EMC3-EIRENE treats main plasma transport, impurity transport, and plasma-neutral interactions. The model applies to trace-impurity conditions in both stellarators and tokamaks, and its ability to resolve fully three-dimensional, non-axisymmetric magnetic geometry makes it particularly well suited to island divertor studies on W7-X and similar devices. Current limitations include the omission of error-field and $\boldsymbol{E}\times\boldsymbol{B}$ drift effects and the substantial CPU time demanded by a full 3D grid, which is much higher than for 1D or 2D transport codes.

3 Edge diagnostic systems on W7-X

Diagnostics are essential tools for monitoring plasma properties, they can provide a wide range of basic plasma parameters. Diagnostic methods can be divided into passive and active measurements. Passive diagnostics measure electromagnetic waves or particles emitted by plasma, which provide many information about plasma density, plasma temperature, impurities content, etc. Typical passive diagnostics include infrared cameras, divertor spectroscopic diagnostics, charge exchange recombination spectroscopy, and so on. Active diagnostics, on the other hand, involve the intentional application of some form of disturbance to the plasma, such as various probes, microwaves and neutral particle beams. These disturbances can cause responses from the plasma, and then the relevant plasma information can be obtained by measuring these response signals. While these methods may disturb the plasma to some extent, they are more flexible and can provide more detailed information compared to passive measurements. Validating diagnostic measurements is both necessary and meaningful, especially for these devices with complex 3D edge structure, like W7-X. However, cross-validation of diagnostics is challenging because the physical quantities measured by different diagnostics often vary, or the measurements are taken from different locations. EMC3-EIRENE is a powerful tool for integrating different diagnostics or diagnostics at different measurement locations. It can self-consistently validate the accuracy of various diagnostics.

This chapter begins with an overview of the key edge diagnostics on W7-X, focusing on divertor spectroscopic and endoscope systems. Synthetic constructions of these diagnostics are then implemented in EMC3-EIRENE, including an original synthetic divertor spectroscopy and endoscope module developed in this work. Comparing multiple synthetic signals with measurements helps choose EMC3-EIRENE inputs, such as the cross-field transport coefficients, so that the simulation agrees more closely with the experiment.

3.1 Edge-related diagnostics

3.1.1 Divertor Spectroscopy Endoscopes

Divertor spectroscopy endoscopes are integrated spectroscopic systems developed by Forschungszentrum Jülich (FZJ) to study edge impurity transport and plasma-wall interactions on W7-X. There are four sets of endoscopes installed at two cross-sections with toroidal angles $\varphi = 126^{\circ}$ and $\varphi = 306^{\circ}$. Figure 3.1 shows the output of the endoscope systems at the lower divertor and the optical path of the endoscope system. As seen in this figure, each cross-section has two sets of endoscopes with partially overlapping fields of view. The object area is approximately 70mm × 70mm. The field of view can be adjusted poloidally by scanning the first mirror in the front of the endoscope tube, which can rotate up to 30 degrees. The light from the plasma is split into two branches by a pyramidal mirror installed on a prism and then each branch is divided multiple times into ultraviolet (UV), visible (VIS) and infrared (IR) branches by several dichroic beam splitters before entering the corresponding detectors. Each endoscope is equipped with five filter cameras, comprising 3 VIS and 2 UV filter cameras, in addition to 1 infrared camera and 5 spectrometers. These spectrometers include 1 IR, 1 VIS, 2 UV and 1 overview spectrometer, providing comprehensive wavelength coverage from 300 to 7000 nm.

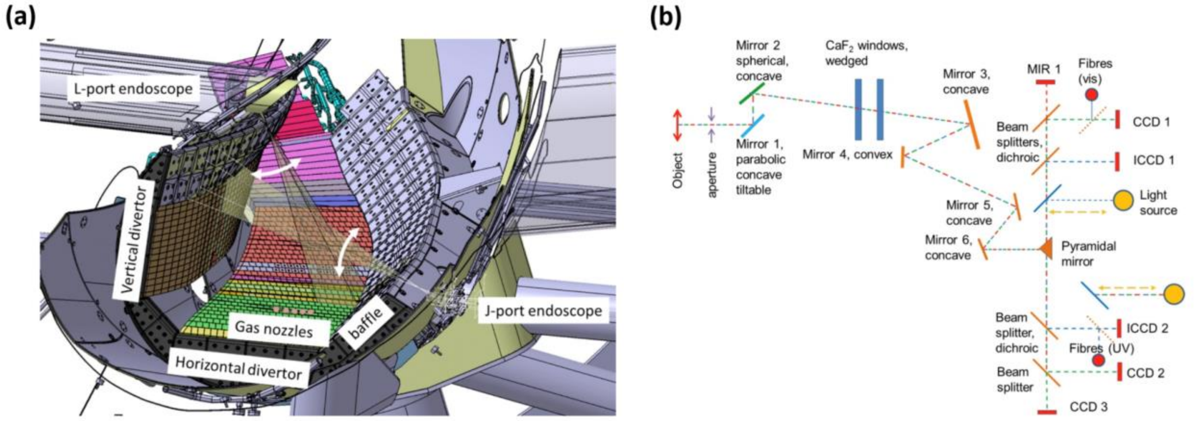

Figure 3.1: (a) Schematic view of the endoscope systems at the lower divertor, (b) Diagram of the optical path and rear-end detectors of the endoscope system. Images adapt from [69].

In the subsequent section, the calibration process for the endoscope system is explained. First, two fundamental physical quantities are introduced: radiant flux Φ and radiance $L_{radiance}$. Radiant flux Φ, also known as radiant power, is the amount of light energy per unit time, measured in watts (W). Radiance represents the radiant power per unit area per unit solid angle, measured in W/m²/sr. For the analysis, a simplified geometry for the light source and detector is assumed, as depicted in Figure 3.2(a). The relationship [70] radiant flux Φ and radiance $L_{radiance}$ can be expressed as:

$$\Phi = \iint L_{radiance} \, cos\theta \, dA_p d\Omega = \iint L_{radiance} \, cos\theta \, dA_p \frac{cos\,\beta dA_d}{s^2} \qquad 3.1.1$$

where A_p is the area of light source, A_d is the area of detector, Ω is solid angle of detector, θ is the angle between the line connecting the light source to the detector and the normal of dA_p and β is the angle between the line connecting the light source to the detector and the normal of dA_d. For small angles θ and β, this relationship can be simplified to

$$\Phi = L_{radiance} A_p \Omega_p = L_{radiance} A_p \frac{A_d}{s^2} \qquad 3.1.2$$

In practical measurements, the radiant power is typically influenced by factors such as the size of the entrance slit, mirrors, or lenses. These effects can be simplified to a virtual slit, as illustrated in Figure 3.2(b). Here, the radiant power can be written as:

$$\Phi = L_{radiance} A_p \Omega_s \quad 3.1.3$$

Since radiance is an invariant in an optical system with no absorption [70], assuming slit A_s is the new light source, the relationship is given by:

$$\Phi = L_{radiance} A_p \Omega_s = L_{radiance} A_s \Omega_m \quad 3.1.4$$

In the calibration process, an Ulbricht sphere is used as the light source, providing a uniform and isotropic light field. Figure 3.3 shows the spectral radiance *S* used in the calibration process, with units in W/m²/sr/nm. If there is a filter in front of detector, the radiance is given by:

$$L_{sphere} = \int_{\lambda}^{\lambda+\Delta\lambda} S(\lambda) T(\lambda)\, d\lambda \quad 3.1.5$$

where $T(\lambda)$ is the transmission of filter, and $\Delta\lambda$ is the bandwidth of filter. Therefore, the radiant power entering to the detector is

$$\Phi_{sphere} = L_{sphere} A_p \Omega_s = L_{sphere} A_s \Omega_m \quad 3.1.6$$

For the detector, the intensity flux in counts unit is

$$\Phi_{detector} = \frac{\sum_1^n N_i}{\Delta t} \quad 3.1.7$$

where N_i is the counts of i-th pixel, *n* is the number of pixels, and Δt is the exposure time. The ratio of Φ_{sphere} and $\Phi_{detector}$ is then:

$$R = \frac{\Phi_{sphere}}{\Phi_{detector}} = \frac{L_{sphere} A_s \Omega_m}{\frac{\sum_1^n N_i}{\Delta t}} \quad 3.1.8$$

Since A_s and Ω_m are fixed for specific systems and difficult to measure directly, a calibration factor *K* can be defined as:

$$K = \frac{R}{A_s \Omega_m} = \frac{L_{sphere}}{\frac{\sum_1^n N_i}{\Delta t}} \quad 3.1.9$$

The calibration factor for each single pixel on the detector is

$$K_i = \frac{L_{sphere}}{\frac{N_i}{\Delta t}} \quad 3.1.10$$

Thus, the radiance for the light from the plasma can be calculated from the detector counts:

$$L_{exp} = \frac{K N_{exp}}{L_{sphere} \Delta t} \quad \text{or} \quad L_{i\,exp} = \frac{K_i N_{i\,exp}}{L_{sphere} \Delta t} \quad 3.1.11$$

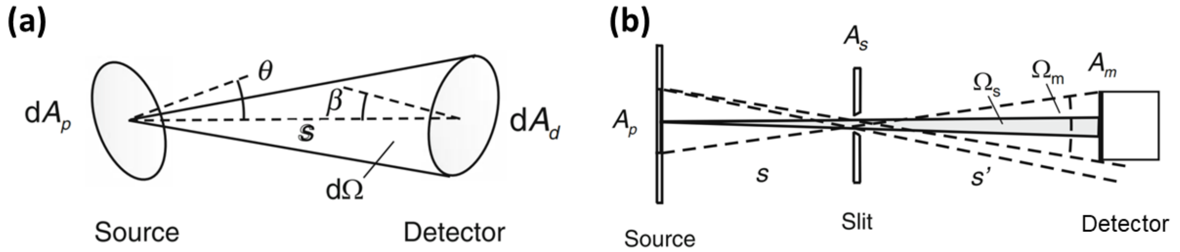

Figure 3.2: (a) Simplified geometry of light source and detector, (b) simplified optical path for each branch of endoscope system. Image adapted from [70].

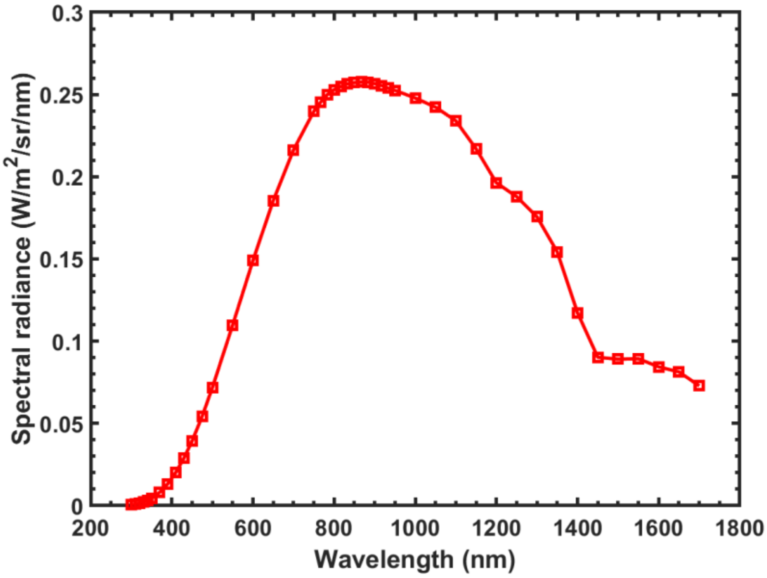

Figure 3.3: Typical spectral radiance of Ulbricht sphere.

3.1.2 Edge and Divertor Spectroscopy

Edge and divertor spectroscopy system comprises two parts: one part observes the upper divertor with both horizontal and vertical lines of sight (LOS) at $\varphi = 12.3^{\circ}$, while the other part focuses on lower divertor with similar horizontal and vertical lines of sight at $\varphi = -12.3^{\circ}$. Figure 3.4 illustrates the lines of sight at the upper divertor, which nearly covers the entire divertor region. The observation spans a wavelength range from UV to IR light, making it suitable for observing hydrogen Balmer series lines and impurity emission lines. Especially it can observe the emission lines from thermal helium beams injected through five divertor nozzles. By analyzing the ratio of these emissions, plasma density and temperature distributions can be calculated. In addition, Edge and divertor spectroscopy system can provide 2D plasma parameters through tomographic reconstruction using crossed lines of sight [71]. On W7-X, edge and divertor spectroscopy diagnostics are used to investigate impurity sources [72] and the radiation characteristics of detached plasmas [73]. The temporal resolution typically ranges from 50 to 500 ms, depending on the detector. The calibration process is similar to that of the divertor spectroscopy endoscopes system. By measuring the signal before and after switching the Ulbricht sphere, the background can be subtracted. Then, using equation 3.1.9, the calibration factor can be calculated.

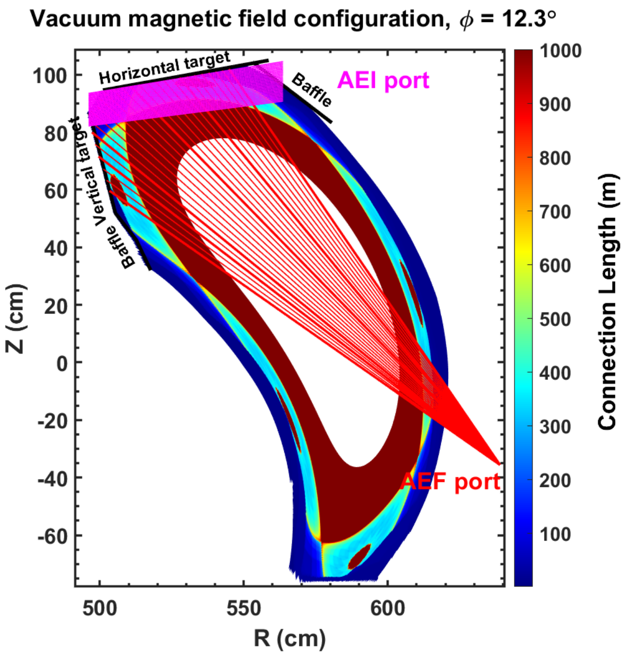

Figure 3.4: Lines of sight of edge and divertor spectroscopy system at the upper divertor with $\varphi = 12.3^{\circ}$. Magenta lines mark the horizontal views, while red lines mark the vertical views.

3.1.3 Multi-Purpose Manipulator

The multi-purpose manipulator (MPM) is a rapid reciprocating probe carrier system designed to observe upstream plasma parameters [74, 75]. The capabilities of the system vary based on the type of probe head being used. For example, a Langmuir probe can measure electron density and temperature, ion saturation current, and floating potential; Mach probes are used for plasma flow measurements; and an RFA probe can determine ion temperature, among other parameters. As depicted in Figure 3.5, the MPM system is installed at the mid-plane with a toroidal angle of $\varphi = 200.8°$ and $Z = -17$ cm. The probe can horizontally enter the plasma at a maximum speed of 3.5 m/s, with the deepest measurement point being roughly 2 cm from the last closed flux surface. The depth is limited by the heat load on probes. During plasma operation, inserting a probe into the plasma can obtain a radial profile of corresponding plasma parameters.

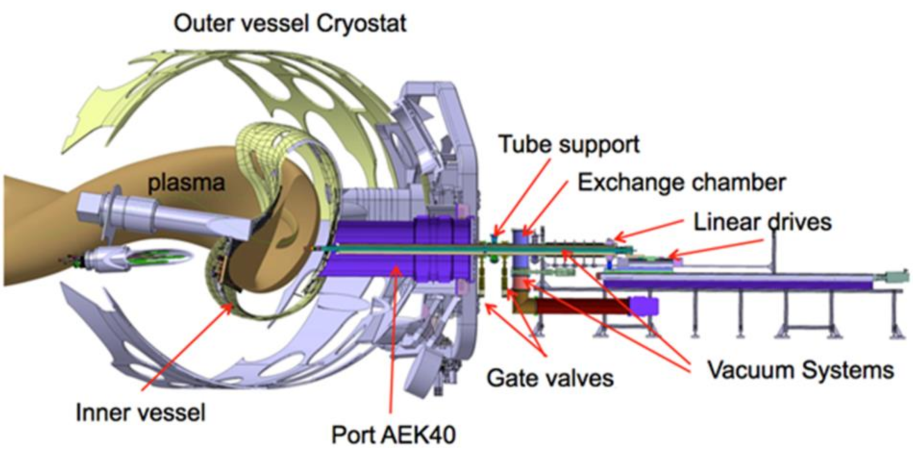

Figure 3.5: Setup of the Multi-purpose manipulator on W7-X. Image from [75].

3.2.4 Other diagnostics

Thomson Scattering: Thomson scattering (TS) is the elastic interaction between electromagnetic waves and free charged particles. This technology is commonly used to measure plasma properties, for example electron density and temperature [76, 77]. When the laser light passes thought the plasma, it interacts with free electrons along its path, causing the light to scatter in different directions. The scattered light contains information about plasma properties. The intensity of the scattered light is proportional to the electron density, while the Doppler broadening of the scattered light spectrum can provide information about the electron temperature. Figure 3.6 shows the schematic view of the Thomson scattering system on W7-X, located at a toroidal angle of $\varphi = 170.5°$. The TS system comprises a light source, scattering volumes, collection optics and detectors. The scattering volumes (shown in red) pass through the core plasma and intersect the magnetic islands in the standard configuration. The system typically achieves temporal and spatial resolutions of 100 ms and 0.5 - 3.5 cm, respectively. The measurement ranges for electron density and temperature are 10^{18}-10^{21} m^{-3} and 10-1000 eV, respectively.

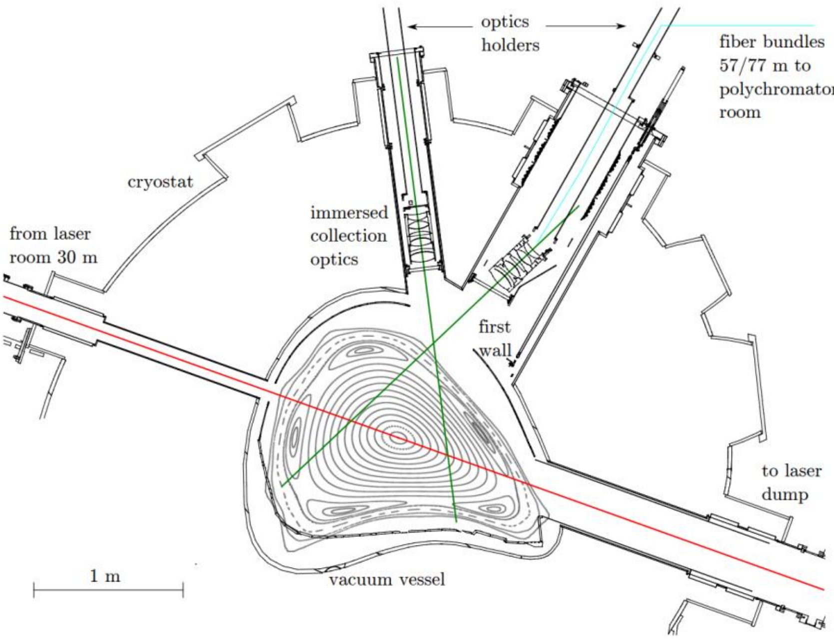

Figure 3.6: Schematic view of the Thomson scattering system on W7-X. Image from [76].

Alkaline Metal Beam: Alkaline metal beam, a beam emission spectroscopy diagnostic technique, is commonly used to measure electron density in the edge [78]. The principle of this diagnostic method is to inject a beam of alkaline metal atoms into the plasma, then these atoms are excited by collisions with electrons in plasma, and subsequently, the excited alkaline metal ions de-excite and emit photons. Since the intensity of the emitted photons is related to the plasma electron density, the electron density can be derived by measuring the photon emission. Figure 3.7 illustrates the schematic view of the alkaline metal beam system on W7-X, which consists of two parts: the alkali beam injector and a spectroscopy system. This system is located at a toroidal angle of $\varphi = 72°$ and $Z = 0.0$ m. it can provide electron density profile from far scrape-off layer to the edge plasma. The measurement radial range is limited by penetration depth of alkaline metal atoms. The temporal and spatial resolutions of this system are 20 μs and 0.5 cm, respectively.

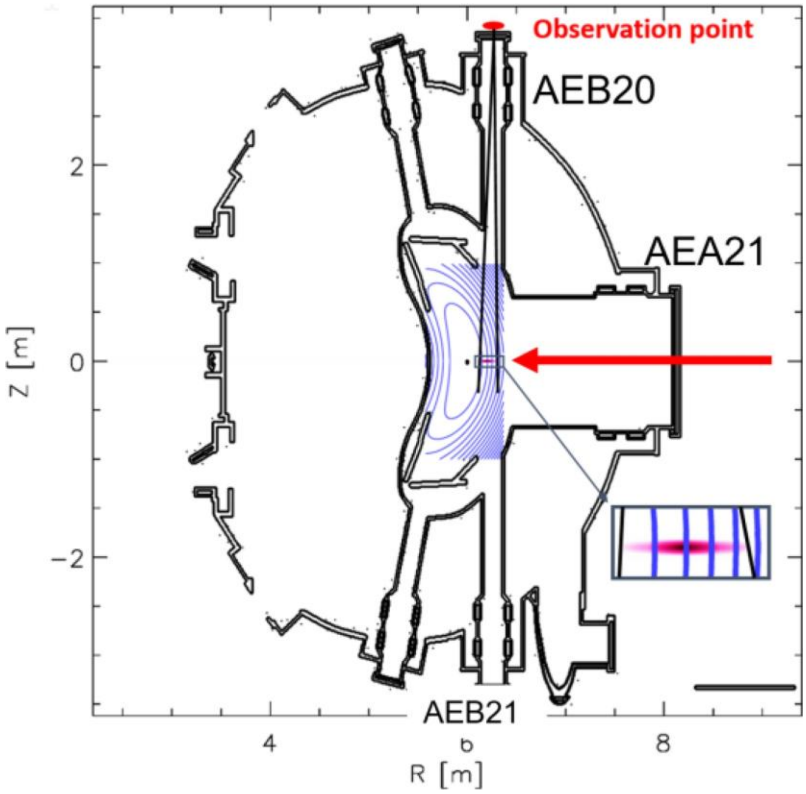

Figure 3.7: Overview of the alkali metal beam system on W7-X. Image from [78].

Divertor Thermography: An infrared (IR) camera is a fundamental diagnostic tool commonly used to measure the surface temperature and heat flux of divertor targets in fusion devices. This technique is based on the principles of black body radiation. Since the infrared radiation emitted by an object depends on its temperature, analyzing this radiation can provide surface temperature information. By understanding the temperature distribution on the divertor targets, it is able to derive the heat flux, considering the heat diffusion characteristics of the targets. The LOS of one IR camera system [79] on W7-X can cover the entire divertor region within a module. For example, Figure 3.8 (b) shows an image captured by the IR camera system. Additionally, all IR cameras collectively monitor 9 out of the 10 divertors. The spatial resolution can reach 3 - 10 mm, and the temporal resolution is approximately 10 μs.

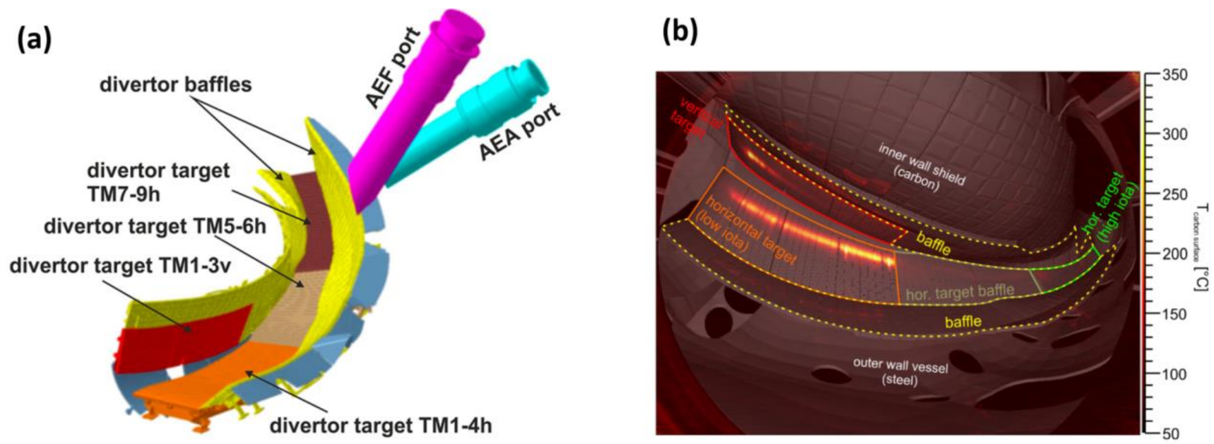

Figure 3.8: (a) Schematic view of the infrared camera system in one lower divertor on W7-X. (b) Infrared image for one divertor. Images adapted from [79].

Divertor Langmuir Probes: As mentioned in the previous section, Langmuir probes can infer electron density and temperature, ion saturation current, and plasma floating potential. On W7-X, divertor Langmuir probes are used to monitor particle and heat fluxes on the divertor, and they are essential tools for determining divertor detachment. In OP 2.1, with the replacement of test divertor units (TDUs) by water-cooled divertors, pop-up divertor Langmuir probes are equipped on W7-X [80, 81]. Figure 3.9 shows the locations of pop-up Langmuir probes on the lower divertor. The upper divertor has the same layout as the lower divertor. On both the lower and upper divertors, there are three layers of probes. Two layers (e.g., TM2h and TM3h) are nearly at the same toroidal angle ($\varphi = -82.4°$ for the lower divertor, $\varphi = -61.5°$ for the upper divertor), while another layer is on the high-iota tail ($\varphi = -56.8°$ for the lower divertor, $\varphi = -87.1°$ for the upper divertor). The temporal resolution of the divertor Langmuir probe system is limited to 1 ms.

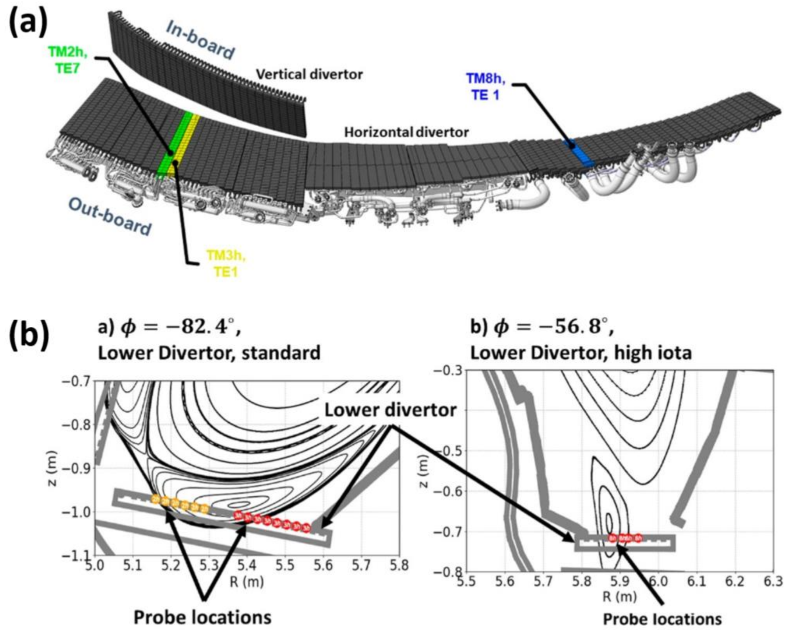

Figure 3.9: (a) Setup of the pop-up Langmuir probes on W7-X lower divertor. (b) measurement locations of the Langmuir probes on corresponding cross-sections. Image adapted from [80].

3.2 Synthetic edge diagnostics for EMC3-EIRENE

The EMC3-EIRENE output directly provides electron density, electron temperature, and particle and heat fluxes on targets, among other parameters. Therefore, for diagnostics that measure these parameters, which are directly available from EMC3-EIRENE simulations, the corresponding synthetic diagnostics are straightforward to develop. However, for synthetic line emissions, it is necessary to derive them not only from basic plasma parameters (including electron density, electron temperature, and impurity content) but also to account for the integration effect of the line of sight. In next section, the detailed process for building spectroscopic synthetic diagnostics are introduced.

3.2.1 Synthetic endoscope cameras

Before introducing the process of building synthetic spectroscopic diagnostics, it is essential to understand the line emissions. Line emission refers to the characteristic spectral lines of a specific atom or molecule, caused by the transition of electrons between energy levels. When an electron transitions from a higher energy level to a lower one, it emits a photon with energy equal to the difference between the two levels. In general, line emission is composed of three parts: excitation, recombination and charge exchange. Thus, the intensity of line emission can be expressed as:

$$\epsilon_{k\to l} = \sum_\sigma PEC^{(exc)}_{\sigma,k\to l} n_e n_\sigma^{z+} + \sum_\rho PEC^{(rec)}_{\rho,k\to l} n_e n_\sigma^{(z+1)+} + \sum_\rho PEC^{(CX)}_{\rho,k\to l} n_H n_\sigma^{(z+1)+} \ \text{(photons m}^{-3}\text{s}^{-1}) \quad 3.2.1$$

where PEC represents photon emissivity coefficient, σ metastable, n_e electron density, n_σ^{z+} density of the z-times ionized ion. Three types of PEC coefficients used in this work are from the generalized collisional-radiative model [82]. The viewing geometry of the divertor endoscope cameras is shown in Figure 3.1. Each camera is equipped with wavelength-selection and intensity filters to isolate the desired spectral line and adjust the light level. The relevant filter settings are listed in Table 3.1. At present the system can record the hydrogen Balmer series, C I (909.48 nm), C III (426.726 nm, 465.025 nm), He I (667.82 nm, 706.52 nm, 728.13 nm) and several molecular bands. These lines were chosen because they serve specific diagnostic purposes and have relatively large photon emissivity coefficients, which improves signal quality. Figure 3.10 gives one video frame from a W7-X discharge: (a) C III (465.0 nm) recorded by the camera in port AEL 30; (b) the same line recorded by the camera in port AEJ 51. The divertor structure is clearly visible in both images. To reconstruct the measured line emissions, the corresponding photon emissivity coefficient data are downloaded from open-ADAS. As an example, Figure 3.11 plots the excitation, recombination and charge exchange PECs for H_γ and CIII at a representative electron density of ~10^{19} m^{-3}. At electron temperatures above roughly 2-3 eV, the PECs for excitation and charge exchange are of comparable size, but the neutral density is already low, so the actual emissivity from charge exchange remains well below that from excitation. Excitation therefore dominates the signal. At lower temperatures recombination becomes the main contributor. Consequently, in attached plasmas on W7-X the H_γ and CIII lines are governed by excitation, whereas in detached conditions recombination plays the leading role.

Next, the calculation of plasma line emission radiance $L_{radiance}$ in EMC3-EIRENE modeling is presented. Figure 3.12 depicts a channel for endoscope cameras LOS. The LOS is assumed to be conical, resulting in a circular cross-section of the observation plane. The LOS is intersected by EMC3-EIRENE grids. Each atom intercepted by the LOS is regarded as a point emitter, so total radiant flux Φ collected by the spectrometer is

$$\Phi = \sum_i \Delta s_i A_i \sum_j \left(\frac{A_{ij}}{A_i}\frac{\sum_\sigma PEC^{(exc)}_{\sigma,k\to l} n_e n_\sigma^{z+} + \sum_\rho PEC^{(rec)}_{\rho,k\to l} n_e n_\sigma^{(z+1)+} + \sum_\rho PEC^{(CX)}_{\rho,k\to l} n_H n_\sigma^{(z+1)+}}{4\pi}\frac{A_s}{s_i^2}\right) \qquad 3.2.2$$

where A_i is the area of observation plane i, s_i is the distance between observation plane i and entrance slit, Δs_i is thickness of observation plane i, A_{ij} is the part of A_i that overlaps cell j, and A_s is the area of slit. Similar to the simplified optical path in Figure 3.2(b), the radiant flux Φ can be expressed as

$$\Phi = L_{radiance} A_p \Omega_s = L_{radiance} A_p \frac{A_s}{s^2} = L_{radiance} A_s \frac{A_p}{s^2} \qquad 3.2.3$$

where A_p is the detector area. Combining equation 3.2.2 and 3.2.3, we get

$$\sum_i \Delta s_i A_i \sum_j \left(\frac{A_{ij}}{A_i}\frac{\sum_\sigma PEC^{(exc)}_{\sigma,k\to l} n_e n_\sigma^{z+} + \sum_\rho PEC^{(rec)}_{\rho,k\to l} n_e n_\sigma^{(z+1)+} + \sum_\rho PEC^{(CX)}_{\rho,k\to l} n_H n_\sigma^{(z+1)+}}{4\pi}\frac{A_s}{s_i^2}\right) = L A_s \frac{A_p}{s^2} \qquad 3.2.4$$

Given that $\frac{A_i}{s_i^2} = \frac{A_p}{s^2} = \pi \sin(\Theta)^2$,

$$\sum_i \Delta s_i \cancel{A_s} \cancel{\frac{A_i}{s_i^2}} \sum_j \left(\frac{A_{ij}}{A_i}\frac{\sum_\sigma PEC^{(exc)}_{\sigma,k\to l} n_e n_\sigma^{z+} + \sum_\rho PEC^{(rec)}_{\rho,k\to l} n_e n_\sigma^{(z+1)+} + \sum_\rho PEC^{(CX)}_{\rho,k\to l} n_H n_\sigma^{(z+1)+}}{4\pi}\right) = L \cancel{A_s} \cancel{\frac{A_p}{s^2}} \qquad 3.2.5$$

So, the line emission radiance $L_{radiance}$ in the modeling is written as

$$L_{radiance} = \sum_i \Delta s_i \sum_j \left(\frac{A_{ij}}{A_i}\frac{\sum_\sigma PEC^{(exc)}_{\sigma,k\to l} n_e n_\sigma^{z+} + \sum_\rho PEC^{(rec)}_{\rho,k\to l} n_e n_\sigma^{(z+1)+} + \sum_\rho PEC^{(CX)}_{\rho,k\to l} n_H n_\sigma^{(z+1)+}}{4\pi}\right) \qquad 3.2.6$$

Because an exact geometric overlap is hard to obtain, a Monte-Carlo approach is used. Random sample points are generated inside slice i, and the weight for grid cell j is defined as

$$\omega_{ij} = \frac{A_{ij}}{A_i} = \frac{N_{ij}}{N_i} \qquad 3.2.7$$

where N_{ij} is the number of points that fall in cell j and N_i is the total number of points in the slice i. Using these weights, the LOS radiance $L_{radiance}$ is

$$L_{radiance} = \sum_i \Delta s_i \sum_j \left(\omega_{ij}\frac{\sum_\sigma PEC^{(exc)}_{\sigma,k\to l} n_e n_\sigma^{z+} + \sum_\rho PEC^{(rec)}_{\rho,k\to l} n_e n_\sigma^{(z+1)+} + \sum_\rho PEC^{(CX)}_{\rho,k\to l} n_H n_\sigma^{(z+1)+}}{4\pi}\right) \qquad 3.2.8$$

With a known LOS of diagnostics, the line emission radiance $L_{radiance}$ (photons $\mathrm{cm^{-2}s^{-1}sr^{-1}}$) can be derive from EMC3-EIRENE outputs (electron density, electron temperature and impurity densities).

Table 3.1: Filter settings of the divertor endoscope cameras.

Allocation of the Installed Filter Wheels							
Endoscope	Branch/Camera	IF wheel	λ/nm	atom, molecule	Camera I_{cam}/I_{total}	intensity filter T	fibre I_{fibre}/I_{total}
AEL30 AEJ30 AEL51 AEJ51	**1A (UV) / Cam1**	1	no filter		35%	70%	15%
		2	410.1738	H I (δ)	35%	70%	
		3	426.726	C III	35%	70%	
		4	430	CD	35%	70%	
		5	434.0466	H I (γ)	35%	70%	
		6	tbd	tbd	35%	70%	
	2A (UV) / Cam2	1	no filter		25%	50%	25%
		2	465.025	C III	25%	50%	
		3	486.1327	H I (β)	25%	50%	
		4	516	C_2	25%	50%	
		5	tbd	tbd	25%	50%	
		6	tbd	tbd	25%	50%	
	1A (vis) / Cam3	1	no filter		25%	50%	/
		2	601.5	D_2	25%	50%	
		3	656.1	H I (α)	25%	50%	
		4	706.52	He I	25%	50%	
		5	tbd	tbd	25%	50%	
		6	tbd	tbd	25%	50%	
	2B (vis) / Cam4	1	no filter		35%	70%	15%
		2	728.13	He I	35%	70%	
		3	909.48	C I	35%	70%	
		4	tbd	tbd	35%	70%	
		5	tbd	tbd	35%	70%	
		6	tbd	tbd	35%	70%	
	1B (vis) / Cam5	1	no filter		25%	/	/
		2	667.82	He I	25%	/	
		3	tbd	tbd	25%	/	
		4	tbd	tbd	25%	/	
		5	tbd	tbd	25%	/	
		6	tbd	tbd	25%	/	
	1A (IR) / MIR		na	na	50%	tbd	/
	2A (IR) / MIR		na	na	/	tbd	50%
							Stand: 13.11.2020

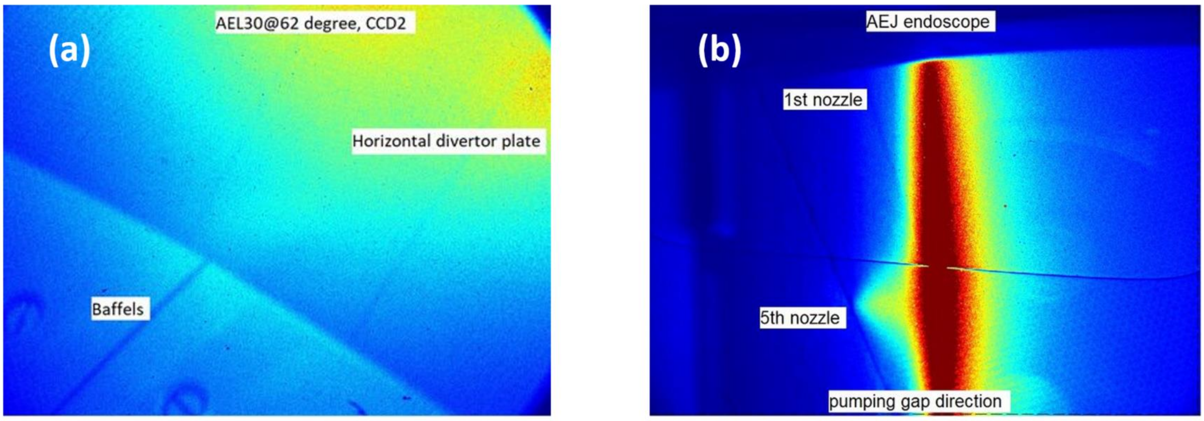

Figure 3.10: Sample C III (465 nm) images from the divertor endoscope system (a) Port AEL 30; (b) Port AEJ 51.

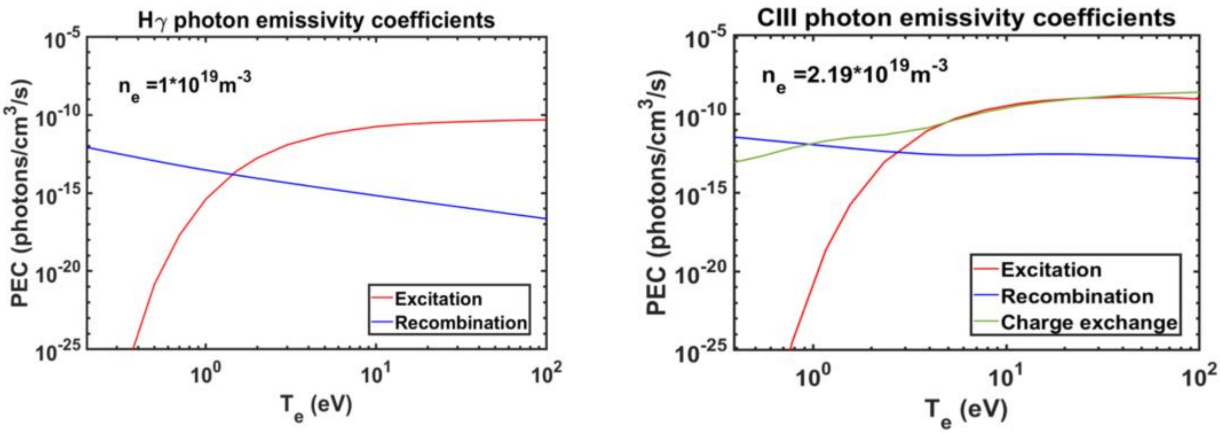

Figure 3.11: Excitation, recombination and charge exchange photon emissivity coefficients (PEC) as a function of electronic temperature under a given electron density for H_γ (434.0466 nm) and CIII (465.025 nm).

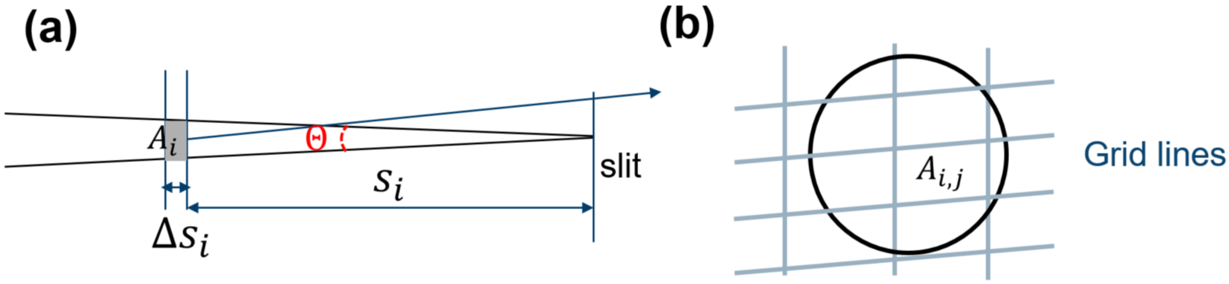

Figure 3.12: (a) Geometry of a single LOS. The circular patch A_i lies a distance s_i from the entrance slit; its thickness along the LOS is Δs_i ; θ is view angle. (b) cross-section of observation plane i. A_{ij} is overlap area between slice i and cell j.

3.2.2 Synthetic divertor spectroscopy

An example spectrum taken with the SP-2750 spectrometer during discharge #20180919.025 is plotted in Figure 3.13. In routine operation the CII (426.83 nm), CIII (465.025 nm) and hydrogen Balmer lines are selected. The radiance L seen by each divertor spectroscopy channel is evaluated with exactly the same formula derived for the endoscope cameras (equation 3.2.8). The procedure is identical except that the spectrometer records a one-dimensional chord-integrated signal, whereas the endoscope cameras provide a two-dimensional image.

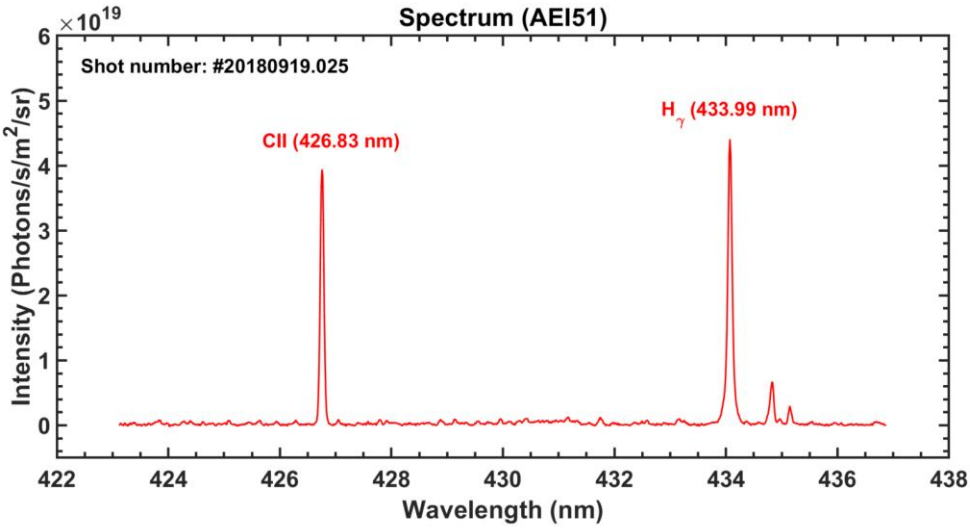

Figure 3.13: Spectrum of one horizontal spectroscopy channel.

3.2.3 Other synthetic diagnostics

As mentioned in Chapter 2, EMC3-EIRENE modeling can provide many useful outputs, for example, electron density, electron temperature, and particle and heat fluxes on targets, and so on. If the coordinates of a diagnostic are known, the code interpolates its results to those locations, producing synthetic signals for direct comparison with measurements from diagnostics like probes, Langmuir probes, IR cameras, or Thomson scattering systems.

3.3 Summary

The focus of this chapter was two-part. First, it presents the main edge diagnostics on W7-X that can be compared with EMC3-EIRENE, such as the divertor spectroscopy endoscope, edge and divertor spectroscopy, Thomson scattering, divertor thermography, divertor Langmuir probes, and others. Calibration procedures for the divertor endoscope and the divertor spectrometers are described in detail. Second, it develops synthetic counterparts of these diagnostics in EMC3-EIRENE post-processing, with particular emphasis on the original method introduced here for reconstructing synthetic spectroscopic signals.

4 EMC3-EIRENE modelling of edge impurity transport constrained by multiple edge diagnostics

A critical challenge in high-power, long-pulse operation of fusion devices is preventing excessive heat loads on plasma-facing components, which can significantly reduce their lifetime [35, 83, 84]. Impurities in the plasma edge can dissipate energy through radiation before the plasma reaches the target plate, helping to reduce heat loads [85-87]. However, excessive impurity penetration into the plasma core can degrade confinement, leading to thermal instabilities such as multifaceted asymmetric radiation from the edge (MARFE) [36-38, 88], and can even cause radiation disruption [39, 89, 90]. W7-X employs a three-dimensional island divertor to control particle and heat exhaust; its complex 3D geometry makes impurity analysis particularly challenging.

The 3D edge plasma transport code EMC3-EIRENE [54, 91] has been widely used to study impurity transport in fusion devices [60-63, 92-95], with extensive applications to W7-X [44, 46, 48, 96, 97]. For quantitative analysis [98] with EMC3-EIRENE, the code inputs must be consistent with the measurements. This is especially true for the cross-field particle and heat diffusivities: EMC3-EIRENE represents anomalous transport with these user-defined coefficients, so their values must be fixed by measurements. This chapter constrains the EMC3-EIRENE inputs with a wide set of upstream and downstream diagnostics. To that end, synthetic reconstructions were added to the post-processing for all relevant diagnostics—most importantly the divertor spectroscopic signals.

The chapter is organized as follows. Section 4.1 describes how the edge diagnostics are used to adjust the cross-field transport coefficients and other inputs, and demonstrates that the optimized EMC3-EIRENE simulations reproduce the experimental data to a satisfactory degree. Section 4.2 examines the evolution of the radiation front during a density ramp, highlighting the similarities and discrepancies between measurement and modelling. Section 4.3 investigates how changes in the 3D magnetic equilibrium, such as β-induced deformation and toroidal current, modify edge transport.

4.1 Optimizing EMC3-EIRENE input parameters using edge diagnostics

EMC3-EIRENE requires several key input parameters, among them: input heating power across the innermost boundary surface (P_{in}), radiation power fraction (f_{rad}), electron density at the separatrix ($n_{e,sep}$), cross-field particle diffusivity $D_{\perp}$, cross-field thermal diffusivity $\chi_{\perp}$

and other quantities. The first three parameters (P_{in}, f_{rad}, $n_{e,sep}$) can be limited directly by experimental measurements. In contrast, the cross-field transport coefficients $D_\perp$ and $\chi_\perp$ cannot be inferred from diagnostics alone; they must be tuned iteratively. The usual procedure is to run EMC3-EIRENE with trial values, generate synthetic diagnostic signals, compare them with the measurements, and then adjust $D_\perp$ and $\chi_\perp$ until good agreement is obtained. The following sections show, step by step, how the experimental data are used to constrain inputs and discuss the transport-coefficient settings that best reproduce the observations.

4.1.1 Experimental data and modeling setup

W7-X is a highly optimized device [99-101] equipped with an island divertor [102, 103] to control particle and heat exhaust, as well as plasma-wall interactions. It is designed to achieve steady-state operation with a high heating power (10 MW) and reactor-relevant plasma parameters for pulse lengths up to 30 minutes [100]. By adjusting the currents in the planar and non-planar toroidal coils, W7-X can produce various magnetic configurations [23], characterized by the rotational transform $\iota_a = n/m$, where n and m represent the toroidal and poloidal mode numbers, respectively. In this work, the standard configuration with an edge rotational transform of 5/5 is chosen.

A typical experimental discharge, including both attached and detached plasmas, is selected to better study impurity transport under different divertor operating regimes. Figure 4.1 shows the time evolution of a typical density ramp-up detachment discharge (#20181010.026). As the line-integrated electron density, measured by reflectometry, increases, a transition from attached to detached plasma occurs at approximately 1.25 seconds, indicated by the rollover of the divertor ion saturation current. To study the transport properties of the attached and detached plasma, we select two time points: $t_1 = 1.2$ s (attachment) and $t_2 = 1.8$ s (detachment), marked by the vertical magenta and blue dashed lines, respectively. At the attached condition ($t_1 = 1.2$ s), the ECRH heating power is 5.3 MW, while the total radiation power detected by the bolometer [104] is approximately 1.4 WM. The radiation fraction $f_{rad} = P_{rad}/P_{ECRH} \approx 26\%$, and the mean effective ion charge derived from bremsstrahlung is low at 1.2. In the detached regime ($t_2 = 1.8$ s), the total radiation power increases significantly to 5.0 MW, with the heating power remaining constant. Consequently, the radiation fraction rises to 94%, and the mean effective ion charge reaches 1.6. After detachment, the average parallel ion saturation current decreases continuously from 11.2 to 3.3 $\mathrm{Acm^{-2}}$. The toroidal plasma current attained from the in-vessel Rogowski coil remains below 1 kA, so toroidal plasma current effects are not considered in simulations. To determine the input parameters for modelling, especially the cross-field transport coefficients, several relevant edge diagnostics are selected. These include: (1) divertor spectroscopy, which detects CIII line emission (465.01 nm) in the divertor region; (2) Thomson scattering (TS) offering radial electron density and temperature measurements; (3) divertor Langmuir probes (LPs), which measure ion saturation currents on the divertor; (4) infrared cameras capturing heat flux data on nine divertors. The locations of these diagnostics in W7-X are shown in Figure 4.2. In the EMC3-EIRENE post-processing, synthetic signals are generated for each of the selected diagnostics according to the methods outlined in Section 3.2.

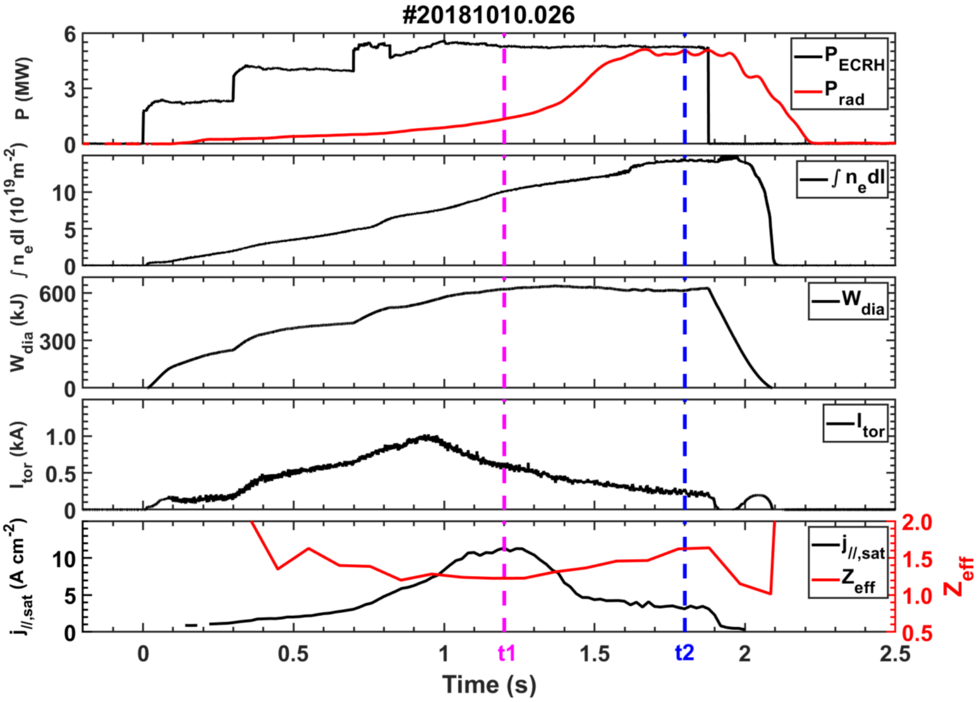

Figure 4.1: Time evolution of key W7-X parameters during a typical density ramp-up detachment discharge (#20181010.026). Dashed lines indicate the two selected modeling times.

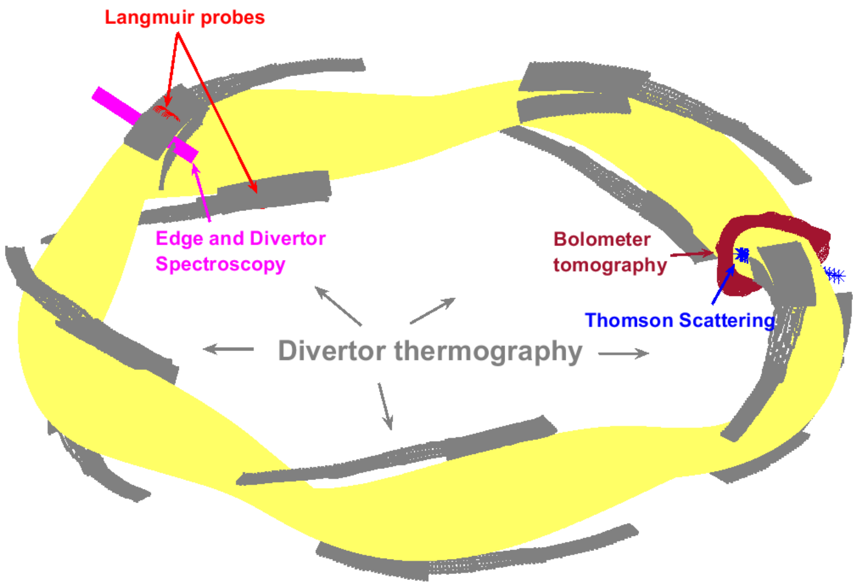

Figure 4.2: The diagram of relevant diagnostic positions and a magnetic surface of standard configuration in W7-X

In the following simulations, the power flux across the innermost boundary surface of the simulated domain is set to 4.8 MW, equally split between electrons and ions. The upstream density at the separatrix is set to 5.0×10^{19} m^{-3} for the attachment case and 7.0×10^{19} m^{-3} for the detachment case, based on experimental measurements. Carbon particles sputtered from the divertor and baffle targets are the only radiators, and the carbon source is controlled by total radiation power. According to experimental data, the carbon radiation fraction is set to 29% for attached plasma and 94% for detached plasma, respectively. Additionally, the Bohm sheath boundary condition is applied to the particle and heat flux at the target plates. Figure 4.3 shows a schematic of the EMC3 and EIRENE computational domains, together with the locations of the divertor targets and the baffle, in a bean-shaped cross-section.

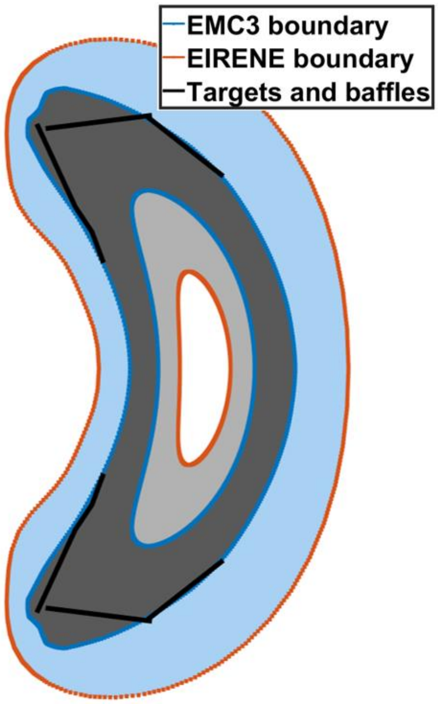

Figure 4.3: Schematic of the EMC3 and EIRENE computational domains at a bean-shaped cross-section. The blue contours outline the EMC3 boundary, the orange contours outline the larger EIRENE boundary, and the black segments indicate the divertor targets and baffles.

For the input-free parameters in EMC3-EIRENE, anomalous cross-field transport coefficients for particles ($D_{\perp}$) and energy ($\chi_{\perp}$) are adjusted to match the measurement profiles mentioned in Section 2. Typically, cross-field transport coefficients are set to be spatially constant and no distinction is made between background plasma and impurities in the simulations. However, with the development of a synthetic diagnostic for the divertor spectroscopy system, we are able to provide diagnostic validation for simulations with spatially non-uniform transport coefficients as well as simulations distinguishing between background plasma and impurity transport coefficients.

Some combinations of simulation input parameters are shown in Table 4.1. Cases 1-4 correspond to attached plasma simulations, while cases 5-8 represent detached plasma conditions. For cases 3-4 and 7-8, which use non-uniform transport coefficients, the distributions of $D_{\perp}$ and $\chi_{\perp}$ in the bean shape cross-section are illustrated in Figure 4.4. The transport coefficient in simulation domain is divided into three regions: the scrape-off layer, remnant magnetic islands and area inside the last closed flux surface (LCFS). Currently, this division is simply based on the relationship with the connection length. In the future, a more reasonable transport coefficient distribution will need to be developed by incorporating more edge diagnostics. The standard magnetic configuration considers only the vacuum field, without accounting for plasma beta effect.

Table 4.1: Radiation fraction, separatrix density and cross-field transport coefficients for different modeling cases.

Case	f_{rad}	$n_{e,sep}$ (10^{19} m^{-3})	$D_{\perp,e}$ (m^2/s)	$D_{\perp,im}$ (m^2/s)	$\chi_{\perp,e} = \chi_{\perp,i}$ (m^2/s)
Case 1	26%	5	0.5	0.5	1.5
Case 2	26%	5	1.0	1.0	0.75
Case 3	26%	5	Figure 4.4 (a)	figure 4.4 (a)	figure 4.4 (c)
Case 4	26%	5	Figure 4.4 (a)	figure 4.4 (b)	figure 4.4 (c)
Case 5	94%	7	0.5	0.5	1.5
Case 6	94%	7	1.0	1.0	0.75
Case 7	94%	7	Figure 4.4 (a)	figure 4.4 (a)	figure 4.4 (d)
Case 8	94%	7	figure 4.4 (a)	figure 4.4 (b)	figure 4.4 (d)

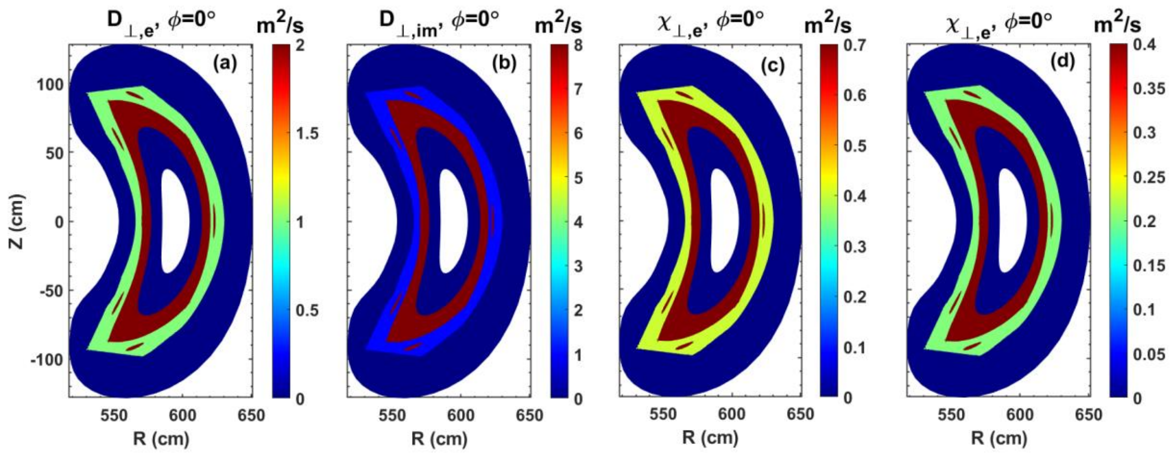

Figure 4.4: Distributions of the cross-field transport coefficients in simulations.

4.1.2 Quantitative comparisons of simulation results and experimental measurements

The attached and detached plasma properties are simulated using EMC3-EIRENE code with the input parameters listed in Table 4.1, and the simulation results are subsequently compared with experimental measurements. Figure 4.5 and Figure 4.6 show the upstream Thomson scattering electron density and temperature profiles along the observation path depicted in Figure 4.2 for both attached and detached plasmas, along with the corresponding simulation data. As seen in these figures, the current TS system is not well-suited for edge comparisons due to incomplete coverage on the inner-side observation path and insufficient resolution in the outer-side channels. In the attached plasma regime (Figure 4.5), the experimental results are roughly consistent with all simulation cases listed in Table 4.1 within the 97.5% confidence interval of the experimental measurements, except Case 1. This suggests that during attachment, the cross-field particle transport coefficient ranges between 0.5 and 2.0 m²/s, while the thermal diffusivity is around 0.75 m²/s. For the detached plasma shown in Figure 4.6, the TS diagnostics provide limited edge measurements. Temperature measurements outside the LCFS are unreliable due to extremely low values in that region. Regarding the density, it is difficult for the simulation results to match the measurements simultaneously both inside and outside the LCFS. Overall, a more precise estimate of the transport coefficients would require the additional diagnostics.

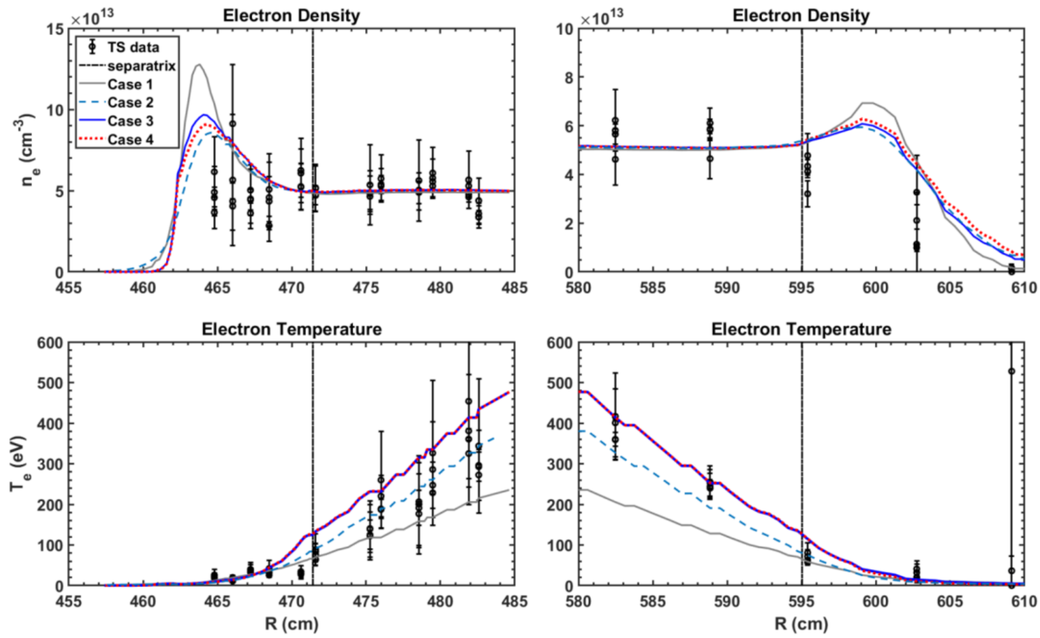

Figure 4.5: Electron density and temperature profiles from TS measurements and four EMC3-EIRENE simulation cases for attached plasma. The case numbers in the legend correspond to those listed in Table 4.1.

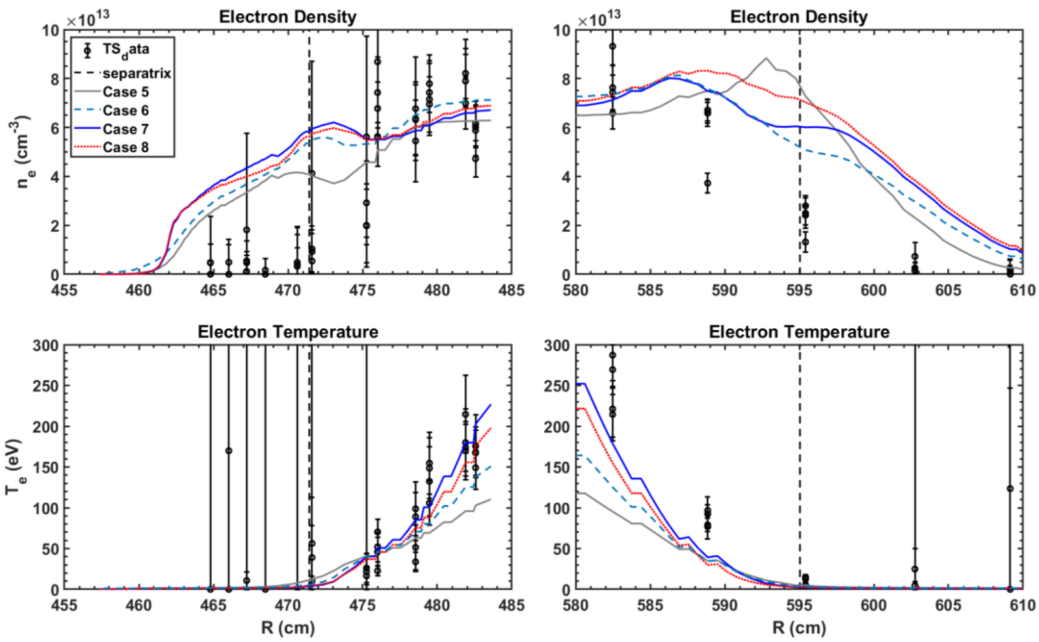

Figure 4.6: Similar to Figure 4.5, but for detached plasma.

Shifting our attention to the downstream data comparison, we begin by analyzing the heat flux measurements on the target plates from the nine infrared cameras and comparing them with the EMC3-EIRENE simulation results. By mapping the heat flux from experimental data and simulations onto the cross-section corresponding to the divertor target probes, Figure 4.7 illustrates the heat flux distribution in both attached and detached regimes. The gray dashed lines indicate the heat flux measurements from the upper and lower divertor plates obtained by infrared cameras, while the black solid line represents the average of these measurements. The colored curves correspond to the simulation results. For the attached plasma, the heat flux measurements are asymmetric—not only between the upper and lower divertor plates but also between divertor target plates at different toroidal modules. This asymmetry is likely caused by error fields and drift effects that are not considered in the EMC3-EIRENE simulations. When comparing the averaged heat flux measurements with the simulation results, the peak position of the measured heat flux aligns closely with the simulations, though the experimental profile is narrower. In contrast, for the detached plasma, the heat flux on the target plate is reduced to the background level of the infrared camera measurements, and the asymmetry observed in the attached regime disappears. Comparisons between experimental measurements and simulation results lose their meaning. Figure 4.8 displays more comprehensively the averaged experimental measurements from nine cameras and simulation results across the entire target plates for attached plasma.

It is obvious that the heat flux width in the experimental measurements is more peaked than in the simulations. However, adjusting the cross-field transport coefficients can reduce the heat flux width in the simulations. Compared to Case 1, Case 2 reduces the heat flux width on both the horizontal and vertical targets by lowering the cross-field thermal diffusivity coefficients. Cases 3 and 4 apply spatially non-uniform cross-field transport coefficients, further reducing thermal diffusivity coefficients in the SOL, resulting in better agreement between the simulations and experimental measurements.

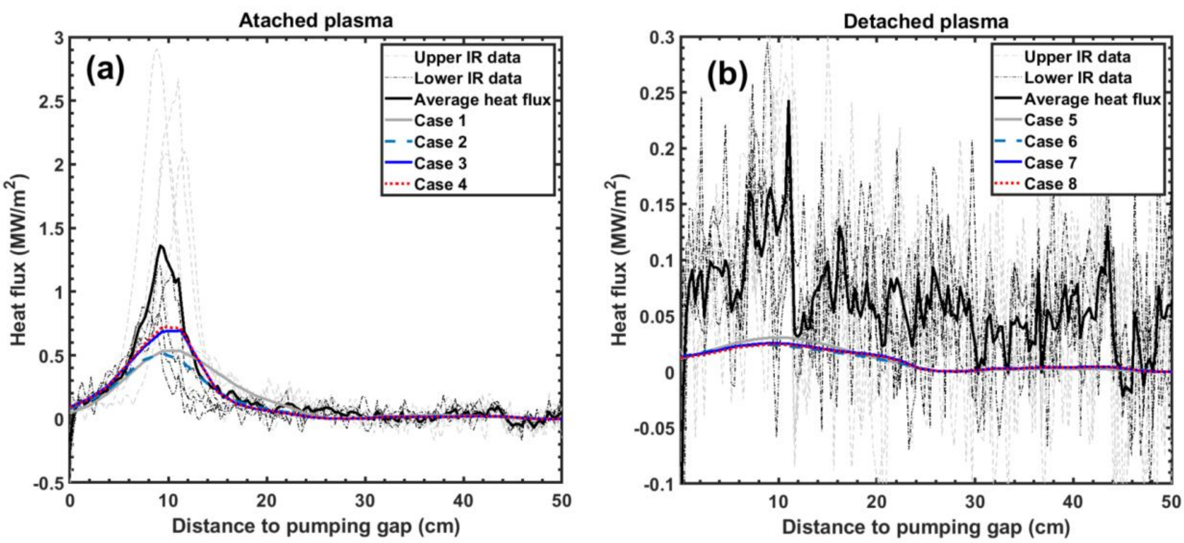

Figure 4.7: Comparison of nine IR-cameras (dotted) and simulated (solid) heat flux profiles at the divertor probe location under conditions of (a) attachment and (b) detachment.

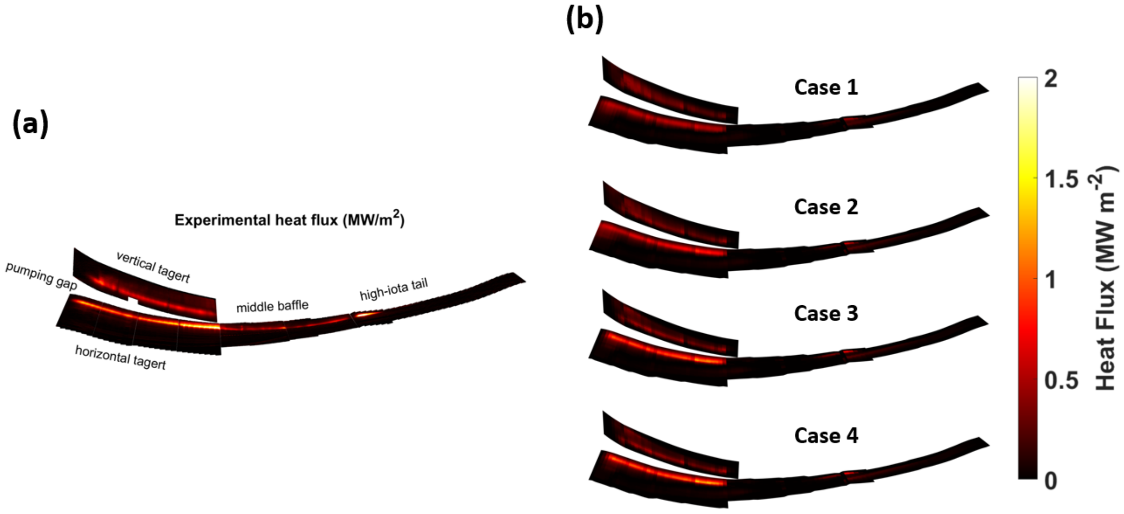

Figure 4.8: Average experimental and simulated heat flux distributions for attached plasma.

The particle flux on the upper and lower horizontal divertor plates is measured using the divertor Langmuir probes (LPs) shown in Figure 4.2. Figure 4.9 illustrates the parallel ion saturation currents obtained from both experimental measurements and EMC3-EIRENE simulations for attached and detached plasmas. From the attached to the detached regime, the parallel ion saturation currents decrease by a factor of about five. In the attached regime, the ion saturation currents measured on the upper and lower divertor plates show a clear asymmetry, with the values on the two plates differing by approximately a factor of two. However, this asymmetry decreases significantly in the detached regime, which is consistent with the trends observed in the heat flux measurements. As the measurement region does not fully cover the peak profile of the particle flux, the current divertor LPs system has limited ability to optimize the simulation input parameters. As shown in Figure 4.9, reducing the thermal diffusivity sharpens the ion saturation currents profiles in the simulations, but the changes within the LPs measurement region remain small.

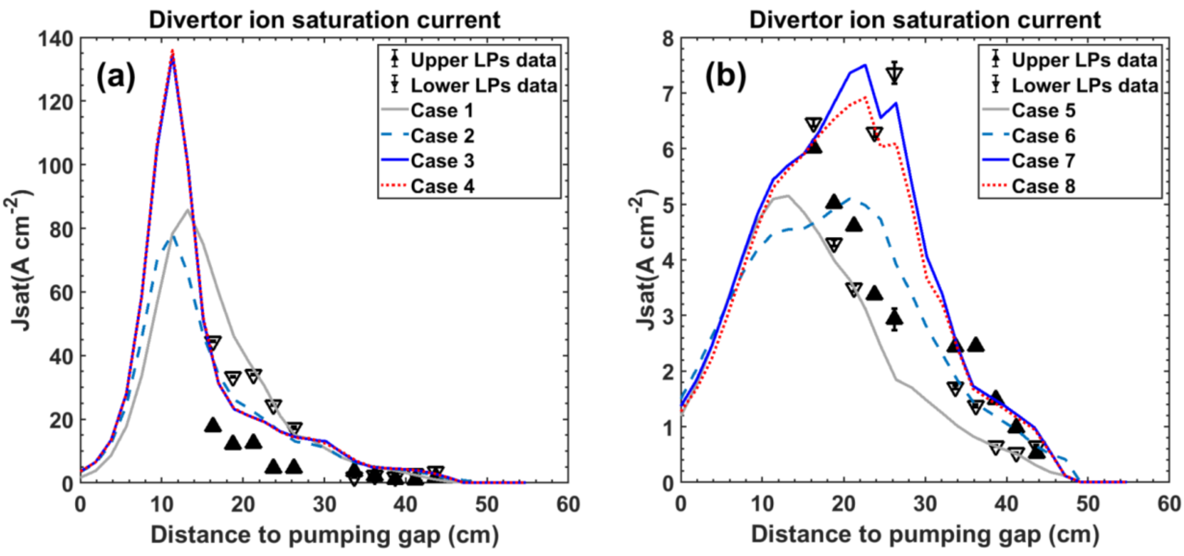

Figure 4.9: The parallel ion saturation currents from experimental measurements and EMC3-EIRENE simulations for (a) attached and (b) detached plasmas.

The comparison between experimental measurements and simulation results shows that the above experimental measurements do not provide a sufficiently accurate assessment of the simulation input parameters, especially regarding the cross-field transport coefficient of the impurity. For example, the differences between simulation cases 2, 3, and 4, as well as cases 6, 7, and 8, are difficult to distinguish based on these above diagnostic measurements. However, the synthetic divertor spectroscopy developed in this work shows high sensitivity to variations in the cross-field transport coefficients of both the background plasma and impurities in the simulations. This sensitivity arises from the dependence of impurity line emissions on the distributions of plasma electron temperature, electron density, and impurity ion density, as illustrated in Equation 3.2.8. To further investigate these transport mechanisms, Figure 4.10 presents the distribution of CIII radiance above the horizontal target, comparing experimental measurements with simulations for both attached and detached plasmas. The experimental measurements show that the peak of the CIII radiance shifts further away from the horizontal target during the transition from attachment to detachment. This shift occurs for two main reasons: first, during detachment, the electron temperature near the target drops significantly, which prevents impurities from ionizing and allows them to penetrate deeper into the plasma. Second, the temperature gradient near the target becomes steeper during detachment, causing a stronger thermal gradient force that drives impurities upstream. In this work, by optimizing cross-field transport coefficients, EMC3-EIRENE successfully reproduces this process in the simulations.

In Figure 4.10(a), the experimental measurements (black dashed line with circles) show a primary CIII radiance peak located about 5.5 cm from the horizontal target, with a smaller secondary peak around 2 cm. In the simulations, Case 1 shows a lower radiance peak compared to the experimental data, and the peak position is closer to the target. By increasing the particle transport coefficient and reducing the thermal transport coefficients, Case 2 achieved peak positioning closer to the experimental data, but the peak intensity remains insufficient. Further optimization using spatially non-uniform transport coefficients to enhance impurity transport within the magnetic island, allows Cases 3 and 4 to nearly match the experimental peak positions. Case 4, which enhanced impurity transport coefficients within magnetic islands, demonstrated peak intensities remarkably consistent with experimental values. However, the peak width in Case 4 is narrower than observed in the experimental data, and the second smaller peak at 2 cm is not reproduced. Additionally, beyond 8 cm, the simulated CIII radiance in Case 4 is slightly higher than the experimental results. These discrepancies may be due to the coarse setting of the spatially non-uniform transport coefficients. The narrower peak width in Case 4 suggests that the region with enhanced transport coefficients within the magnetic island needs to be extended. The reasons for the failure to reproduce the secondary peak at 2 cm in Case 4, and the slightly higher CIII radiance beyond 8 cm will be discussed further in the following sections. For the detached plasma case (Figure 4.10(b)), the experimental measurements show the radiance peak penetrates further into the plasma, with the peak position near 12 cm, and a broader peak width compared to the attached plasma case. In Case 8, by applying spatially non-uniform cross-field transport coefficients and increasing impurity transport, the simulation achieves a good match with the experimental measurements in terms of peak position, width, and intensity.

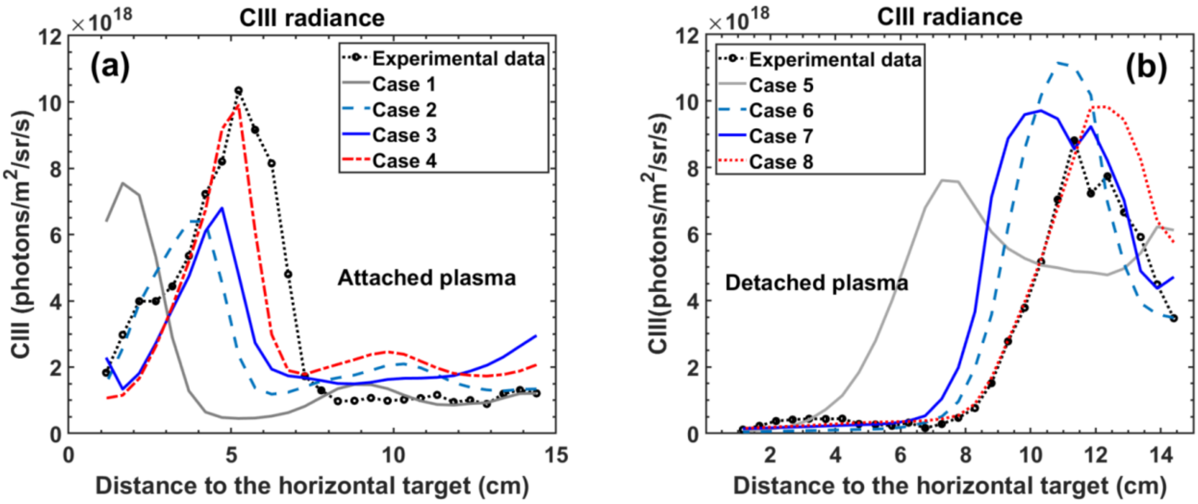

Figure 4.10: The distribution of CIII radiances above the horizontal target from experimental measurements and simulations in (a) attached and (b) detached plasma.

To more intuitively demonstrate the agreement between the simulation results and experimental measurements, a quantitative method is needed. Since this study involves comparing multiple diagnostic measurements with simulation results, it is necessary to define a dimensionless parameter to describe the quality of the match. The dimensionless factor used here is expressed as a weighted average:

$$W = \frac{\sum_{i=1}^{n} \omega_i R_i^2}{\sum_{i=1}^{n} \omega_i} \quad 4.1.1$$

where ω_i represents the weight of the i-th diagnostic measurement, R_i^2 is the R-squared value for the i-th diagnostic measurement, n is the number of diagnostic system used here. The R-squared R_i^2 is defined as:

$$R_i^2 = 1 - \frac{\sum_j (Y_j - X_j)^2}{\sum_j (Y_j - \bar{Y})^2} \quad 4.1.2$$

where Y_j denotes the experimental measurement at the j-th point for the i-th diagnostic, X_j is the corresponding simulation data, and $\bar{Y}$ is the average of the experimental measurements for the i-th diagnostic.

For attached plasmas, due to the large error bars in the TS data near the edge, we assign it a weight of 0.1, while other diagnostics are given a weight of 1. For detached plasmas, since the particle flux and heat flux measurements on targets approach background levels, their weights are also set to 0.1, with the TS data remaining at 0.1 and all other diagnostics assigned a weight of 1. According to the definition, the closer the weighted average is to 1, the better the match between the experimental and simulation results. Table 4.2 presents the weighted averages for the eight simulation cases. It shows that for the attached plasma, Case 4 is the closest to 1, indicating the best match. For the detached plasma, Case 8 shows the best match. Case 4 and

Case 8 nearly reproduce the transition from the attached to the detached regime as observed in the diagnostics presented in this study.

Table 4.2: The weighted averages for the eight simulation cases.

	Attachment				**Detachment**			
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8
W	-0.31	0.53	0.64	0.76	-7.27	0.61	0.40	0.76

We use the attached plasma as an example to discuss the necessity of applying spatially non-uniform cross-field transport coefficients and enhancing impurity cross-field transport coefficients within magnetic islands in simulations. A new simulation, Case 9, is introduced, differing from Case 2 only by increasing the uniform impurity cross-field transport coefficient to 8.0 m^2/s. Figure 4.11 shows the CIII radiance profiles and the CIII line emission distribution in the cross-section where the divertor spectroscopy diagnostics are located for Cases 1, 3, 4, and the newly added Case 9. As described in Equation 5, the CIII radiance is obtained by integrating the CIII line emission along the line of sight. In Case 9, the CIII radiance closely matches the experimental measurements in terms of peak height, position, and width, and successfully reproduces the small secondary peak observed around 2 cm from the target in the experimental data. This confirms our earlier assumption that the region of enhanced transport coefficients within the magnetic island in Case 4 needs to be expanded. However, Case 9 significantly overestimates the CIII radiance profile beyond 8 cm from the target. In Figure 4.11(b), a comparison of the CIII line emission distributions for Cases 1, 4, and 9 helps explain the differences observed in their respective CIII radiance profiles. In the CIII line emission distribution, we focus on three specific regions, marked with red, orange, and green circles, referred to as Areas 1, 2, and 3. These regions correspond to the main peak, secondary peak, and the region beyond 8 cm from the target in the CIII radiance profiles. Compared to Case 1, the spatially non-uniform transport coefficients used in Case 4 enhance impurity transport within the magnetic island, resulting in stronger CIII line emission in Area 1 and shifting the emission peak further from the target. This occurs because the increased impurity transport allows C^{2+} ions to penetrate more deeply into the magnetic island. However, the enhanced impurity transport in Area 3 also increases the CIII line emission there, leading to slightly higher CIII radiance beyond 8 cm compared to experimental measurements. In Case 9, the higher, spatially uniform impurity cross-field transport coefficients significantly increase the CIII line emission across all regions, helping to reproduce the main and secondary peaks observed in the experimental data. However, this setting also greatly increases the CIII line emission in Area 3. In conclusion, adopting spatially non-uniform transport coefficients, as demonstrated in Case 4, combines the advantages of both Cases 1 and 9, resulting in a closer match with the experimental data. However, further refinement of Case 4 is necessary to achieve better agreement with the experimental results.

To better match the CIII measurements, it was found that impurity cross-field transport coefficients within the magnetic islands needs to be several times higher than the commonly used range of 0.5-1.5 m^2/s in W7-X simulations. As shown in Figure 4.11(a), compared to Case 3, Case 4 increase the impurity cross-field transport coefficients in remnant islands to be 8.0 m^2/s, causing the CIII radiation to move further away from the horizontal target and increase in intensity, better matching the experimental measurements. The enhanced transport in Case 4 increases impurity penetration, shifting the CIII emission farther from the target.

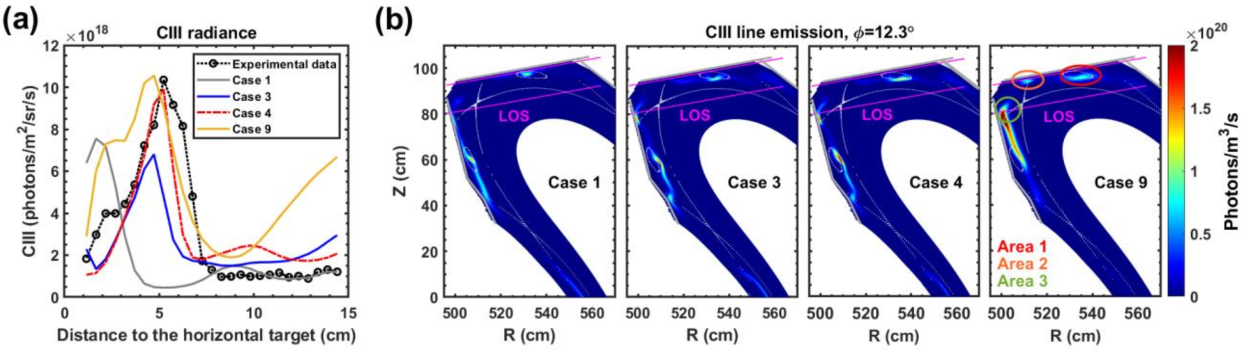

Figure 4.11: (a) The CIII radiance profile and (b) the CIII line emission distribution in the cross-section where the divertor spectroscopy diagnostics are located, for Cases 1, 3, 4, 9.

Figure 4.12 compares experimental measurements and simulated results of the radiation distribution in a triangular cross-section for both attached and detached plasmas. The experimental measurements (top panels) exhibit distinct characteristics under different plasma conditions. For the attached plasma (top-left), the radiation region is primarily located outside the LCFS and concentrated near the O-point of the magnetic island. In contrast, for the detached plasma (top-right), the radiation region extends inside the LCFS and shifts toward the X-point of the magnetic island. Both attached and detached cases in the experimental data display significant up-down asymmetry, which may be caused by error fields and drift effects.

The simulated results (bottom panels) are compared with the experimental data. For the attached plasma (bottom-left, Case 4), the simulation captures the general radiation distribution, with the primary radiation region located outside the LCFS. However, it fails to reproduce the pronounced up-down asymmetry observed in the experiment and significantly overestimates the radiation inside the right-hand side of the magnetic island. For the detached plasma (bottom-right, Case 8), the simulation shows the radiation region extending inside the LCFS, which is consistent with experimental observations. However, the penetration depth of the radiation in the simulation exceeds that observed in the experiment. Furthermore, the finer details of the up-down asymmetry visible in the experimental data are not captured in the simulation. These comparisons suggest that certain factors contributing to the asymmetry, such as the drift effects and error fields, may not be included in the current simulation.

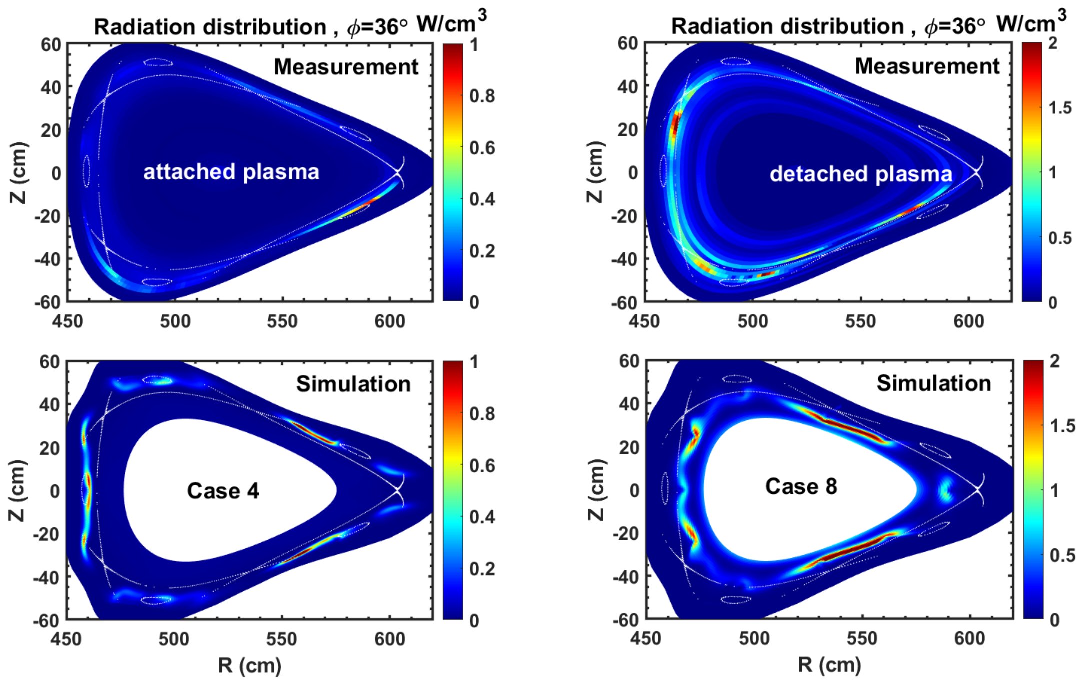

Figure 4.12: Comparison of experimental and simulated radiation distribution at the triangular cross-section for attached and detached plasmas.

Overall, at low density the drift effects are pronounced, so the simulation-measurement agreement is unsatisfactory; however, if profiles from different up/down or toroidal positions are mapped and then averaged, the agreement becomes acceptable. At higher density, the drift effects weaken, and the direct agreement between simulation and diagnostics correspondingly improves.

4.2 Radiation front migration with rising radiation fraction

As reported by Krychowiak [71], in the W7-X standard configuration the experimentally observed C II (712 nm) and C III (674 nm) line radiation shifts from the strike point toward the magnetic X-point as the radiated-power fraction f_{rad} increases. The simulations reproduce this overall migration, but they also predict a non-negligible localized region of C II and C III emission near the O-point during the transition—an intensity that is absent in the measurements. The divertor spectroscopy system (Figure 4.13) provides horizontal and vertical viewing channels across the upper divertor at the toroidal angle $\varphi = 12.3°$; the vertical view, in particular, allows direct discrimination between X-point and O-point line emission. Accordingly, the synthetic divertor spectroscopy diagnostic developed in Section 3.2 enables one-to-one comparison between simulated and measured views to investigate the simulation-experiment discrepancy.

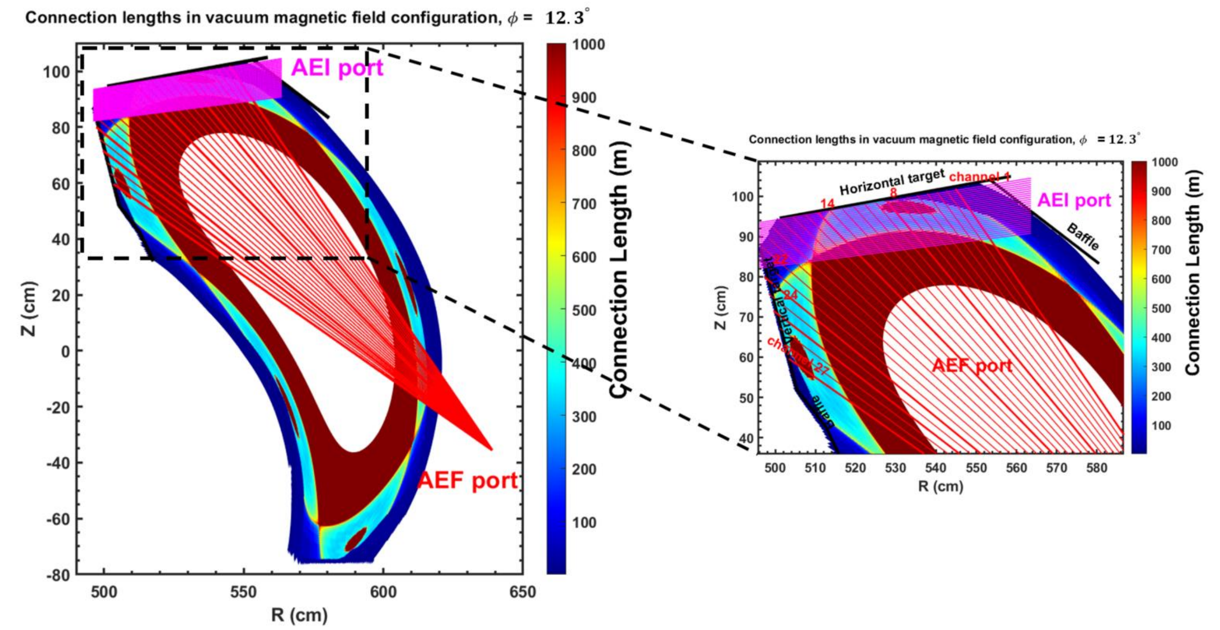

Figure 4.13: Lines of sight of the divertor-spectroscopy system at the upper divertor for the toroidal angle $\varphi = 12.3^{\circ}$. Magenta lines indicate horizontal sightlines, and red lines indicate vertical sightlines.

In the EMC3-EIRENE simulations of discharge 20181010.026 (standard magnetic configuration) selected in Section 4.1, we likewise find that as the radiated-power fraction f_{rad} increases there is a non-negligible C II (426 nm) and C III (465 nm) emission near the O-point. Figure 4.14 shows the modeled C II and C III distributions obtained at different f_{rad} using the cross-field transport coefficients $D_{\perp} = 0.5$ m²/s and $\chi_{\perp} = 1.5$ m²/s. Consistent with the measurements, the dominant C II emission region migrates from the strike point toward the magnetic X-point as f_{rad} increases; however, during this migration simulations also develop appreciable O-point emission. For C III, the simulated O-point emission is even more pronounced, in contrast to the experiment where no comparable O-point C III signal is observed. To investigate this discrepancy, we reproduce discharge 20181010.026 at $f_{rad} = 0.26$ and perform scans of key EMC3-EIRENE inputs, including cross-field transport coefficients, f_{rad}, separatrix density, and the initial energy of sputtered carbon atoms, and compare the resulting synthetic signals with the measurements.

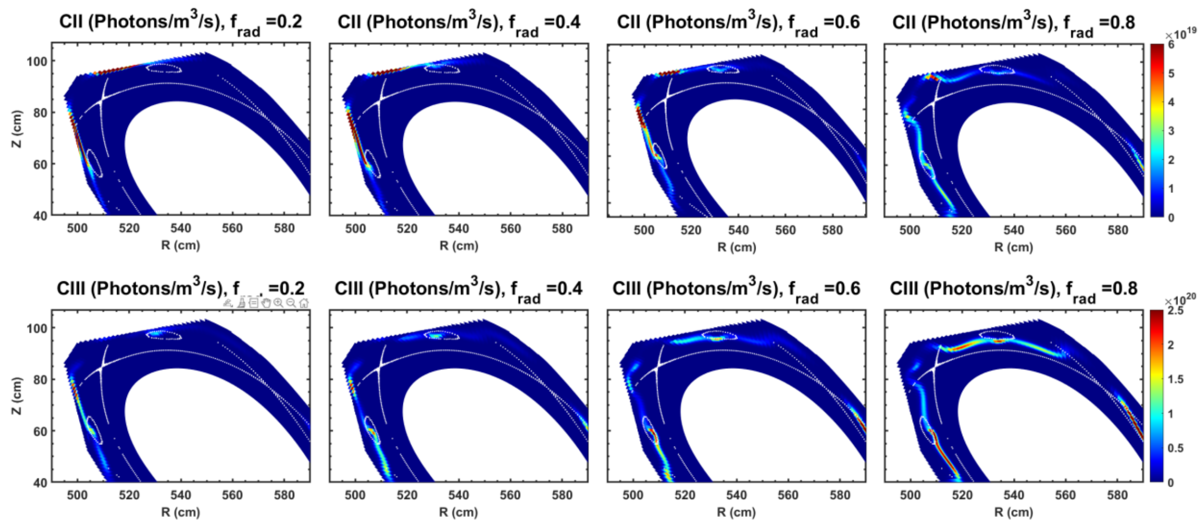

Figure 4.14: Distributions of C II (426 nm) and C III (465 nm) at different radiated power fractions f_{rad}, obtained from simulations using experimental inputs.

Radiated power fraction scan

With the separatrix electron density fixed at $n_{e,sep} = 5.0 \times 10^{19}$ m^{-3} and cross-field transport coefficients $D_{\perp} = 0.5$ m²/s and $\chi_{\perp} = 1.5$ m²/s, we take the measured radiated-power fraction $f_{rad} = 0.26$ as a reference. Accounting for the bolometer uncertainty, we scan f_{rad} from 0.15 to 0.40. Figure 4.15 compares simulated and measured C II (711.8 nm) and C III (674.1 nm) emission in the vertical view. For C II, the simulated peak location and amplitude agree well with the measurement; modestly increasing f_{rad} improves the match near the strike point but slightly enhances emission near the O-point. For C III, the simulated intensity is about an order of magnitude lower than measured, and the simulations predict a pronounced O-point peak that is absent in the data. The persistence of O-point C III emission across $f_{rad} = 0.15 - 0.4$suggests that it is unlikely to arise from a mis-specified f_{rad} due to bolometer error.

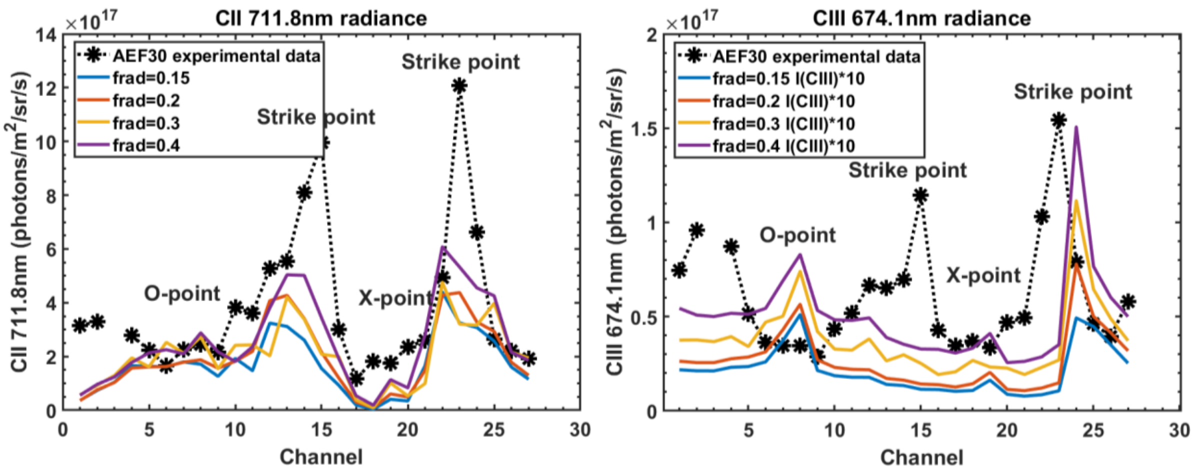

Figure 4.15: Comparison between simulated and measured C II and C III emission in the vertical view, with fixed separatrix electron density and cross-field transport coefficients.

Separatrix electron density scan

Next, we scan the separatrix electron density $n_{e,sep}$, while holding $f_{rad} = 0.4$, $D_{\perp} = 0.5$ m²/s and $\chi_{\perp} = 1.5$ m²/s. As shown in Figure 4.16, at low $n_{e,sep}$, C II and C III emission near the O-point is weak; as $n_{e,sep}$ increases, O-point emission becomes stronger, and for C III the dominant emission region shifts from the strike point to the O-point. This is because, at higher $n_{e,sep}$, T_e in the divertor region decreases, leading to deeper impurity penetration and stronger C II and C III emission near the O-point. Figure 4.17 shows the spatial distributions of T_e, C^+ and C^{2+} for different $n_{e,sep}$: C^+ is concentrated near the magnetic island where T_e is a few eV, whereas C^{2+} is localized in regions around ∼10 eV. As shown in Figure 4.18, at $n_e = 1.0 \times 10^{19}$ m^{-3} the collisional-radiative model predicts that the C^+ abundance peaks near 2 eV, whereas C^{2+} peaks around 5 eV. This energy dependence accounts for the correlation between the C^+ or C^{2+} spatial patterns and T_e in Figure 17. Comparison with other diagnostics indicates that while $n_e = 2.0 \times 10^{19}$ m^{-3} reproduces the absence of O-point C II and C III emission, it is lower than the Thomson scattering density. This indicates that we may need to adjust other input parameters to modify the electron-temperature distribution. Next, we scan the cross-field transport coefficients.

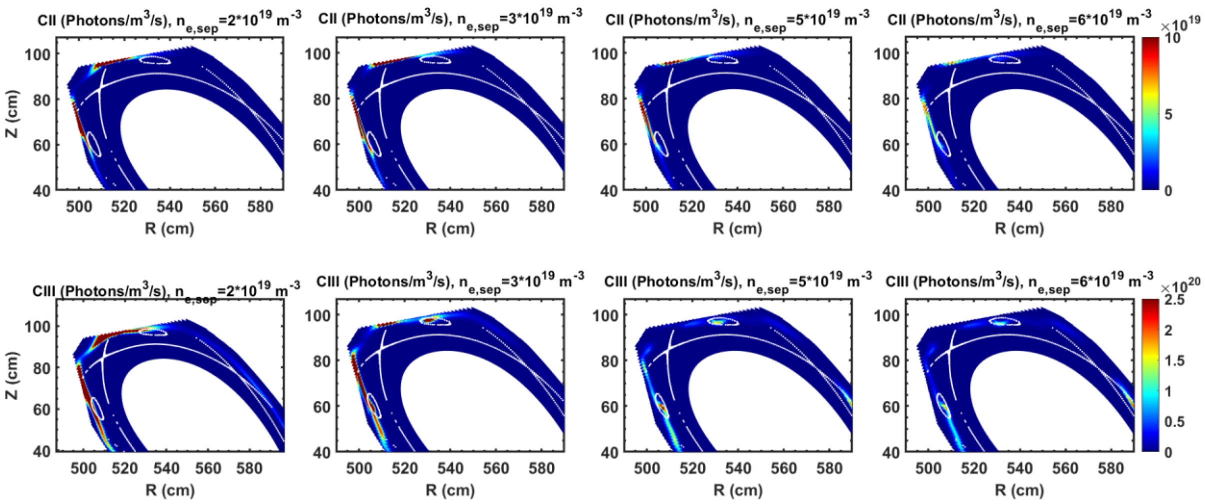

Figure 4.16: Simulated C II and C III emission distributions at $f_{rad} = 0.4$, with fixed cross-field transport coefficients, for varying separatrix electron density.

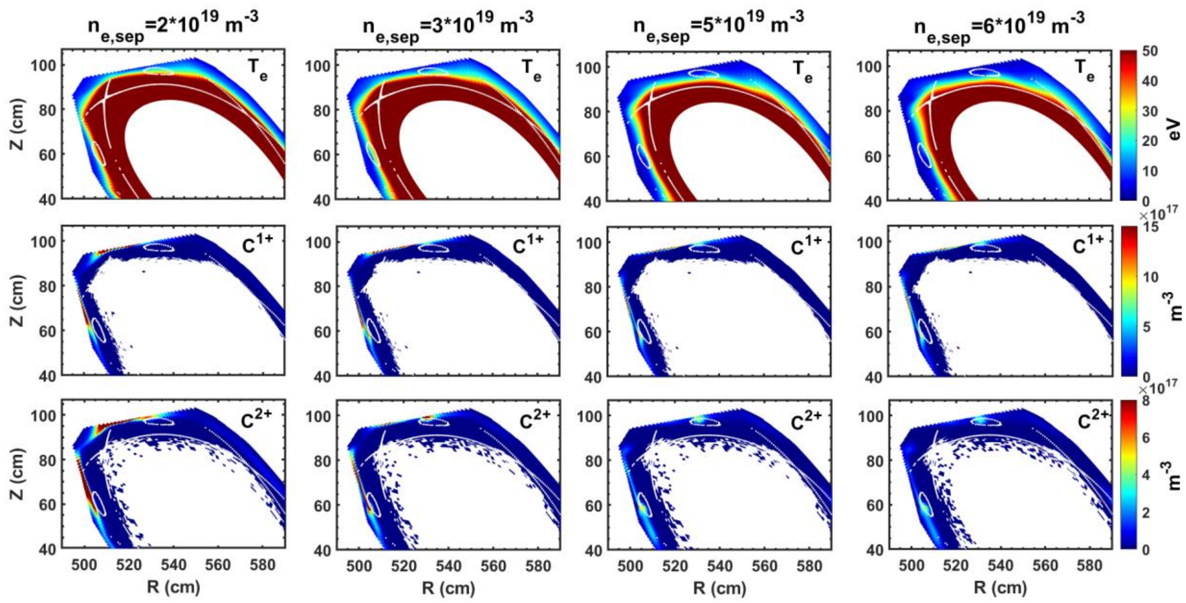

Figure 4.17: Spatial distributions of Electron temperature, C^{+} and C^{2+} at different separatrix densities.

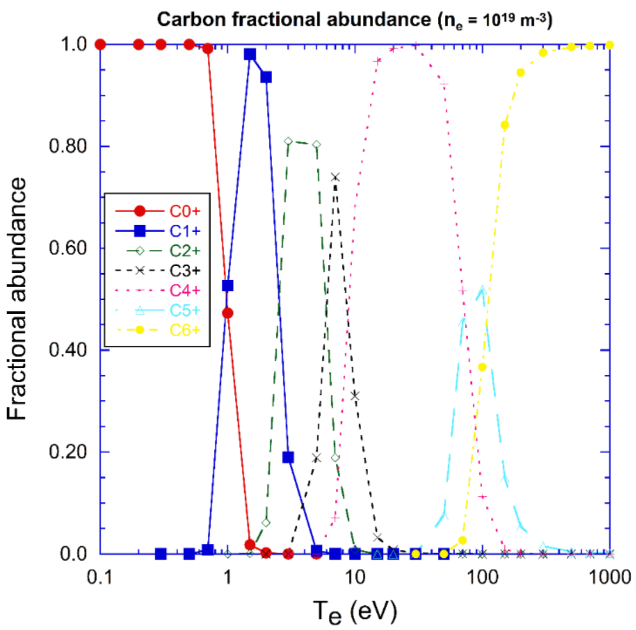

Figure 4.18: Collisional-radiative model predictions for the fractional abundances of carbon charge states at a fixed electron density.

Cross-field transport coefficient scan

At fixed $P_{tot} = 6$ MW, $f_{rad} = 0.26$, and $n_{e,sep} = 5.0 \times 10^{19}$ m^{-3}, we scan the cross-field transport coefficients $D_{\perp}$ and $\chi_{\perp}$. Table 4.3 lists the cases considered. As summarized in Figure 4.19: case 1 adopts the commonly used set in W7-X; case 2 reduces heat transport inside the remnant magnetic island; case 3 reduces both particle and heat transport inside the island; and case 4 reduces heat transport but increases particle transport inside the island. Figure 4.20 shows the simulated T_e, C II and C III emission for each case under the same P_{tot}, f_{rad}, and $n_{e,sep}$. Despite large variations in $D_\perp$ and $\chi_\perp$that substantially modify T_e around the island and affect impurity penetration, O-point C II and C III emission remains non-negligible. This indicates that adjusting cross-field transport to modify T_e or impurity penetration alone is insufficient to eliminate the simulated O-point emission.

Table 4.3: Radiation fraction, separatrix density and cross-field transport coefficients for different modeling cases.

Case	f_{rad}	$n_{e,sep}$ (10^{13} cm^{-3})	$D_{\perp,e} = D_{\perp,imp}(\mathrm{m}^2/\mathrm{s})$	$\chi_{\perp,i} = \chi_{\perp.e}(\mathrm{m}^2/\mathrm{s})$
Case 1	26%	5	0.5	1.5
Case 2	26%	5	Figure 4.19(a)	Figure 4.19(b)
Case 3	26%	5	Figure 4.19(c)	Figure 4.19(d)
Case 4	26%	5	Figure 4.19(e)	Figure 4.19(f)

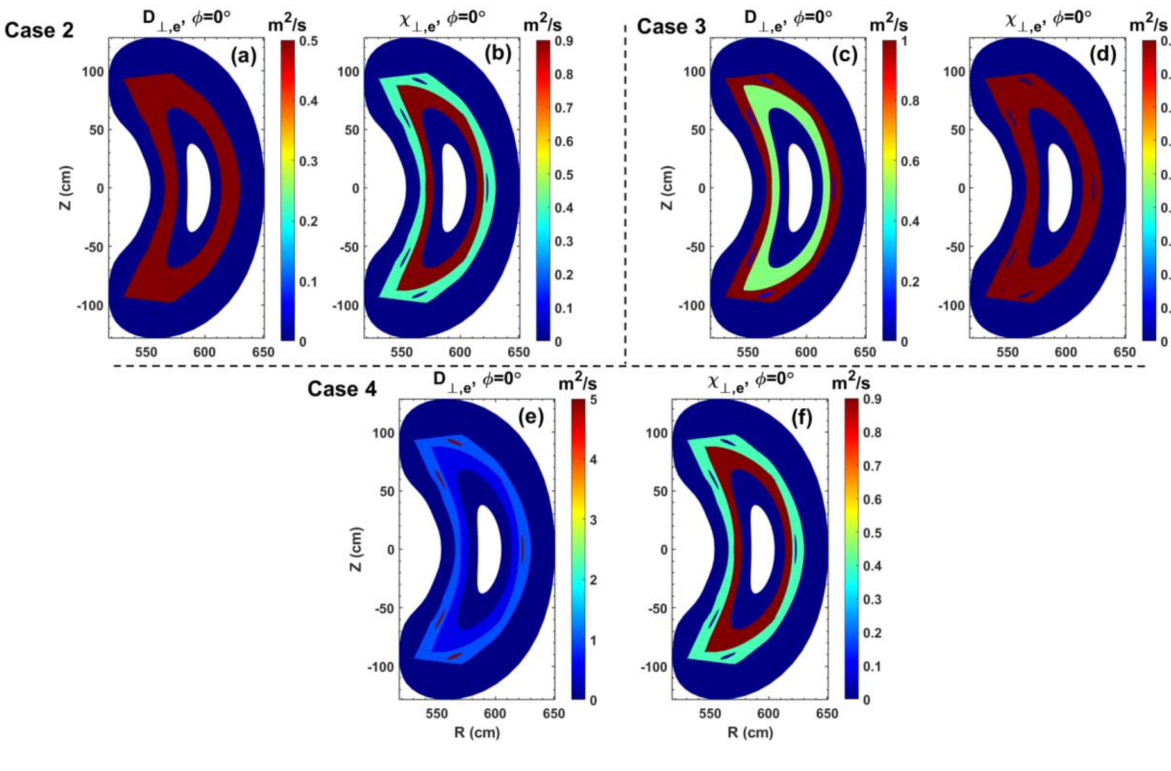

Figure 4.19: Spatial distributions of the cross-field transport coefficients for cases 2-4.

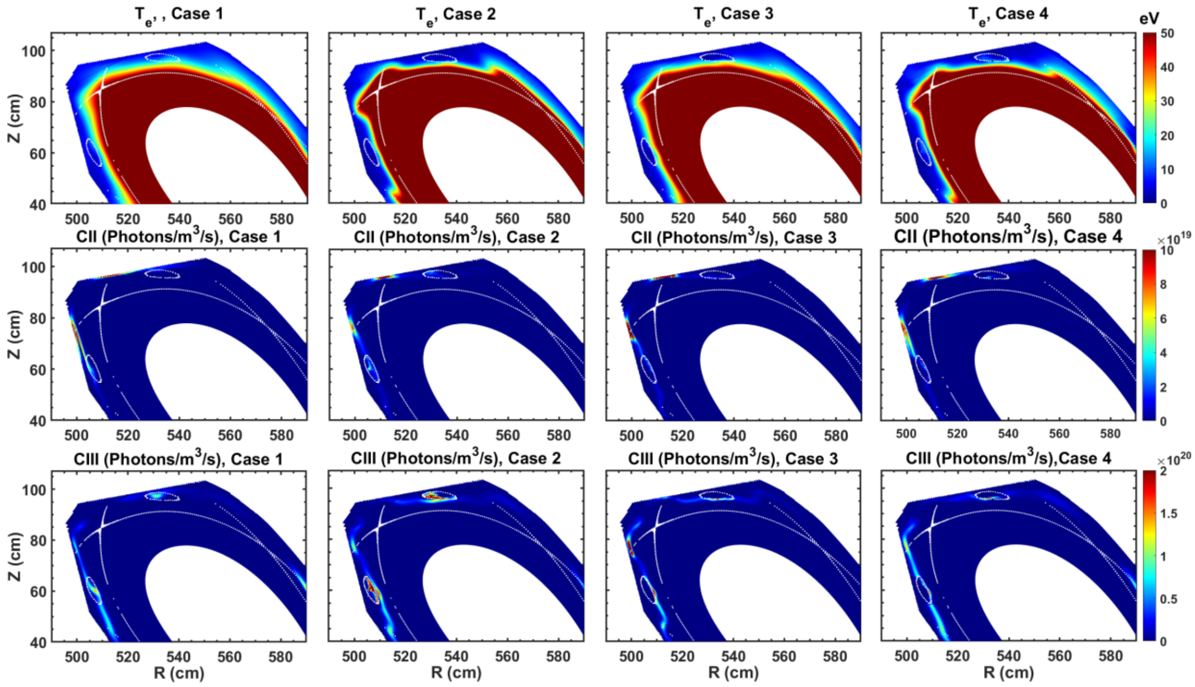

Figure 4.20: Simulated electron temperature, C II and C III emission distributions for cases 1-4.

Sputtered-carbon initial energy scan

We also scanned the initial energy (E0) of sputtered carbon, E0 (0.01 eV, 1.0 eV, and 10 eV). Figure 4.21 shows the C II and C III distributions for these three cases at $f_{rad} = 0.6$, $n_{e,sep} = 4.0 \times 10^{19}$ m^{-3}, $D_\perp = 0.5$ m²/s and $\chi_\perp = 1.5$ m²/s. As E0 increases, the impurity penetration depth increases slightly, but the O-point distributions of C II and C III are nearly unchanged. In other words, changing E0 does not reduce the O-point C II and C III emission.

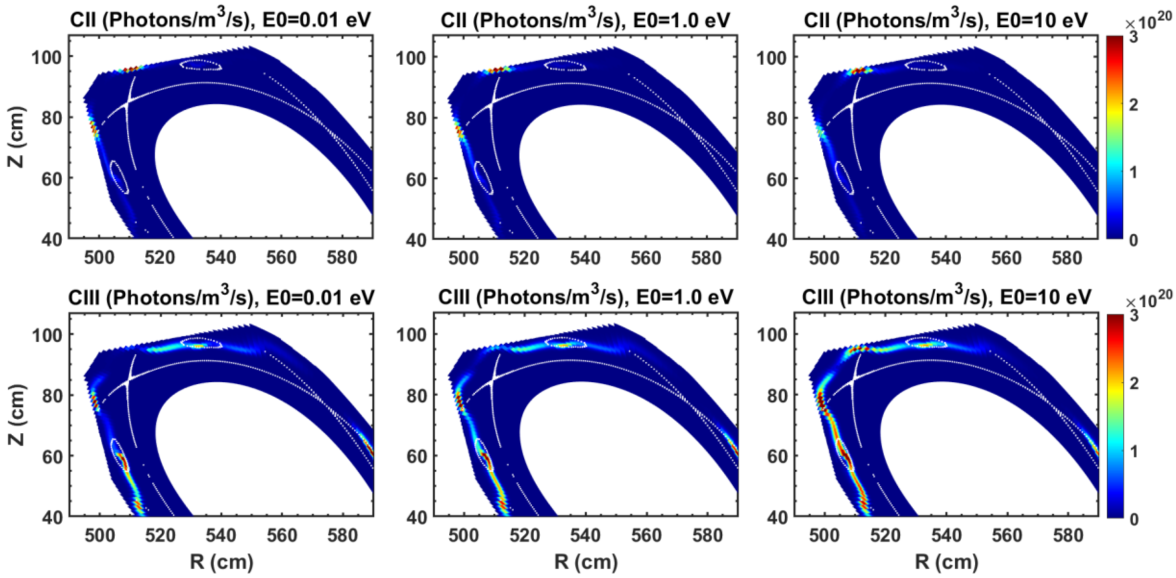

Figure 4.21: Simulated C II and C III emission distributions for varying sputtered-carbon initial energy.

Overall, reproducing the experimentally observed absence of O-point C II and C III emission in the standard configuration as f_{rad} increases remains challenging for EMC3-EIRENE. Within the parameter ranges explored above, this behaviour was not reproduced. A possible cause is the omission of error-field and drift physics in the present model version.

4.3 Impact of 3D magnetic equilibrium on edge transport

Previous studies show that finite plasma beta can affect edge stochastization [47, 105, 106], and toroidal currents can shift the magnetic islands in the W7-X SOL [107]. In order to obtain more reliable output parameters or prediction results, the equilibrium effects (including plasma beta and current effects) on edge magnetic topologies have been considered in EMC3-EIRENE by integrating with HINT code. Below we consider one discharge and reproduce the measurements under three magnetic-equilibrium assumptions: (i) the vacuum field, (ii) the field including the toroidal plasma current, and (iii) the field including both the toroidal current and finite plasma beta.

We selected discharge 20180919.025 because it exhibits a toroidal plasma current that can modify the edge magnetic topology. Figure 4.22 presents the time evolution of the main parameters; at the chosen instant (t = 6 s) the electron-cyclotron heating power was 4.0 MW, the total radiated power 0.72 MW, and the line-integrated electron density 5.5 × 10^{19} m^{-2}. The mean effective ion charge had reached Z_{eff} = 1.3, the diamagnetic energy was about 470 kJ, and the plasma toroidal current roughly 3.0 kA. Combining the electron density and temperature from Thomson-scattering measurements with the ion temperature from X-ray Imaging Crystal Spectroscopy (XICS) yields an on-axis beta of approximately 2 % at this moment. For validating the EMC3-EIRENE input parameters we use the same diagnostic set as before: divertor spectroscopy detecting C II (465.01 nm) and H_γ (433.99 nm) emission in the divertor region; Thomson scattering (TS) providing radial profiles of electron density and temperature; divertor Langmuir probes (LPs) measuring ion-saturation currents at the divertor targets; infrared cameras capturing heat-flux distributions on nine divertors.

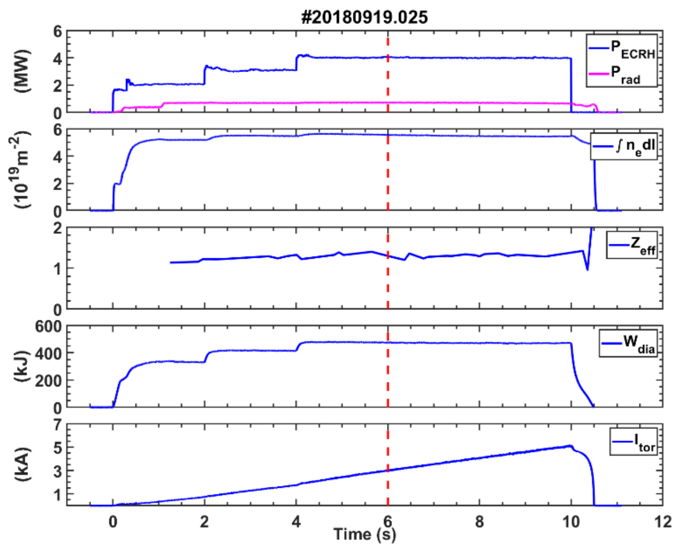

Figure 4.22: Time evolution of key W7-X parameters during discharge #20180919.025. Dashed line indicates the selected modeling time point.

In the modelling setup, a power flux of 4 MW is imposed on the innermost boundary of the computational domain and is split evenly between electrons and ions. Carbon atoms sputtered from the divertor and baffle targets are taken as the only radiator; their sputtering rate is controlled by radiated power. The radiation fraction is fixed at 18 %, consistent with the experimental value. The upstream separatrix electron density and the cross-field transport coefficients for particles ($D_\perp$) and energy ($\chi_\perp$) are subsequently adjusted to reproduce the measurement.

First, under vacuum-field conditions we scanned the upstream separatrix electron density at 2.0, 2.2, 2.5 × 10^{19} m^{-3} to identify an appropriate value. Throughout the scan the cross-field transport coefficients were fixed at the values routinely employed on W7-X, $D_\perp = 0.5\ \mathrm{m}^2/\mathrm{s}$ and $\chi_\perp$=1.5 m^2/s, corresponding to cases S1-S3 in Table 4.4. Figure 4.23 compares the synthetic signals from these three cases with the measured diagnostics. Case S2 ($n_{e,sep} = 2.2 \times 10^{19}\ \mathrm{m}^{-3}$) reproduces the experimental data best overall; however, the simulated electron-temperature profile is still lower than the measurement inside the last closed flux surface, implying that the chosen $\chi_\perp$ is too high there. Consequently, we fixed $\mathrm{n_{e,sep}} = 2.2 \times 10^{19}\ \mathrm{m}^{-3}$ and performed a second scan of the cross-field transport coefficients, allowing different values inside and outside the LCFS (cases S4-S6 in Table 4.4). Figure 4.24 shows the resulting synthetic signals versus experiment. The parameter set of case S4 yields the closest overall agreement, with the simulated temperature profile now matching the measurements well.

Table 4.4: Input separatrix electron densities and cross-field transport coefficients for the individual cases

Case	$n_{e,sep}$ (10^{19} m^{-3})	$D_\perp$ (m^2/s)	$\chi_\perp$ (m^2/s)
S1	2.0	0.5	3*D
S2	2.2	0.5	3*D
S3	2.5	0.5	3*D
S4	2.2	0.3 (r_{eff}<1), 0.4 (r_{eff} ≥1)	3*D
S5	2.2	0.8 (r_{eff}<1), 0.3 (r_{eff} ≥1)	3*D
S6	2.2	0.3 (r_{eff}<1), 1.0 (r_{eff} ≥1)	3*D

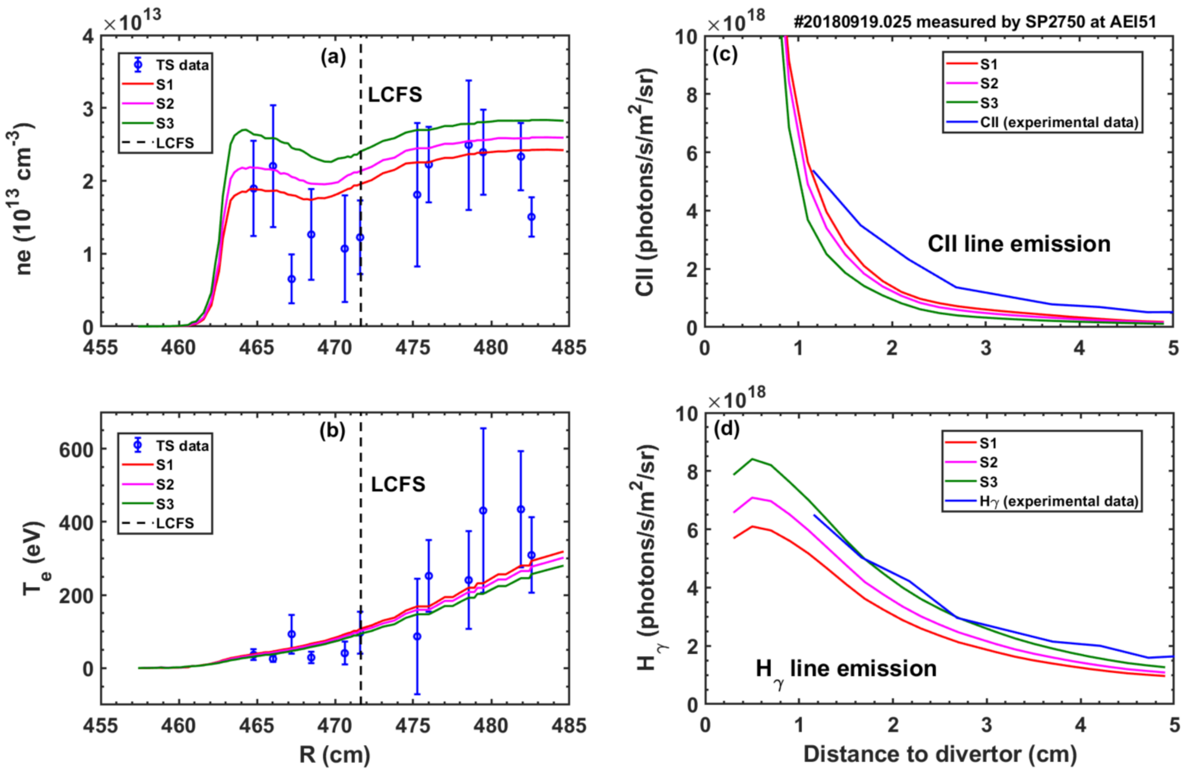

Figure 4.23: Experimental data and simulation results in case S1-S3: (a) electron density, (b) electron temperature, (c) CII line emission, (d) H_{γ} line emission.

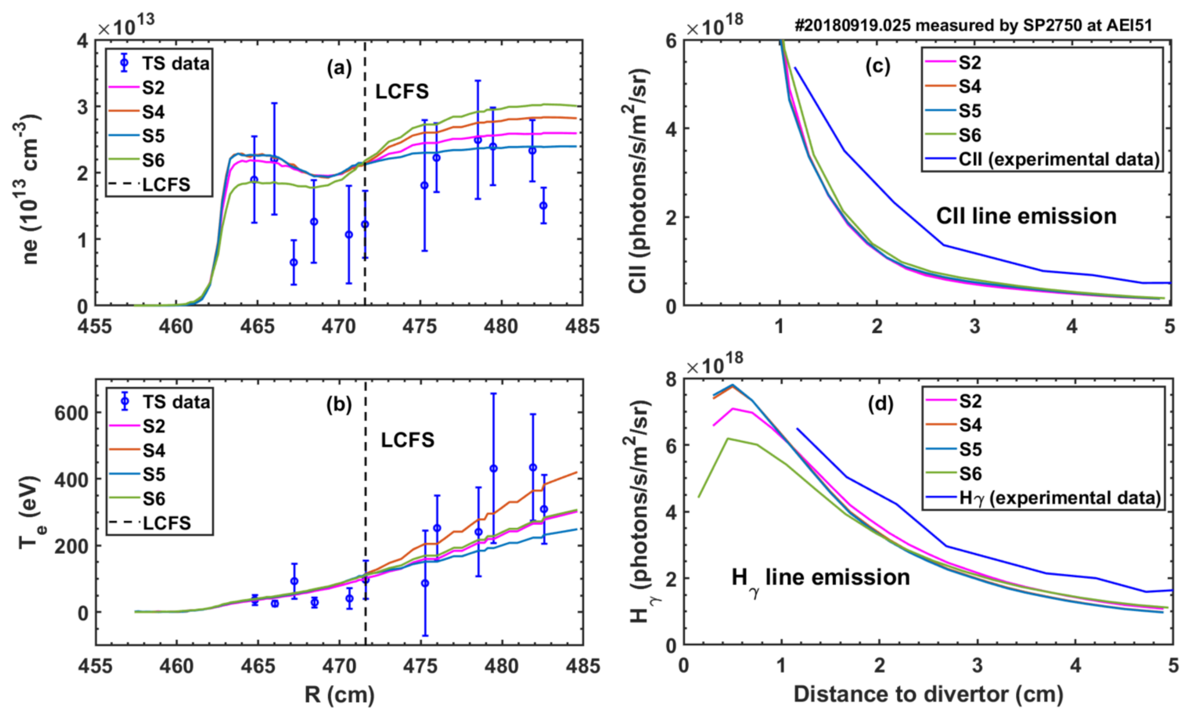

Figure 4.24: Experimental data and simulation results in case S4-S6: (a) electron density, (b) electron temperature, (c) CII line emission, (d) H_{γ} line emission.

Using the S4 input-parameter set, we next examine how plasma-equilibrium effects modify edge transport, with particular emphasis on impurity behavior. Figure 4.25 compares synthetic and experimental signals for three magnetic configurations: (i) the vacuum field, (ii) the field including the measured toroidal current I_φ ($\approx$ 3 kA), and (iii) the field including both the toroidal current and an on-axis β_{ax} = 2 %. The principal difference among the three cases is the electron density profile, which in turn governs the distribution of the H_γ line emission; the electron temperature profiles remain almost identical. Including both toroidal plasma current and β increases the boundary density and therefore yields a H_γ pattern that agrees best with the measurements.

We further examine measurements at the divertor targets, namely the heat flux and ion saturation current distributions. Figure 4.26 shows heat flux distribution on the divertor targets: all simulations underestimate the absolute level, yet the vacuum and I_φ-only cases differ little from each other—apart from a strike-line that shifts farther from the gap when I_φ is included. The shift results from the larger iota, which drives the magnetic island radially inward. When both I_φ and β are included, the heat-flux profile becomes markedly more peaked, most clearly on the vertical target. Consequently, the simulated heat-flux pattern with beta effects agrees more closely with the measurements. A likely explanation is that β-induced changes in magnetic topology produce a more peaked footprint. A similar trend appears in Figure 4.27: the measured and simulated divertor ion-saturation currents show little difference between the vacuum and I_φ cases, whereas the I_φ+ β configuration produces a much sharper peak.

To illustrate how magnetic topology affects impurity transport, Figure 4.28 shows the connection length and C^{3+} density at a toroidal angle of 12.33° for all three configurations. Comparing the vacuum field with the $I_\varphi = 3$ kA case indicates that the toroidal current pulls the remnant magnetic island inward, but the C^{3+} distribution almost keeps unchanged. In contrast, adding $\beta_{ax} = 2$ % alters the edge magnetic topology and increases field-line stochastization. These changes modify the electron density and temperature distributions, recycling flux, and impurity sources, promoting deeper impurity penetration. As a result, the C^{3+} shifts inward toward the remnant island center in the combined I_φ+ β equilibrium.

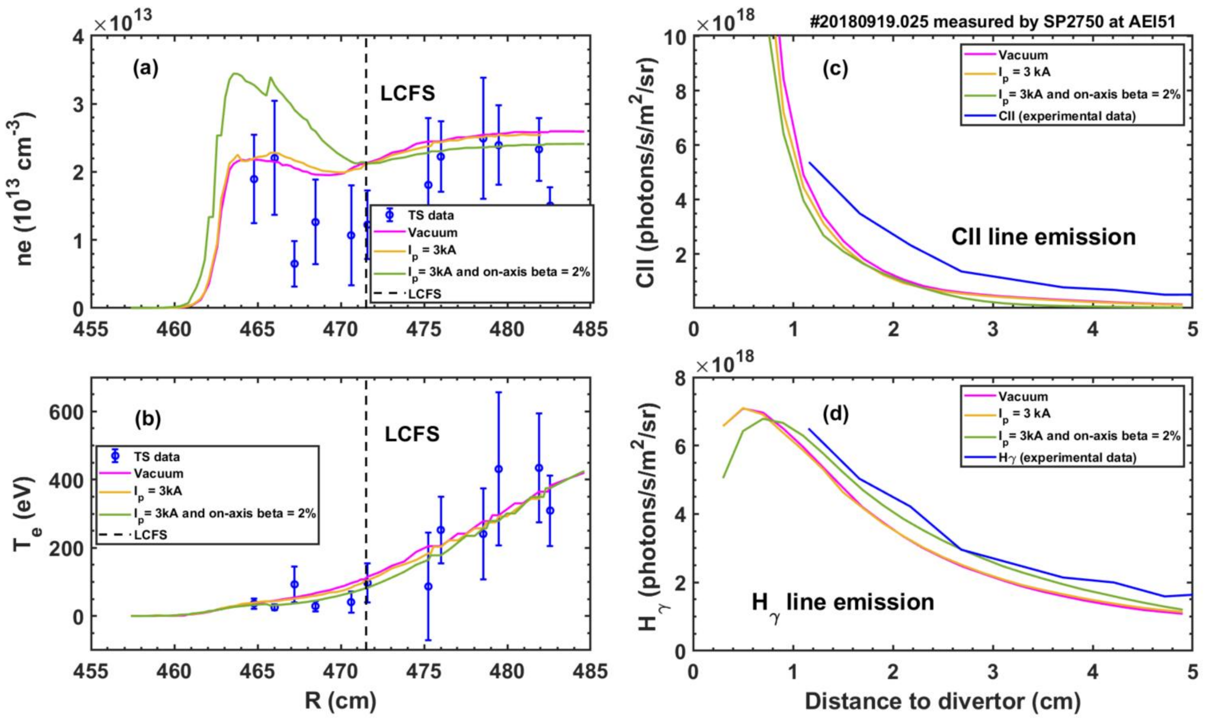

Figure 4.25: Experimental data and simulation results at three different magnetic configurations.

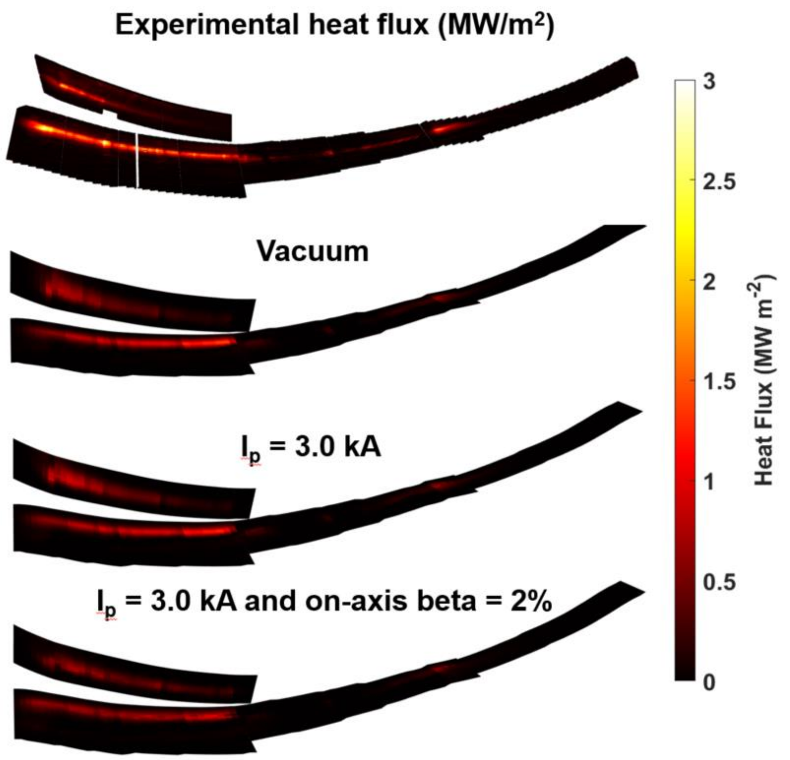

Figure 4.26: Experimental heat flux and simulation results at three different magnetic configurations.

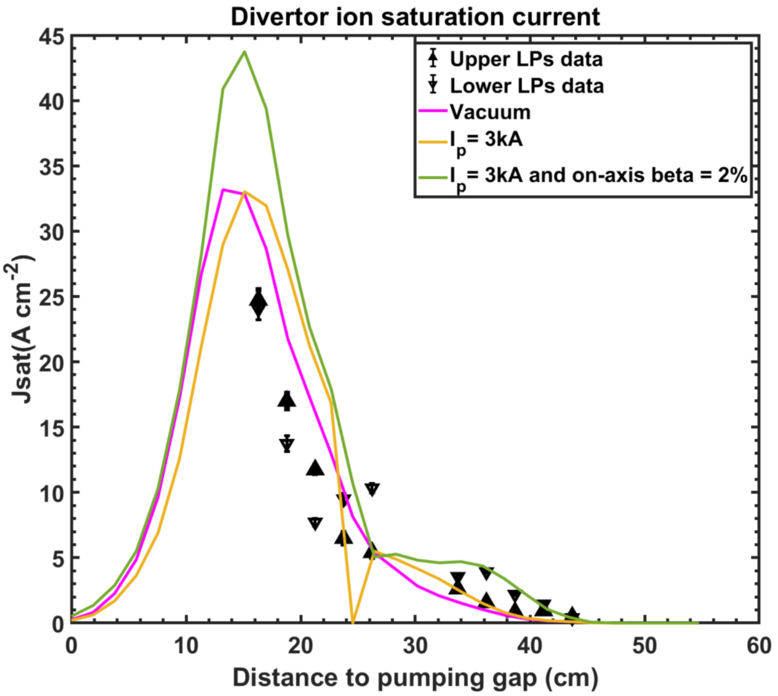

Figure 4.27: measured and simulated divertor ion saturation current at three different magnetic configurations.

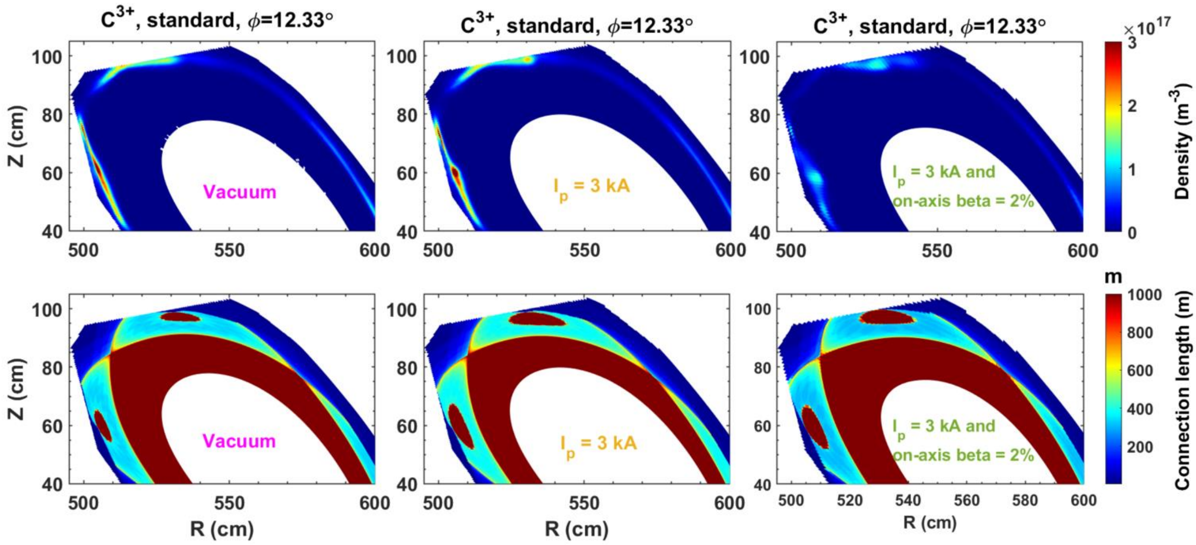

Figure 4.28: C^{3+} distributions and connection lengths in different magnetic configurations.

4.4 Summary

A synthetic reconstruction of divertor spectroscopy has been developed to study the transport behavior of background plasma and intrinsic carbon on W7-X. By comparing the experimentally measured line emission radiance with synthetic diagnostic results, we can more precisely evaluate the cross-field transport coefficients in EMC3-EIRENE simulations, especially with regard to impurity transport. We found that using spatially non-uniform transport coefficients in simulations is able to improve the agreement with multiple local diagnostics.

In density ramp-up discharges with the standard configuration, the measured C II or C III emission front moves from the strike point toward the X-point as f_{rad} increases. The simulations reproduce this trend but also show a non-negligible O-point peak in C II or C III emission. Extensive scans of EMC3-EIRENE inputs indicate that, as f_{rad} rises and the remnant island cools, impurity accumulation near the O-point is difficult to avoid. This may be due to the present model's omission of error-field and drift effects.

By integrating HINT with EMC3-EIRENE, we further assess how equilibrium effects (plasma beta and toroidal plasma current) influence impurity transport. The toroidal current pulls the remnant island inward while leaving the impurity pattern largely unchanged. In addition, beta-induced changes in the edge magnetic topology modify impurity sources and transport, thereby altering impurity penetration depth.

5 Bayesian neural network inference of EMC3-EIRENE inputs from edge diagnostics

Aligning EMC3-EIRENE with edge diagnostics is challenging. Manual iterative adjustment of EMC3-EIRENE inputs—especially the cross-field transport coefficients—requires many full simulations and makes the workflow slow and costly. The principal original contribution of this chapter is the development of a surrogate-assisted Bayesian inversion method that fast infers EMC3-EIRENE input parameters consistent with diagnostic measurements and their uncertainties. The workflow is as follows: an feed-forward neural network (FNN) surrogate, trained on an EMC3-EIRENE dataset, provides a fast mapping from inputs to synthetic diagnostic signals at the measurement sightlines; the surrogate is then coupled with a Bayesian framework to obtain posterior distributions of the inputs. Running EMC3-EIRENE with the maximum-a-posteriori (MAP) inputs reproduces the diagnostics well while greatly reducing computational cost compared with manual trial-and-error.

This chapter is arranged as follows. First, the challenges in optimizing EMC3-EIRENE input parameters are outlined. Second, the background of machine learning are reviewed. Third, a Bayesian neural-network framework is developed to infer EMC3-EIRENE input parameters from experimental measurements. Finally, the framework's accuracy is assessed using a W7-X discharge.

5.1 Challenges in optimizing EMC3-EIRENE input parameters

Edge plasma transport modeling plays a crucial role in understanding edge plasma behavior and optimizing the design of fusion devices. This is particularly important for devices with complex three-dimensional edge structures, such as W7-X, where boundary diagnostics are limited to localized regions, complicating transport analysis. The EMC3-EIRENE can effectively handle configurations with complex edge structures, and has been widely applied in both tokamaks and stellarators for interpreting experimental observations and conducting predictive simulations. However, aligning EMC3-EIRENE simulations with experimental diagnostic measurements often requires manual tuning of input parameters (especial the free input parameters cross-field transport coefficients) and running multiple simulations. Each simulation is compared to the experimental data to find the best match, making the process both time-consuming and computationally demanding. To overcome this problem, a major challenge is quickly establishing a quantitative physical relationship from EMC3-EIRENE input parameters to their corresponding synthetic diagnostics. These synthetic diagnostics represent simulated values at the same locations as the experimental measurements. Fortunately, a neural network-based surrogate model presents a promising solution.

In recent years, machine-learning techniques have gained broad traction in fusion research, powering advances in plasma control [108-110], disruption predictions [111-113], and the acceleration of large-scale simulations [114-116], due to its exceptional capabilities in classification, regression, and pattern recognition. In this chapter, we introduce a method that combines a FNN surrogate model with Bayesian inference via Dynamic Nested Sampling [117] to quickly identify EMC3-EIRENE input parameters that yield synthetic diagnostics consistent with experimental data on W7-X. The FNN is initially trained on an EMC3-EIRENE simulation dataset, fully learning the physical relationships between each input parameter and its corresponding synthetic diagnostic data. After that, Bayesian inference then evaluates how well those surrogate-generated signals match the experimental measurements and their measurement uncertainties to obtain the posterior distributions of the EMC3-EIRENE input parameters. The entire procedure runs in only a few minutes on an ordinary desktop PC. With this method, the many full EMC3-EIRENE runs that conventional workflows demand are replaced by a single final run, while the resulting simulation remains tightly constrained by the diagnostic data and the underlying physics captured by EMC3-EIRENE.

5.2 Introduction to machine learning

Although EMC3-EIRENE is powerful for simulating and predicting edge plasma transport, its widespread application is severely limited by its high computational demands and the stringent expertise required for its operation. In recent years, machine learning techniques have gained popularity in fusion research due to their remarkable capability to capture complex nonlinear patterns. In this thesis, several neural network surrogate models are developed that leverage advanced machine learning techniques to significantly accelerate EMC3-EIRENE simulations by drastically reducing computational overhead. Furthermore, these surrogate models can rapidly generate EMC3-EIRENE simulation outputs consistent with experimental observations based on edge plasma diagnostic measurements. This provides a fast and effective tool for impurity transport studies that incorporate experimental edge data, greatly cutting computational demands and streamlining the overall research workflow.

This section introduces the neural-network tools used. It first outlines the fundamentals of fully connected feed-forward and convolutional neural networks, then explains the key hyper-parameters that govern network performance. Search strategies and early-stopping criteria for hyper-parameter optimization are described next. The section closes with recent examples of neural-network applications in fusion research.

5.2.1 Neural-network fundamentals

Artificial neural networks (ANNs) mimic biological neural systems, making them similar in both architecture and function. For biological neurons, they receive input signals from other neurons through their dendrites, process these signals, and then produce output signals to other neurons. In ANNs, an artificial neuron receives input signals from the previous layer, performs a weighted sum, adds a bias, and then passes the result through an activation function to the next layer. ANNs is a fundamental and widely used machine learning approach known for their capability in processing data, particularly in capturing complex nonlinear relationships in data. The following sections will introduce the neural network architectures employed in this study, namely fully connected feed-forward neural networks and convolutional neural networks, as well as the hyperparameter optimization algorithms used to enhance model performance.

5.2.1.1 Forward neural networks

In conventional neural networks, the most commonly used layer type is the fully connected feed-forward layer (FNN), where each neuron in one layer is connected to every neuron in the adjacent layer, while there are no connections among neurons within the same layer. A typical FNN is composed of several layers: the input layer, hidden layers, and the output layer. The input layer consists of input data that you want to feed into your neural network. Hidden layers are located between input layer and output layer. The number of layers and number of neurons per layer are important hyperparameter in FNN, as they define the structure and complexity of the network. Their selection needs to be determined based on the specific problem. The output layer is comprised of the predictions you want to generate. Depending on the specific task, the neurons in the output layer may behave differently. For multi-class classification, the output layer might consist of multiple neurons with a softmax activation function. For regression, it could have one or more neurons with a linear activation function. Figure 5.1(a) shows a 3-layer FNN network. This network comprises three input neurons, two hidden layers with five neurons each, and a single output layer containing one neuron.

In order to understand what neural networks are, we need to know how perceptrons (artificial neurons) work. Figure 5.1(b) illustrates the working principle of a neuron. As shown, for neuron j in layer l, it receives input signals x_1^{l-1}, x_2^{l-1}, …, from all neurons in the previous layer l-1. The corresponding weights between neuron j and all neurons in the previous layer are ω_1^l, ω_2^l, …, and the bias of neuron j is b_j^l. The weighted sum z_i^l of neuron j in layer l is given:

$$z_i^l = \sum_i^n \omega_i^l x_i^{l-1} + b_j^l \qquad 5.2.1$$

Weights and bias are two key parameters in calculating weighted sums z_i^l. Weights determine the impact of the input signal by amplifying or diminishing it. Bias, on the other hand, is a constant term that provides additional flexibility to the model by allowing it to produce a non-zero output when the input is zero. All of them can be adjusted by the backpropagation algorithm during the training process.

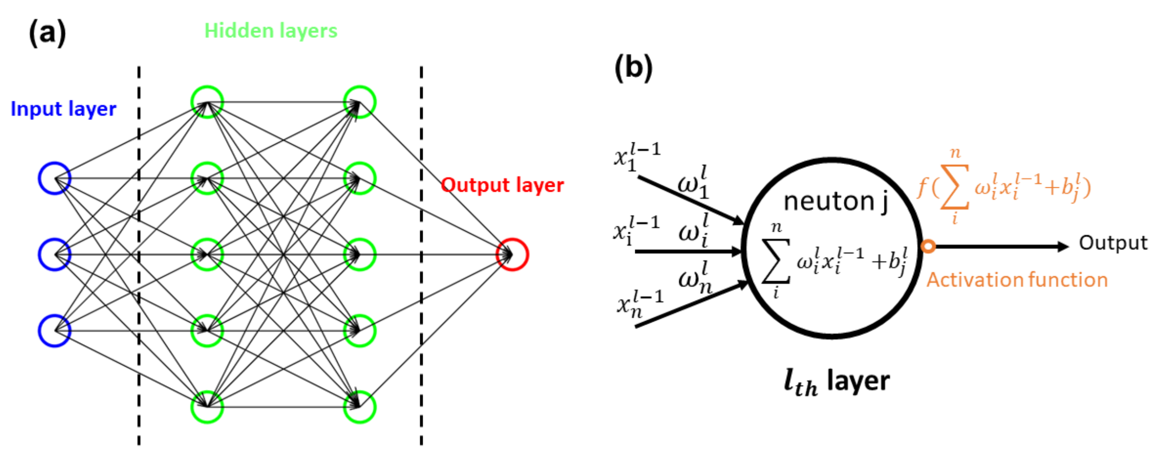

Figure 5.1: (a) Schematic diagram of a 3-layer fully connected neural network with three inputs, two hidden layers of five neurons each, and a single output layer of one neuron. (b) diagram illustrating the working principle of a neuron.

The above weighted sum are just linear transformations of the input signals. If there are only linear transformations in the network, then no matter how many layers there are in the network, the effect is the same as a single-layer network, which is unable to capture the complex non-linear relationships in the data. Therefore, after calculating the weighted sum, an activation function needs to be applied to introduce non-linearity into the network so that it can model complex patterns. The output of neuron *j* in layer l to neuron in the next layer can be represented as:

$$y_i^l = \Phi(z_i^l) = \Phi(\textstyle\sum_i^n \omega_i^l x_i^{l-1} + b_j^l) \qquad 5.2.2$$

where $\Phi(z_i^l)$ represents the activation function, and y_i^l is the output of neuron *j* in layer l. The output y_i^l of neuron *j* is the same for all neurons in the next layer *l*+1, but the weights ω_i^{l+1} associated with each connection to the neurons in layer *l*+1 are different. Repeating this process, the predicted values in the output layer can be obtained based on the initial input signals. In the example above, the output layer includes one neuron, so there is only one predicted values in this model. Actually, the output layer can contain multiple neurons. The neural network model used in the next chapter of this thesis is designed for predicting multiple outputs. The size of a neural network is typically described by several key factors: the number of layers, the number of neurons per layer, and the number of parameters. The total number of parameters in a neural network is the sum of its weights and biases. In the example, the number of parameters is equal to 56 ($[3 \times 5] + [5 \times 5] + [5 \times 1] = 45$ weights and $5 + 5 + 1 = 11$ biases).

The activation function is an important part of neural networks. Typically, it is differentiable with respect to the weighted sum because gradient-based optimization methods, such as backpropagation used during training, rely on these differentials. Commonly used activation functions include Sigmoid Function [118], Hyperbolic Tangent (Tanh) [119], Rectified Linear Unit (ReLU) [120] and so on. Figure 5.2 illustrates these activation functions. The Sigmoid Function is one of the earliest used function in neural network. It is defined as:

$$\sigma(z) = \frac{1}{1-e^{-z}} \tag{5.2.3}$$

with an output range between 0 to 1. This makes it useful for binary classification tasks where the output needs to be probabilistic. The Tanh function is somewhat like the sigmoid function but produces values ranging from -1 to 1. It is expressed as:

$$tabh(z) = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}} \tag{5.2.4}$$

compared to the sigmoid function, Tanh functions usually converge faster during training. However, for both Sigmoid Function and Tanh (bounded activation functions), they suffer from the vanishing gradient problem when the weighted sum z is either very large or very small. It because when these activation functions saturated, their gradients are very small. During backpropagation, the application of the chain rule results in the multiplication of these small gradient values. Consequently, the gradient decreases during backpropagation between layers, resulting in minimal or negligible weight updates. This can significantly slow down the learning process and may even cause the network to stop learning. ReLU is commonly used activation function in modern neural networks because it is simple and effective. It introduces sparsity by activating only a fraction of neurons, which helps to reduce the computational load [120]. It is formulated as:

$$f(z) = max(0, z) \tag{5.2.5}$$

ReLU faces a similar problem of 'dying ReLU' when the inputs to the ReLU function are negative. If the output of a neuron is zero, its gradient during backpropagation is also zero, which means that the weights associated with that neuron are not updated during training.

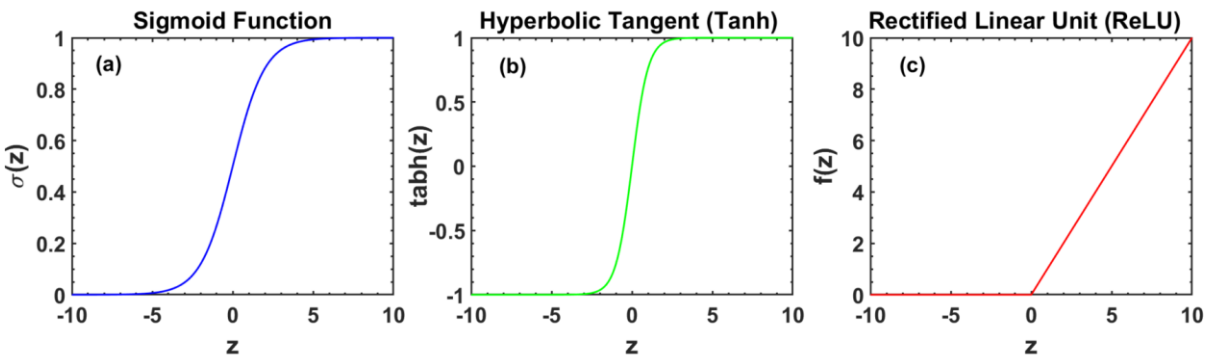

Figure 5.2: Commonly used activation functions: (a) Sigmoid Function, (b) Hyperbolic Tangent (Tanh), (c) Rectified Linear Unit (ReLU).

It is worth noting that activation functions can vary from layer to layer. Different layers in the network may use different activation functions depending on the specific needs of the model. Especially, the activation function of output layer depends on the tasks. For classification tasks, a commonly used activation functions is Softmax [121], which transforms the output of the network into a probability distribution over all possible classes, where the sum of all probabilities equals 1. For regression tasks, a linear activation function or no activation function is usually used, as they allow the output to be a continuous value.

The previous parts discussed the forward propagation in which predictions are made based on the input data. However, this process lacks the adjustment and optimization of weights and biases. In neural networks, this optimization is done through a process called backpropagation. Before introducing backpropagation, we need to define the loss function. The loss function is used to describe the difference between the predicted value and the actual value. In neural networks, backpropagation adjusts the weights and biases by minimizing the loss function. There are many types of loss function used to convert the difference between the predicted value and the actual value to a single value. The choice of loss function depends on the specific task at hand. For regression problems, the most commonly used loss function is the Mean Squared Error (MSE) given by

$$L_{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_{pred,i} - y_{target,i})^2 \quad 5.2.6$$

where n is the number of samples, $y_{pred,i}$ is the predicted value from neural networks, and $y_{target,i}$ is the actual target value. Another loss function, Mean Absolute Error (MAE), is also widely employed in regression tasks. For classification problem, the aim is to predict a discrete label or class. The most commonly used loss function to describe the difference between the predicted and actual probability distribution is cross-entropy loss. The definition of categorical cross-entropy for multi-class classification is written as

$$L_{cross-entropy} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{C} y_{ij}\log(\widehat{y_{ij}}) \quad 5.2.7$$

where C is the number of classes, y_{ij} is a binary indicator (0 or 1) if class label j is the correct classification for observation I, $\widehat{y_{ij}}$ is the predicted probability of class j for observation i.

After calculating the error using loss function, backpropagation is used to calculate the gradients of the loss function with respect to weights and biases in the network. The algorithm uses the chain rule to propagate these gradients from output layer to hidden layers and finally to the input layer [122]. Subsequently, the optimizer is used to update the weights and biases of the NN model based on the gradients computed during the backpropagation process. Common used optimizers include Stochastic Gradient Descent (SGD) [123] and Adaptive Moment Estimation (Adam) [124]. SGD is known for its efficiency and simplicity, especially when dealing with large datasets. It updates the weights and biases in each iteration according to

$$w' = w - \eta\nabla_w L \quad 5.2.8$$

$$b' = b - \eta\nabla_b L \quad 5.2.9$$

where w and b are the weights and biases, η is the learning rate, and $\nabla_w L$ and $\nabla_b L$ are the gradient of the loss function with respect to the weight and bias, respectively. Here, an important parameter batch need to be introduced. Batch refers to a subset of the training dataset used to update the weights of the model during training. In each iteration, SGD updates weights and biases based on a small batch, instead of the whole dataset. Therefore, this method helps to conserve computational resource. However, SGD may face difficult convergence problems.

Adam generally converges faster than SGD, because it can automatically adjust the learning rate. It integrates two powerful optimizers: momentum [122] and root mean square propagation (RMSprop) [125]. The process of Adam's parameter update is written as:

$$w' = w - \eta \frac{m_w}{\sqrt{v_w}+\varepsilon} \tag{5.2.10}$$

where m_w is the exponentially decaying average of past gradients, v_w represents the exponentially decaying average of squared gradients, and ε is a small constant to avoid division by zero.

Here is the training process of artificial neural networks. Additionally, one frequently discussed issue during training is overfitting. In artificial neural networks, the primary methods to prevent overfitting include regularization, dropout, early stopping, among others. Regularization [126] refers to adding constraints on the model parameters during training to improve the network's generalization ability. A common method is to add a regularization term to the original loss function:

$$L_{total} = L_{loss} + \lambda\Omega(\omega) \tag{5.2.11}$$

where λ controls the strength of the penalty and $\Omega(\omega)$ is a chosen norm of the weights. Commonly used regularizations include L1 and L2. L1 regularization uses $\Omega(\omega) = \sum_i |\omega_i|$. In back-propagation Its gradient adds a fixed push $\lambda sgn(\omega_i)$, so some weights are pushed to zero, leaving a sparse set of non-zero weights. L2 regularization uses $\Omega(\omega) = \sum_i \omega_i^2$. It add a term $2\lambda\omega_i$ to the gradient, gently pulling every weight toward zero and making the output smoother. Dropout improves generalization by randomly turning off a fraction of neurons during each forward pass, so the network cannot depend too much on any single neuron. Early stopping halts training once the validation loss stops falling, preventing the model from over-learning the training data and thus avoiding overfitting.

5.2.1.2 Convolutional Neural Network

A Convolutional Neural Network (CNN) [127] is a type of deep neural network widely used for processing data with spatial structures, such as images. CNNs have achieved significant successes across various fields, including computer vision and natural language processing [128]. CNNs typically consist of an input layer, convolutional layers, pooling layers, fully connected layers, and an output layer. Similar to fully forward neural networks, the input and output layers correspond directly to the desired inputs and outputs of the model. The convolutional layers use convolution kernels that slide over the input data to extract important features. Pooling layers reduce the spatial dimensions of these extracted features through operations such as max pooling [129]. This dimensionality reduction not only decreases computational complexity but also introduces translation invariance, enhancing the model's robustness to minor spatial shifts in the input. Fully connected layers then map the high-dimensional features obtained from the preceding convolutional and pooling layers into the desired output dimensions. These layers enable the network to perform classification or regression tasks effectively. Other common components of CNNs, such as activation functions, backpropagation algorithms, and hyperparameter optimization methods, are similar to those used in standard fully forward neural networks and thus will not be detailed here.

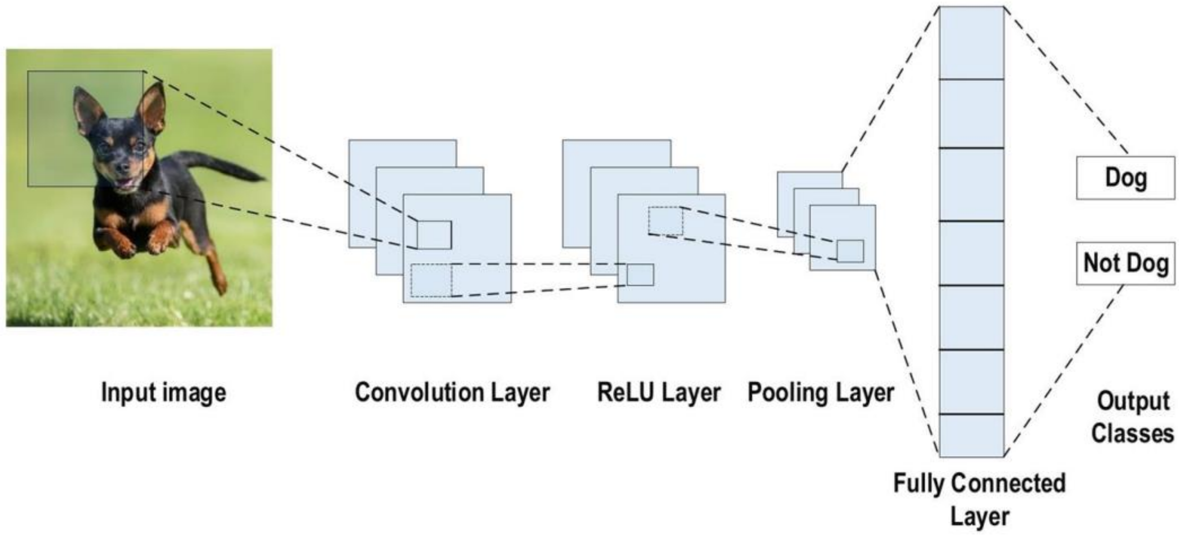

Figure 5.3: Typical architecture of a convolutional neural network designed for image-classification. Image from [130].

Figure 5.3 presents a typical CNN architecture designed for image classification tasks. As shown, the input image first undergoes feature extraction through a convolutional layer composed of multiple groups of convolutional kernels. Each kernel slides over the input image with a specified stride and performs element-wise dot product operations to capture spatial features. Padding is sometimes applied to prevent the loss of edge-related information. This convolution operation effectively extracts relevant features from the input image. After convolution, the resulting feature maps are passed through a nonlinear activation function such as ReLU, which introduces non-linearity into the model and enables it to learn more complex relationships. These activated features are then passed into a pooling layer, which reduces the spatial dimensions while retaining important information. Common pooling methods include max pooling [131] and average pooling [132]. This downsampling step not only improves computational efficiency but also introduces a degree of spatial invariance. The pooled feature maps are then flattened into a one-dimensional vector and passed through a fully connected layers, which map the high-dimensional features to the needed output dimension.

Compared to traditional fully forward neural networks, CNNs exhibit superior performance in image processing tasks due to their high accuracy and significantly reduced processing time. This efficiency arises from their ability to effectively extract and condense meaningful features from complex spatial data.

5.2.2 Hyperparameters optimization

Hyperparameter optimization is an important part of neural networks. This is because the performance of neural network models is not only related to the model architecture itself, but also depends on the selection of hyperparameters. Hyperparameters refer to parameters that need to be set in advance during model training. Compared with model parameters, hyperparameters will not be further updated during training, but are determined by manual or automatic search algorithms. The goal of hyperparameter optimization is to abandon the time-consuming process of manually selecting hyperparameter combinations and realize automatic or semi-automatic search for the optimal hyperparameter combination to improve model performance.

During the training process, the choice of hyperparameters significantly impacts model performance. Given the enormous data generated by EMC3-EIRENE simulations, manual hyperparameter tuning is practically infeasible due to its extensive time consumption. Therefore, advanced automated hyperparameter optimization algorithms such as Random Search [133], Asynchronous Successive Halving Algorithm (ASHA) [134], and Bayesian Optimization with Hyperband (BOHB) [135] are employed to efficiently and effectively identify optimal hyperparameter combinations. The optimized hyperparameters are subsequently used to build more accurate and efficient neural network models, enabling precise predictions.

5.2.2.1 Key hyperparameters

During the training of neural networks, the main hyperparameters include network architecture parameters (such as the number of layers, the number of neurons per layer, and the size and number of convolution kernels), learning rate, batch size, optimization algorithm, activation functions, and regularization parameters (such as dropout rate and regularization coefficients). Among them, network architecture parameters directly determine the complexity of the model. A simple network structure may fail to accurately capture features from complex data, while increasing the model complexity can enhance its learning capacity at the cost of higher computational demands. The learning rate controls the speed at which the model parameters are updated. An excessively large learning rate results in overly aggressive updates, which may cause instability in the loss and even prevent convergence. Conversely, a learning rate that is too small leads to very slow updates, which may cause the model to get stuck in a local minimum and significantly prolong the training process. Therefore, an appropriately chosen learning rate is crucial for both convergence and training stability. Batch size determines how many data samples are used in each parameter update. Smaller batch sizes generally offer better generalization capability but slow down the convergence of the loss function and increase computational cost. Larger batch sizes can improve training efficiency by leveraging parallelism but may reduce the model’s generalization performance. The optimization algorithm, such as stochastic gradient descent (SGD) [136] or Adam, governs how the model parameters are updated during training. Activation functions determine the model’s ability to capture nonlinear patterns in the data. To prevent overfitting, regularization techniques or dropout strategies are commonly applied.

Overall, the combination of hyperparameters is highly diverse, and manually tuning them is often time-consuming and impractical. Therefore, it is necessary to incorporate automated search algorithms and scheduling strategies to identify optimal hyperparameter configurations efficiently.

5.2.2.2 Search strategies

To effectively identify the optimal combination of hyperparameters, various automated search algorithms and scheduling strategies have been developed in recent years. The following introduces several commonly used schedulers in neural network hyperparameter tuning. Grid Search Scheduler is a classical approach, typically used when the number of hyperparameters is relatively small. It systematically explores all possible combinations through exhaustive enumeration. While this method theoretically guarantees finding the optimal configuration, it is extremely computationally expensive, especially as the dimensionality of the hyperparameter space increases. For models with more hyperparameters, a natural extension is to adopt a sampling-based approach, such as the Random Search Scheduler [133]. This scheduler randomly samples hyperparameter combinations from the search space. Although it does not guarantee a global optimum, it often provides a good trade-off between performance and efficiency, especially in high-dimensional spaces where full enumeration is impractical.

The introduction of the Hyperband Scheduler [137], based on the Bandit strategy, further improves search efficiency. It dynamically allocates training resources and progressively eliminates underperforming hyperparameter configurations, ensuring that computational efforts focus on the most promising trials. Building on Hyperband, the Asynchronous Successive Halving Scheduler (ASHA) [134] enables asynchronous execution of multiple trials. This allows early termination of poorly performing configurations and redistributes resources toward more promising candidates, significantly improving parallel efficiency and overall search speed. The Bayesian Optimization with Hyperband Scheduler (BOHB) [135] combines the strengths of Bayesian optimization and Hyperband. Bayesian optimization models the performance distribution over the hyperparameter space, guiding the search toward more promising regions. Combined with Hyperband, it is possible to allocate computing resources more efficiently and search the hyperparameter space more intelligently, which enables this method to achieve good model performance while paying less computing resources. Table 5.1 summarizes the advantages and disadvantages of these four schedulers. While each scheduler is suited to different scenarios, the incorporation of early stopping techniques across these methods has significantly accelerated the hyperparameter search process overall.

Table 5.1: Comparison of common hyperparameter scheduling strategies.

Scheduler	Advantages	Disadvantages	Best use case
Grid Search	Simple, exhaustive	Extremely expensive in high-dimensional spaces	Small search spaces
Random Search	Efficient in high dimensions	Cannot guarantee finding the global optimum	High-dimensional search spaces
ASHA	parallel execution, efficient resource use	performance may depend on search space size	Large-scale parallel tuning
BOHB	smart search, efficient resource use	Computational overhead from Bayesian modeling	Medium-to-large spaces with limited budget

5.2.3 Neural-network applications in fusion research

Neural-network models are valued for their strong ability to handle data and are usually used for three kinds of tasks: regression, classification, and pattern recognition. Regression networks predict continuous numbers from a set of input features. Before training, inputs and targets are scaled; common choices are feed-forward and convolutional networks, and losses such as mean-squared error or R-squared are used. In fusion research, regression networks, for example, act as fast surrogates for heavy edge-plasma simulations [115, 116] and give real-time predictions for diagnostics like divertor heat-flux profiles [138]. Classification networks choose one label from a fixed list. Data must be label-encoded first, and accuracy or similar scores are used for evaluation. Typical fusion uses include disruption prediction [111, 112] and automatic identification of operating regimes [139]. Pattern-recognition networks look for special structures or events in time-series data, with performance often measured by mean average precision. Examples are real-time optical boundary detection on EAST [140] and automatic detection of solitary edge bursts in toroidal devices [141].

As machine-learning techniques become more deeply integrated into fusion research, two main lines of progress are expected—one focused on experiments and the other on modelling. For experiments, machine learning is already improving plasma control, for example by providing real-time feedback on tokamak instabilities [110, 111] and regulating magnetic fields [109]; continued work should make operations more automated and far less dependent on operator experience. For modelling, fast surrogate models tied directly to diagnostic data will allow quick interpretation of experimental phenomena, while other surrogates will be used to explore untested operating spaces and to give first predictions for new devices. All in all, machine learning is making plasma control smarter, simulation analysis faster, and reducing the reliance on human expertise, providing a powerful tool for optimizing the operation and design of future fusion devices.

5.3 Bayesian inversion of EMC3-EIRENE inputs using edge diagnostics

5.3.1 Experimental diagnostics and dataset generation

W7-X is a quasi-isodynamic stellarator with five-fold symmetry and an island divertor. Its various magnetic configurations are created by tuning the currents in the planar and non-planar coil sets, and are typically identified by the edge rotational transform, $\iota_a = n/m$, where n and m denote the toroidal and poloidal mode numbers, respectively. ecause the EMC3-EIRENE dataset was built for the standard magnetic configuration ($\iota_a = 5/5$), the present study considers only discharges operated in that setting. To constrain the EMC3-EIRENE input parameters, four routinely available edge diagnostics are employed: (1) Divertor spectroscopy, recording C II emission at 426.8 nm in the divertor region; (2) Thomson scattering (TS), providing radial profiles of electron density and temperature; (3) Divertor Langmuir probes (LPs), measuring ion-saturation current on the targets; (4) Divertor thermography, capturing heat flux data on divertors. Together, these instruments characterise the scrape-off layer (SOL) both upstream and downstream. Figure 5.4(a) marks the positions of these diagnostics on W7-X, while Figure 5.4(b) displays the horizontal line of sight of the divertor-spectroscopy system at port AEI 51, spanning 1 cm to 14.5 cm above the horizontal target. Because the current EMC3-EIRENE physical model omits error fields (which are also not exactly known experimentally) and drift effects, it cannot reproduce the experimentally observed up-down and toroidal asymmetry. Hence, ion-saturation currents from the upper and lower Langmuir probes are averaged before comparison with simulations. A similar averaging is applied to the infrared-camera heat-flux data, selecting the cross-section that aligns with the probe locations.

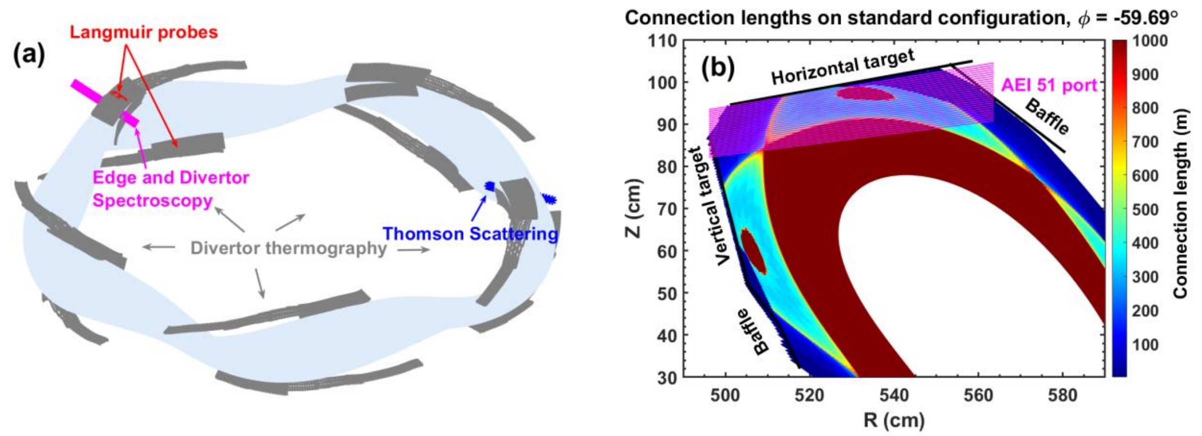

Figure 5.4: (a) Locations of the edge-diagnostic systems on W7-X, (b) horizontal line of sight (LOS) for the divertor spectroscopy system.

A training set of 260 EMC3-EIRENE runs in the W7-X standard configuration forms the basis for the neural-network surrogate. In these cases the radiation power fraction f_{rad}, separatrix electron density $n_{e,sep}$, and cross-field particle $D_{\perp}$ and thermal diffusivities $\chi_{\perp}$ are scanned over 0.2-0.8, $2-7\times10^{19}$ m^{-3}, 0.3-1.1 m²/s and 0.6-1.5 m²/s, respectively, while all other inputs (for example a fixed 6 MW of heating power) remain unchanged. The same cross-field particle diffusivities are assigned to hydrogen and carbon, and all transport coefficients are taken to be uniform in space. Carbon sputtered from divertor and baffle plates is the sole radiator; its source strength is adjusted to match the target total radiation power. For the FNN surrogate model, the inputs are the varied EMC3-EIRENE parameters, and the outputs are the synthetic signals produced in each simulation. As shown in Figure 5.4, we develop synthetic diagnostics for multiple edge diagnostic systems. Most diagnostics can be read directly from EMC3-EIRENE simulation outputs, but the divertor CII 426.8 nm emission is path-integrated and depends on electron temperature, electron density, and the line of sight of divertor spectroscopy. As explained in Section 3.2, the C II line-integrated radiance is obtained from the EMC3-EIRENE electron temperature and density profiles together with ADAS photon-emissivity coefficients (PECs):

$$L_{CII} = \sum_i \Delta s_i \sum_j \left(\omega_{ij} \frac{\sum_\sigma PEC^{(exc)}_{\sigma,k\to l} n_e n_\sigma^{z+} + \sum_\rho PEC^{(rec)}_{\rho,k\to l} n_e n_\sigma^{(z+1)+} + \sum_\rho PEC^{(CX)}_{\rho,k\to l} n_H n_\sigma^{(z+1)+}}{4\pi} \right) \quad 5.3.1$$

where L_{CII} is the line-emission radiance, the calibrated quantity routinely reported by spectroscopic diagnostics. Δs_i is the length of the i-th LOS segment, and ω_{ij} is the fraction of the overlap volume between grid cell j and LOS segment i to the total volume of segment i. PEC^{exc} , PEC^{rec} and PEC^{CX} are the photon emissivity coefficients for excitation, recombination and charge exchange progresses, respectively. n_e is the electron density, n_σ^{z+} is the density of the z-times-ionized ion in metastable state σ, n_H is the neutral-hydrogen density, ρ is metastable state, and k, l label the upper and lower atomic levels. The overlap factor ω_{ij} is estimated with a Monte Carlo sampling: N_i random points are thrown into LOS segment i; the ratio $\frac{N_{ij}}{N_i}$, where N_{ij} is the number of points falls inside grid cell j, yields ω_{ij}. This method avoids geometric intersection calculations. Because the various diagnostics span very different numerical ranges, the synthetic data need to be standardized or normalizied before training. Without this rescaling, features with large absolute values would dominate the gradients during back-propagation [142, 143], biasing the learning process and slowing or even preventing convergence. Here a simple max-value normalization is used: each feature is divided by its maximum across the data set, placing every variable in the [0, 1] interval. This prevents features with large absolute values from dominating the gradient updates and improves network convergence.

5.3.2 neural-surrogate-assisted Bayesian inference framework

This section sets out a two-stage scheme for quickly identifying EMC3-EIRENE input parameters that match the experimental diagnostic data. The procedure consists of (i) building and training a neural-network surrogate and (ii) performing Bayesian inference with Dynamic Nested Sampling. Figure 5.5 summarizes the end-to-end workflow and the network layout. In essence, the neural net provides an ultrafast proxy: given any candidate set of EMC3-EIRENE inputs, it returns the synthetic diagnostic data that a full simulation would generate. The Bayesian layer then compares these surrogate predictions with the measurements and, through the stated error bars, evaluates the likelihood of each candidate. The result is a probability-weighted score telling how well each input vector reproduces the experimental measurement data. Together, the two stages explore input parameter space rapidly yet still yield robust, uncertainty-aware estimates of the optimum inputs. Next, we would like to describe the two stages in detail.

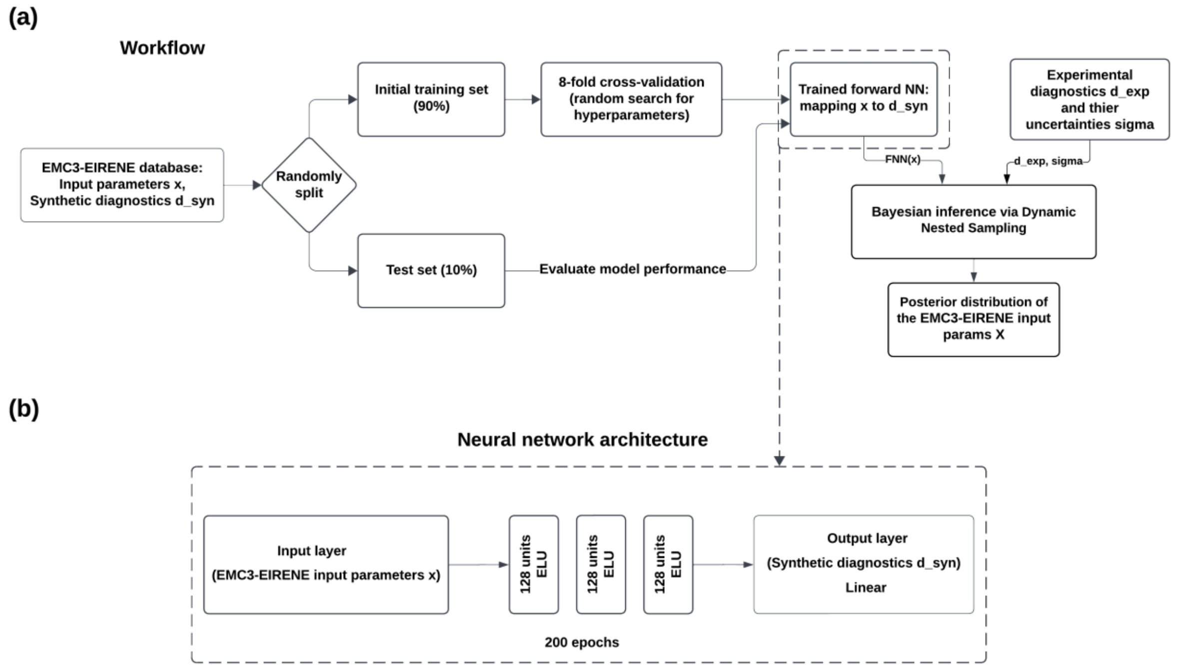

Figure 5.5: (a) Overall workflow of the surrogate-assisted Bayesian-inference method; (b) schematic of the feed-forward neural network.

5.3.2.1 Feed-forward neural network

The surrogate used is a feed-forward neural net composed of an input layer, several hidden layers and a linear output layer, with full connections between adjacent layers. The inputs are the EMC3-EIRENE input parameters; the hidden layers apply nonlinear transforms to mine higher-order features; the outputs are the synthetic diagnostic values. This network can capture the complicate relationships between the EMC3-EIRENE inputs and their associated synthetic diagnostic data. Before training, all quantities were scaled by their maximum values, and the full data set was shuffled and split into an initial training pool (90 %) and an independent test pool (10 %). Hyperparameters were tuned on the 90 % pool via eight-fold cross-validation: seven partitions for fitting and one for validation in each fold, with mean-squared error (MSE) as the loss. The average validation loss obtained from the eight-fold cross-validation served as the selection criterion. Figure 5.6 summarizes the 800 random-search trials: each trial's hyperparameter settings are shown, and the rightmost column records the corresponding average validation loss. The trial with the lowest bar in this column yields the optimal hyperparameter combination. Table 5.2 lists the search ranges explored in the 800 trials and the resulting best-performing settings. The main hyperparameters considered are the learning rate, number of hidden layers, number of units per hidden layer, batch size, L2 regularization strength, activation function, and total training epochs. In the random-search procedure, the learning rate and L2 coefficient were sampled on a logarithmic scale, whereas all other hyperparameters were drawn uniformly at random from their respective ranges. The final network has three hidden layers of 128 neurons, each using an ELU activation [144]; the output layer is linear. Other optimal settings are: L2 = 2.1×10^{-6}, learning rate = 0.0017, 200 epochs, Adam optimizer [124]. Figure 5.7 shows that both training (red) and validation (blue) losses flatten after ~100 epochs at $\approx$ 0.002, indicating that the net has learned the input-output mapping without over-fitting. With the hyperparameters fixed, the model was retrained on the full 90 % pool, then assessed on the hold-out 10 % pool, achieving an MSE $\approx$ 0.002.

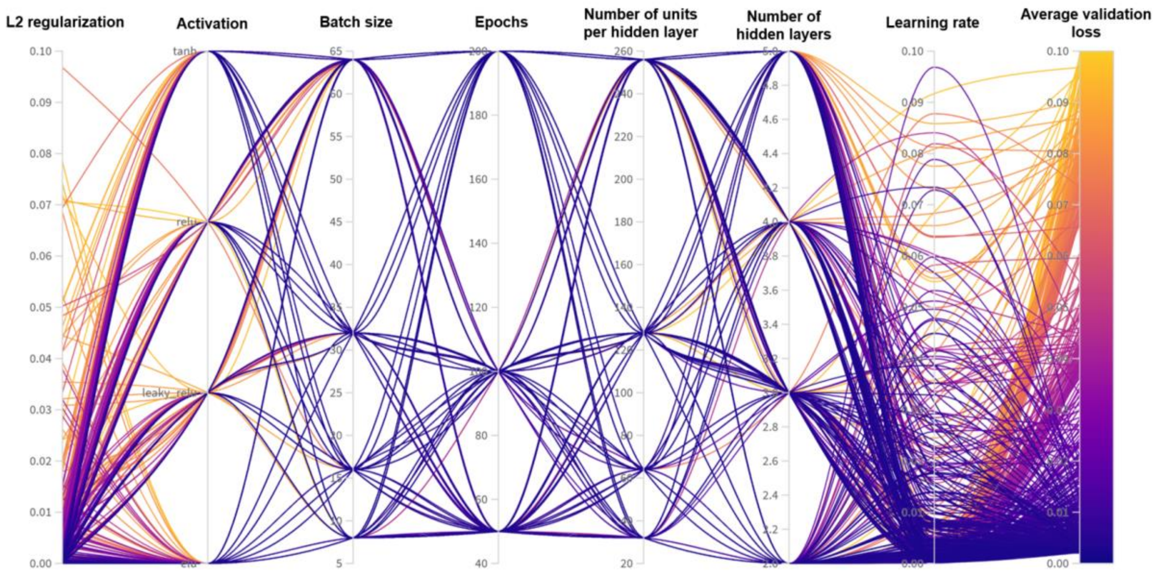

Figure 5.6: Hyperparameter settings for each of the 800 trials.

Table 5.2: Hyperparameter search ranges for neural network optimization and optimized values.

Hyperparameters	Search space	Optimized values
Learning rate	0.0001-0.1	**0.0017**
Number of hidden layers	2, 3, 4, 5	**3**
Units per hidden layer	32, 64, 128, 256	**128**
Batch size	8, 16, 32, 64	**16**
L2 regularization coefficient	10^{-6} - 10^{-2}	**2.1e-06**
Optimizer	Adam	**Adam**
Activation function	ELU, ReLU, Leaky, Tanh	**ELU**
Epochs	50, 100, 200	**200**

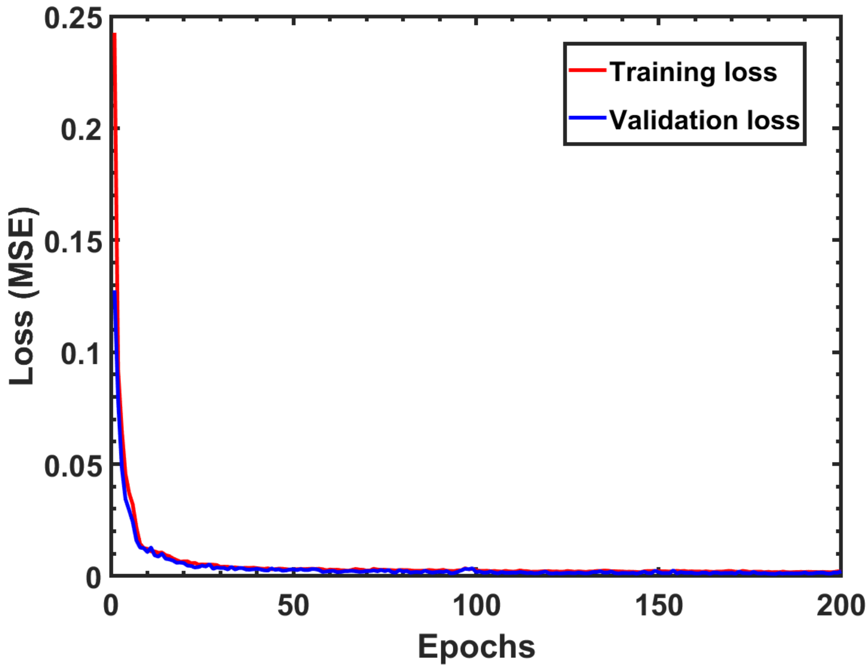

Figure 5.7: Mean training and validation losses during 8-fold cross-validation using the chosen hyperparameters.

A critical issue is whether the existing dataset is large enough to yield dependable predictions. To test this, the network architecture and the hyper-parameters found earlier were held fixed while only the trainset fraction, the share of data assigned to training, was varied. Figure 5.8 plots the test loss against the trainset fraction; for each fraction ten independent trainings were carried out. As the fraction rises from 0.05 to 0.95, the test loss steadily falls and eventually converges. To pinpoint the saturation point, the mean-loss curve was fitted with an exponential decay, and the extra drop that would result from enlarging the training set by an additional 5 % was calculated. When this incremental benefit is below 1 % of the current loss, the gain is regarded as insignificant. On that basis, convergence occurs at a training fraction of roughly 0.55. Beyond about 55 % of the complete dataset, further samples bring virtually no additional reduction, showing that the dataset is already dense enough across the sampled input-parameter domain. Moreover, as the training share increases, the scatter among the ten test-loss curves shrinks, indicating more consistent and stable performance (i.e. reduced sensitivity to the random split). To evaluate the model's accuracy after the test loss has converged, a probability-density function (PDF) of the relative prediction error was calculated for the 0.10 hold-out test set obtained with a trainset fraction of 0.9. As displayed in Figure 5.9, the PDF is close to Gaussian and peaks near zero, indicating virtually unbiased predictions. The probability of an absolute relative error larger than 10 % is below about 15 %, confirming that the surrogate's performance is satisfactory.

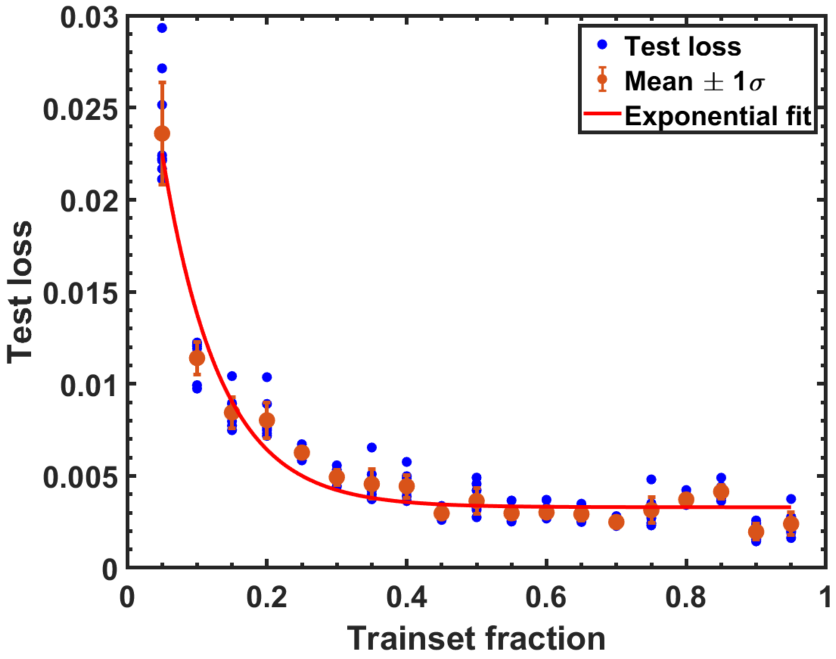

Figure 5.8: Test loss versus trainset fraction. Blue markers show the test-loss values from ten independent training runs at each train-set fraction; orange symbols (with error bars) give the mean ± 1 σ of those runs; the red line is an exponential fit to the mean losses.

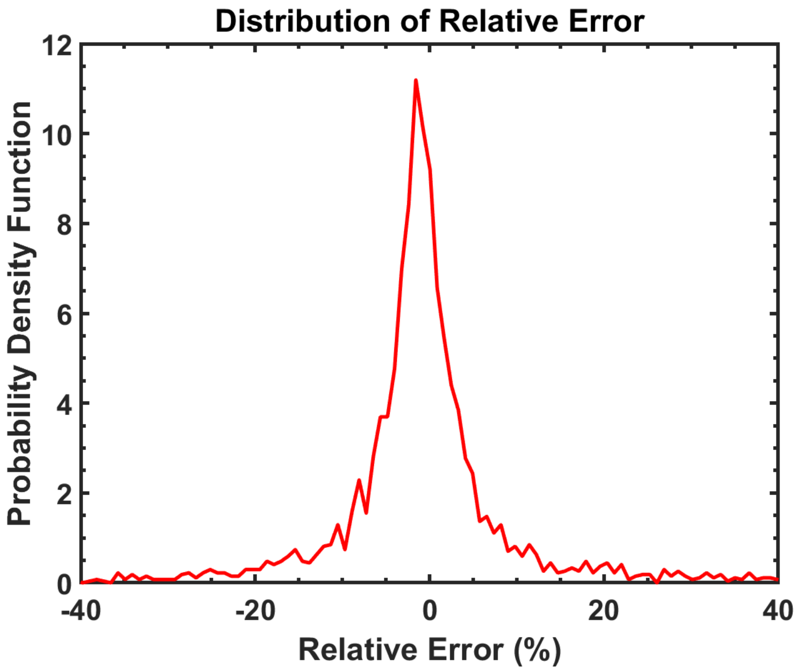

Figure 5.9: Probability-density distribution of relative error (%). The horizontal axis shows the relative error in percent, and the vertical axis gives the corresponding probability density.

5.3.2.2 Bayesian parameter estimation for determining best-matching EMC3-EIRENE input parameters

Once the FNN surrogate is trained, it is embedded in a Bayesian inference framework that employs Dynamic Nested Sampling [117] to determine the EMC3-EIRENE inputs most compatible with the measurements. Bayesian inference yields the posterior probability distribution of the input vector θ given the experimental data D and the surrogate model M. Bayes' theorem states

$$P(\theta|D,M) = \frac{P(D|\theta,M)P(\theta|M)}{P(D|M)} \quad 5.3.2$$

here $P(\theta|D,M)$ is the posterior of input parameters θ; $P(D|\theta,M)$is the likelihood, quantifying the match between surrogate-predicted and experimental signals for a specific θ; $P(\theta|M)$ is the prior, reflecting knowledge before seeing *D*; $P(D|M)$is the evidence, obtained via

$$P(D|M) = \int P(D|\theta,M)P(\theta|M)d\theta \quad 5.3.3$$

The search domain is widened beyond that of the original dataset to $f_{rad} = 0.1 - 0.9$, $n_{e,sep} = 1 - 8 \times 10^{19}\ m^{-3}$, $D_{\perp} = 0.1 - 1.3\ m^2/s$, $\chi_{\perp} = 0.3 - 1.8\ m^2/s$. Truncated-Gaussian priors are assigned to f_{rad} and $n_{e,sep}$ because experiments restrict their plausible ranges, whereas uniform priors are adopted for the cross-field particle diffusivity $D_{\perp}$and thermal diffusivity $\chi_{\perp}$. Surrogate predictions for inputs that lie just outside the original grid still agree well with EMC3-EIRENE runs, validating this moderate expansion (see Figure 5.12). In the likelihood, experimental uncertainties are taken to be Gaussian. Besides the four diagnostics already listed, the bolometric radiated power fraction $f_{rad,exp}$ is included to tighten the constraint on f_{rad}; omitting it allows f_{rad} to drift away from the measured value. Dynamic Nested Sampling (via the dynesty library [145, 146]) allocates samples adaptively according to the evolving likelihood and prior, evaluates the evidence, and returns accurate posteriors. Relative to classic Nested Sampling [147, 148], the dynamic variant uses samples more economically, improving both parameter estimation and evidence calculation.

5.3.3 Performance evaluation of the proposed method using experimental data

In this section, the practicality of the surrogate-assisted Bayesian framework is tested by comparing EMC3-EIRENE results with measurements from a density ramp-up discharge (#20181010.028) performed in the standard magnetic configuration. Because this shot includes several divertor operating regimes, it serves as a stringent benchmark for the method. Figure 5.10 plots the temporal histories of key plasma parameters. As the line-integrated electron density rises, the bolometer-derived radiated power follows suit. When the density reaches the detachment threshold ($\sim 11.7 \times 10^{19} m^{-2}$) and stays there, the radiated power continues to climb, eventually establishing a stable detached regime. After 6.5 s, hydrogen fueling through the divertor nozzles is switched off, the density begins to fall, and the radiation level steadily diminishes. Three times (t1, t2, t3) are selected corresponding to radiated-power fractions f_{rad} = 20 %, 50 %, and 80 %, to test the method under distinct plasma scenarios.

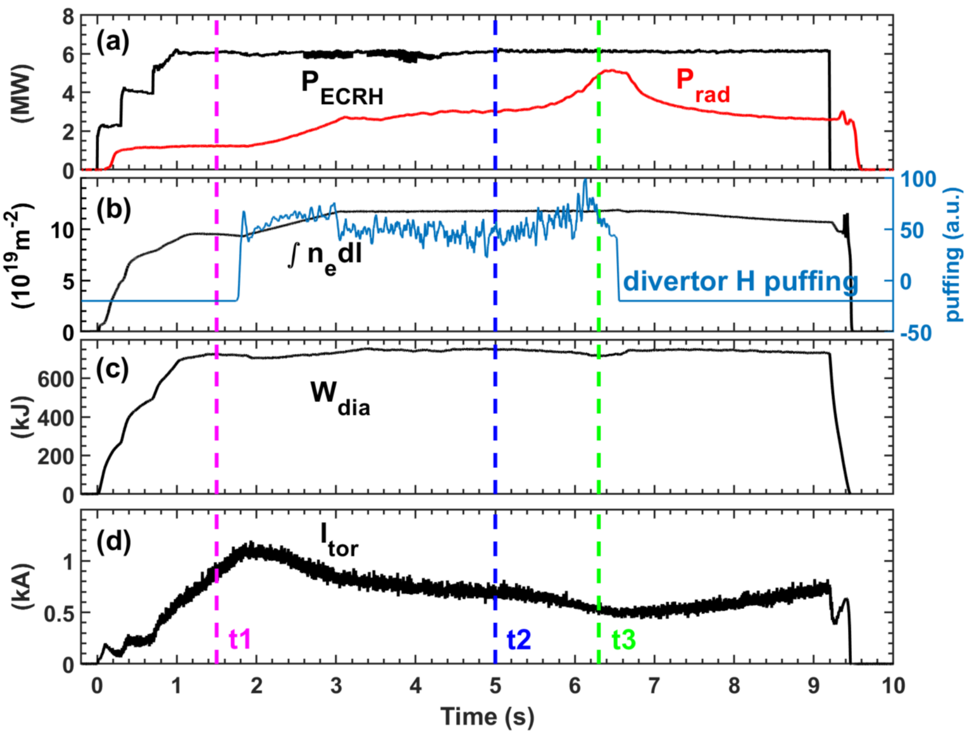

Figure 5.10: Time traces of key W7-X parameters for density ramp-up discharge #20181010.028. Vertical dashed lines mark the three evaluation time points.

Before performing the Bayesian inference, uncertainties for every diagnostic are quantified. Figures 5.12-5.14 compile detailed profiles at each chosen f_{rad}. Because the W7-X Thomson scattering system is not optimized for edge measurements, its edge error bars are sizeable. Five consecutive TS snapshots (30 Hz) around each selected time are averaged. Total TS uncertainty σ_{TS} combines measurement error and short-time variability $\sigma_{TS} = \sqrt{\sigma_{measurement}^2 + \sigma_{time}^2}$, where $\sigma_{measurement}$ is the mean stated error of the five shots and σ_{time} is their standard deviation. For target ion-saturation current and heat-flux signals, large upper-lower asymmetries are very larger than instrumental noise, so the standard deviation of those asymmetries is taken as the uncertainty. For the divertor C II radiance, the uncertainty has not been precisely quantified; a conservative 30 % relative uncertainty is therefore assumed. Similarly, the bolometer-derived radiation fraction uncertainty is set at 20%.

Figure 5.11 displays corner plots of the posterior distributions for EMC3-EIRENE input parameters (f_{rad}, $n_{e,sep}$, $D_{\perp}$ and $\chi_{\perp}$) obtained using the neural-surrogate-assisted Bayesian inference method under the three selected radiation fraction conditions. Diagonal plots show probability distribution histograms, with the central dashed lines indicating median parameter values and side dashed lines marking one-standard-deviation (σ) uncertainty intervals. The specific median values and corresponding uncertainties (+σ and −σ) for each parameter are explicitly indicated above each histogram. Below the diagonal, joint distributions between parameter pairs are illustrated. Contours close to circular indicate weak correlations between corresponding parameters, whereas elliptical contours with distinctly inclined major axes suggest stronger linear correlations. The joint-posterior panels are almost perfectly circular, indicating that the four input parameters are only weakly correlated with one another. A noteworthy trend appears in the cross-field transport coefficients: the particle diffusivity $D_{\perp}$ steadily falls as the radiation fraction increases, whereas the thermal diffusivity $\chi_{\perp}$ decreases at intermediate radiation levels and then rises again once deep detachment is established.

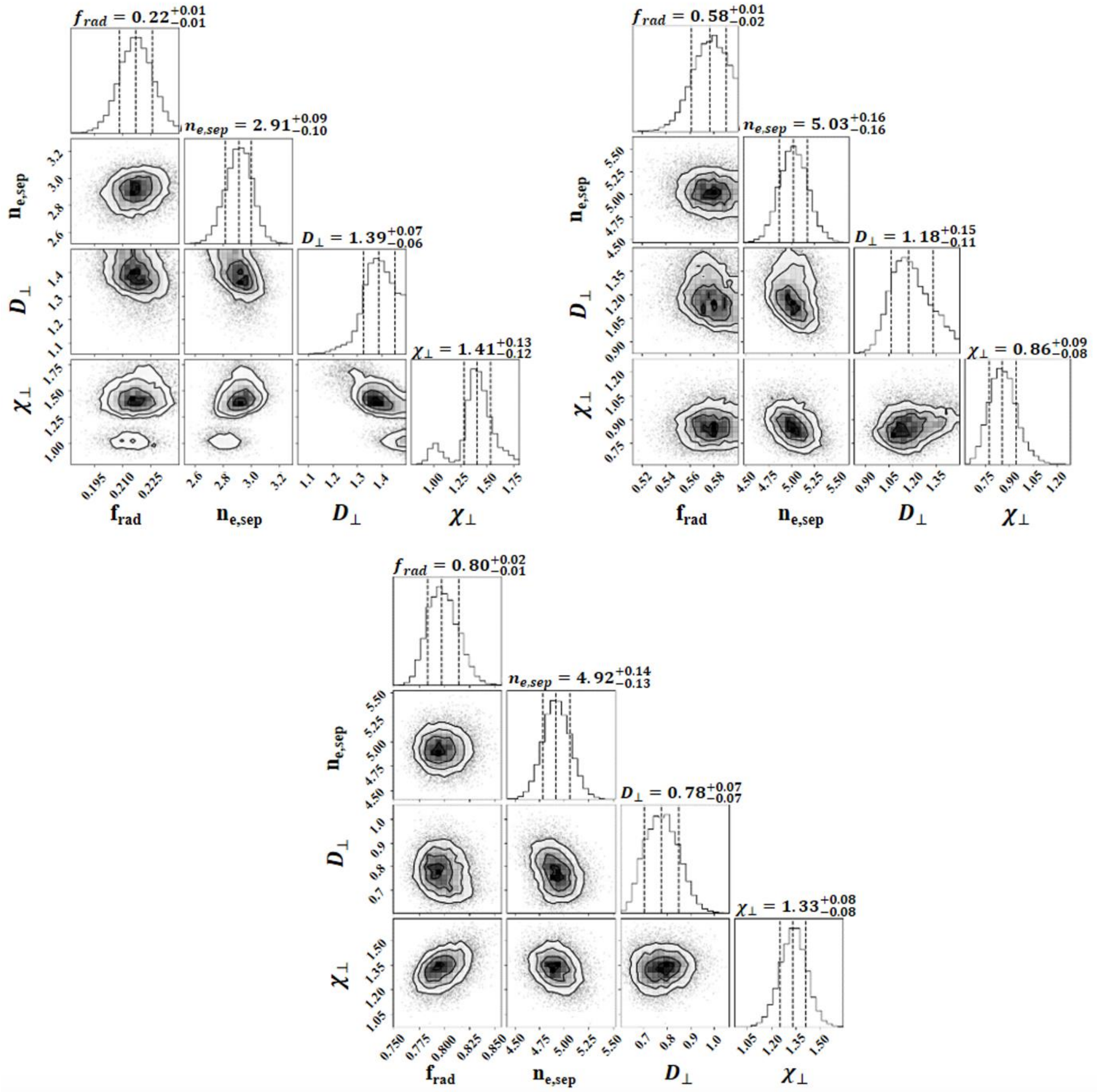

Figure 5.11: Corner plots of posterior distributions of EMC3-EIRENE input parameters at different radiation fractions.

To further validate these predictions, EMC3-EIRENE simulations were conducted using the input parameters corresponding to the maximum a posteriori (MAP) estimate, which is the parameter set with the highest posterior likelihood, and the simulation results were then compared with the experimental data. For radiation fractions of approximately 20%, 50%, and 80%, the MAP estimates were determined to be $f_{rad} = 0.23, 0.60$, and 0.79; $n_{e,sep} = 2.92, 5.00$, and $4.87 \times 10^{19} m^{-3}$; $D_{\perp} = 1.34, 1.09$, and 0.82 m²/s; $\chi_{\perp} = 1.37, 0.82$, and 1.34

m²/s. Figure 5.12 compares EMC3-EIRENE simulations, FNN surrogate predictions, and experimental data for the 20% radiation fraction scenario, showing profiles of electron density n_e, electron temperature T_e, horizontal divertor heat flux at a toroidal angle of -82°, divertor ion saturation current j_{sat} from divertor Langmuir probes, and divertor CII radiance measured by the AEI 51 spectroscopy system. In each plot, black dots or solid lines represent experimental measurements, blue crosses indicate predictions from the FNN surrogate model, and red lines correspond to EMC3-EIRENE simulation results obtained using the MAP estimates. The FNN predictions and the EMC3-EIRENE simulations agree well, demonstrating the FNN surrogate model's good performance in capturing the relationship between the EMC3-EIRENE input parameters and synthetic diagnostic data. Additionally, the EMC3-EIRENE results match the experimental measurements satisfactorily for n_e, T_e, j_{sat}, and CII radiance, although the simulated heat flux has a broader profile than the observed data.

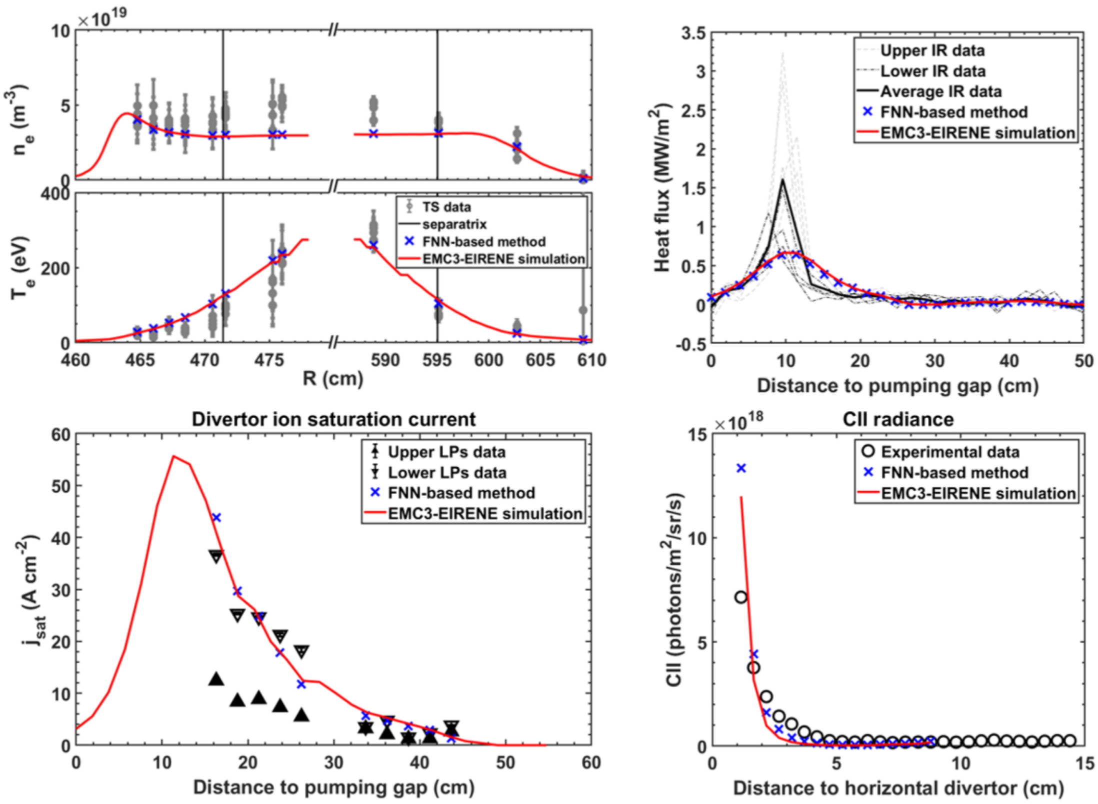

Figure 5.12: Comparison between experimental data, FNN surrogate predictions, and EMC3-EIRENE simulations using MAP input parameters at 20% radiation fraction.

For the scenario with a 50% radiation fraction, Figure 5.13 shows that both the target heat flux and ion saturation current decrease slightly, accompanied by a reduction in asymmetry. The surrogate model maintains strong agreement with EMC3-EIRENE simulations, further validating its applicability at intermediate radiation levels. Comparison of EMC3-EIRENE simulations with experimental data shows good agreement for TS n_e, T_e, divertor heat flux, and divertor ion saturation, but a discrepancy is observed in the spatial location of the CII radiance peak. At an 80% radiation fraction, divertor heat flux and ion saturation current decrease significantly, and impurity ion penetration increases, shifting the CII radiance further away from the target. Figure 5.14 shows a slight increase in the difference between the EMC3-EIRENE simulations and FNN predictions, which may be attributed to relaxed convergence criteria for higher radiation fraction cases in the training dataset (where a relative change of less than 1% is considered converged, compared to less than 0.7% for lower f_{rad} cases). Similar to the 50% f_{rad} scenario, the spatial mismatch between simulated and experimental CII radiance peaks persists, aligning with observations reported in previous studies [71]. One possible explanation for this discrepancy is the assumption of uniform cross-field transport coefficients in EMC3-EIRENE. In reality, spatial variations in these coefficients can modify the balance between parallel and perpendicular transport, thereby influencing impurity distributions [43, 149]. Another potential factor might be the absence of drift effects in EMC3-EIRENE. Significant radiation asymmetry is observed in experiments [45, 150], but the lack of drift effects in EMC3-EIRENE may prevent the simulation from reproducing the CII radiance. Overall, within the capabilities of EMC3-EIRENE, the proposed method achieves satisfactory agreement between simulated and measured diagnostics across different radiation fractions. Nevertheless, comprehensive validation of synthetic divertor spectroscopy predictions requires further investigations.

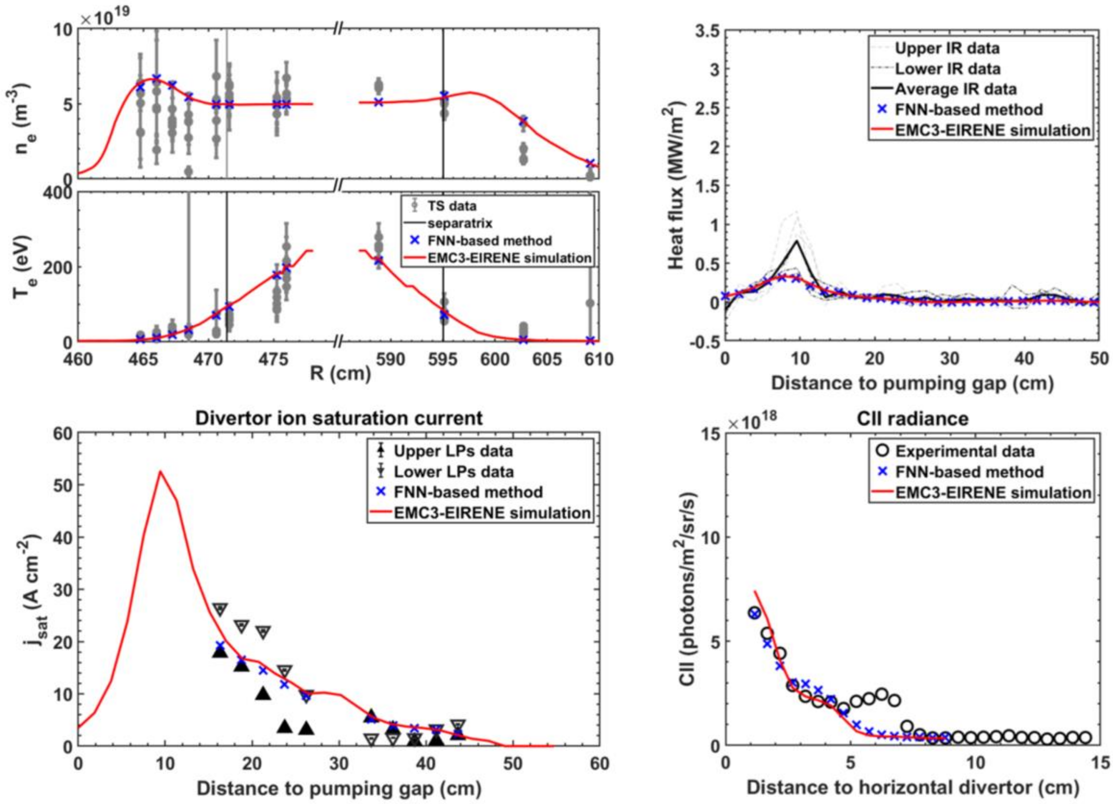

Figure 5.13: Comparison between experimental data, FNN surrogate predictions, and EMC3-EIRENE simulations at 50% radiation fraction.

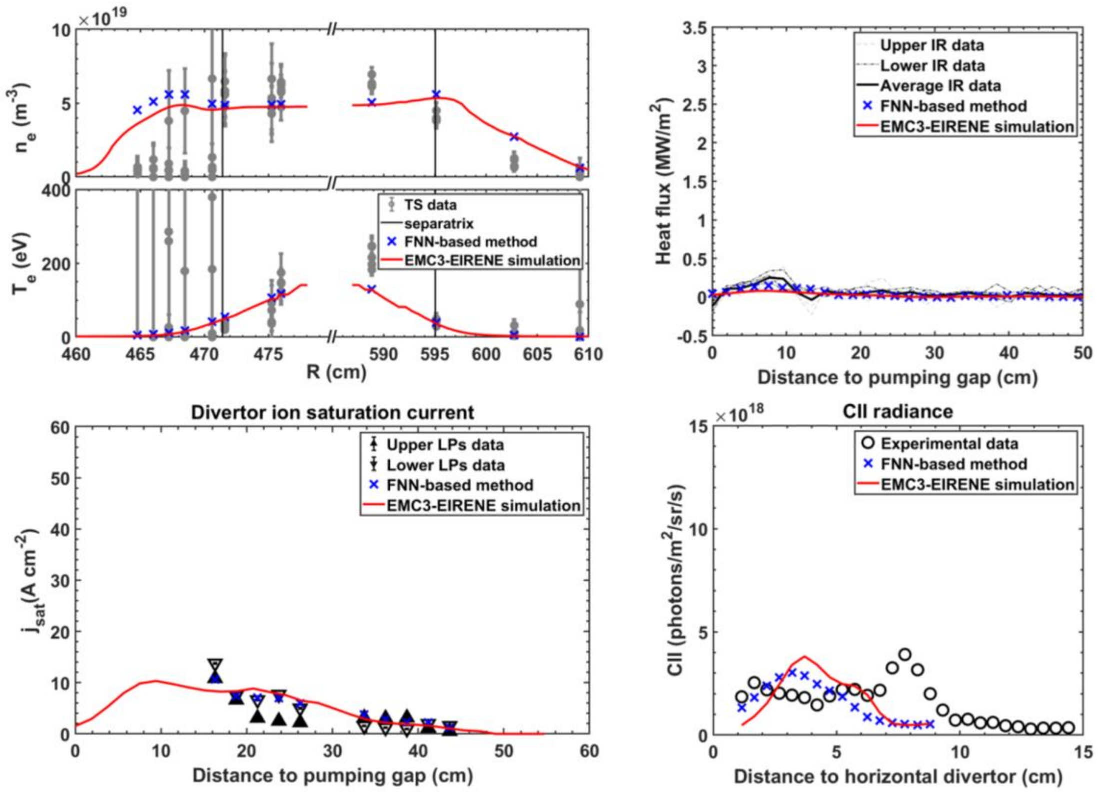

Figure 5.14: Comparison between experimental data, FNN surrogate predictions, and EMC3-EIRENE simulations at 80% radiation fraction.

In short, the neural-network surrogate gives rapid predictions across the input-parameter space, and the Bayesian framework compares those predictions with the experimental data while fully accounting for the measurement uncertainties. Together, these two tools make it possible to find the EMC3-EIRENE inputs that best match the observations, especially the otherwise unknown free parameters cross-field transport coefficients. Full EMC3-EIRENE simulations conducted with MAP estimates of input parameters at different radiation fraction levels showed satisfactory agreement with experimental observations, highlighting the effectiveness of the proposed method. This work is focused on developing and validating the method's feasibility. The current EMC3-EIRENE dataset is still being expanded and therefore does not yet cover the full W7-X operating space, for example the complete range of heating powers, and different magnetic configurations. Future work will expand the dataset and extend the range of synthetic data, thereby broadening the applicability of the neural network-based method.

5.4 Summary

A neural-surrogate-assisted Bayesian inference approach is proposed to rapidly optimize the EMC3-EIRENE input parameters so that the resulting synthetic diagnostics closely match W7-X measurements. The optimized inputs include not only free input parameters like cross-field transport coefficients, but also corrections to experimental quantities with relatively large error bars, such as the separatrix electron density and the total impurity radiation power. Regarding the detailed work, a feed-forward neural network is first trained on an EMC3-EIRENE dataset to provide an instant mapping from input parameters to synthetic diagnostic signals corresponding to experimentally observed physical quantities. This surrogate is then embedded in a Bayesian framework that uses Dynamic Nested Sampling to obtain the posterior distributions of the input parameters. The likelihood compares the surrogate-predicted signals with the experimental data and explicitly includes the measurement uncertainties. Running EMC3-EIRENE with the resulting maximum-a-posteriori parameters reproduces the diagnostics with good fidelity. The approach greatly reduces the computer time required, eliminates labor-intensive manual tuning, and can be transferred to other edge-transport codes.

6 Machine-learning accelerated edge plasma and impurity transport analysis

This chapter presents the third main contribution of the thesis: using machine-learning techniques to speed up edge transport studies. As shown in Chapter 4, the conventional workflow to study edge-transport physics is to constrain EMC3-EIRENE with diagnostic data through manual iterative adjustments of the input parameters. Chapter 5 introduced a Bayesian neural-network framework that rapidly infers inputs consistent with the measurements (and their uncertainties), but a full EMC3-EIRENE run is still required to obtain the 3D plasma parameters for transport analysis. On W7-X, one reasonably fine-resolution run typically costs about 2000 CPU core-hours.

To replace full runs, an EMC3-EIRENE neural-network surrogate is developed that maps inputs to full-code outputs. The combination of the surrogate and the Bayesian neural network inference from Section 5.3 yields EMC3-EIRENE-AI, an end-to-end pipeline that produces fast, code-consistent edge-plasma predictions directly from multiple diagnostic measurements. The workflow comprises three steps:

1. Diagnostic measurements → EMC3-EIRENE input inference

   A neural-surrogate-assisted Bayesian framework uses the experimental signals (with their uncertainties) to infer the most probable EMC3-EIRENE input parameters.

2. EMC3-EIRENE input → output prediction

   A fine-tuned neural-network surrogate replaces the full code, predicting 3D EMC3-EIRENE outputs within a few milliseconds

3. EMC3-EIRENE-AI

   Combining steps 1 and 2 closes the loop, producing near-instant plasma reconstructions that remain consistent with both the measurements and the underlying EMC3-EIRENE physics model.

With this approach, an edge-impurity transport analysis that used to consume several thousand core-hours can now be completed on a desktop computer in only a few minutes. The sections that follow detail steps 2 and 3.

6.1 Deep-learning surrogate for 3D edge-plasma transport on the W7-X

This section presents a neural-network surrogate for W7-X edge transport simulations, trained on an EMC3-EIRENE dataset that spans nearly the entire operating-parameter space of the W7-X standard configuration. The model first uses an autoencoder to compress EMC3-EIRENE outputs into a low-dimensional latent space representation, then a neural regressor maps EMC3-EIRENE input parameters to those latent vectors, and finally both components are fine-tuned together to predict outputs directly from inputs. Hyperparameters are tuned with the Asynchronous Successive Halving Algorithm (ASHA), whose early-stopping mechanism discards unpromising trials and shortens the optimization process. Benchmark tests show that the surrogate clearly outperforms a conventional multilayer perceptron, particularly in detached-plasma conditions. Across all test cases, only 7.25 % of grid cells display an absolute relative error larger than 10 % when compared with the full EMC3-EIRENE simulations. For example, in predicting the electron temperature T_e, only 7.25% of grid cells across all test cases exhibit an absolute value of the relative error greater than 10% when compared with full EMC3-EIRENE simulations. The cells with >10% relative error lie mainly at the outermost edge, where T_e is very small. Consequently, although the relative error can exceed 10% there, the absolute error remains small. Leave-one-value-out evaluation indicates high accuracy within the interpolation domain, with minor degradation when extrapolating. Relative to full EMC3-EIRENE runs, the surrogate provides over a 10^8- fold speedup, enabling large-scale parameter scans or real-time feedback control based on 3D transport simulations to become feasible in the future.

6.1.1 Importance of fast 3D edge-plasma transport prediction

The W7-X is an advanced stellarator optimized for neoclassical transport, fast particle confinement and magnetohydrodynamic (MHD) stability, whose intrinsic three-dimensional (3D) magnetic field structure imposes high demands on simulation. The 3D edge transport code EMC3-EIRENE, which couples the fluid model EMC3 with the kinetic neutral particle transport code EIRENE, is well suited for handling such geometries and its application has substantially deepened our understanding of impurity transport, radiative power and particle exhaust in W7-X. However, the complexity of the EIRENE simulation leads to significant hardware demands, limiting its application in scenarios where real-time feedback or large-scale parameter scans are needed. Although simplified codes such as the diffusive field line tracing (DFLT) [151] and EMC3-Lite [152] offer improved computational efficiency, omitting impurity transport and particle-neutral interactions makes them unsuitable for high-density plasma simulations.

To address this problem, machine learning technology provides a promising solution by learning input-output relationships from a precomputed dataset, enabling rapid output predictions. In fusion research, several approaches have been recently developed to accelerate related simulations. One strategy is to replace full solvers with direct input-output mappings, as shown by one-dimensional codes DIV1D [153, 154], 1D UEDGE [155, 156], and the two-dimensional code SOLPS-ITER [40, 116, 157]. Another approach focuses on accelerating specific time-consuming steps of a workflow [115, 158, 159]. As is well known, if the upfront cost of constructing a simulation dataset is acceptable and cost-effective, machine learning methods that can rapidly and accurately provide predictive results consistent with the underlying physics will offer significant value.

Here, we construct neural network-based surrogates for the computationally intensive 3D EMC3-EIRENE code. For the complex island divertor configuration of W7-X [27, 160], a simple fully connected MLP trained on converged EMC3-EIRENE simulations cannot adequately capture the nonlinear input-output relationship, resulting in non-negligible errors when predicting detached conditions. We therefore adopt an improved architecture [154, 156, 161]: an autoencoder first compresses EMC3-EIRENE outputs into latent space, an MLP then projects the inputs into this latent vectors, and the two networks are finally fine-tuned together. The resulting surrogate achieves markedly higher accuracy than the simple MLP, with relative errors against EMC3-EIRENE simulations kept below 10%. The jointly fine-tuned network excels at interpolation, maintaining high accuracy within the training domain, but its performance declines when applied to extrapolation.

The rest of this section proceeds as follows. First, the simulation dataset, network architecture, and hyperparameter-tuning procedure are described. Next, the fine-tuned surrogate is assessed in terms of prediction accuracy, interpolation and extrapolation capability, and computational efficiency.

6.1.2 Data and Methods

6.1.2.1 Dataset generation

A comprehensive dataset was generated using EMC3-EIRENE code. In these simulations, carbon sputtered from divertor targets or baffles was specified as the sole radiation source, with its source strength controlled by the total radiated power. Cross-field particle diffusivities for hydrogen (bulk plasma) and carbon were assumed equal and spatially uniform, as were the electron and ion heat diffusivities. Five primary input parameters were systematically scanned in the EMC3-EIRENE simulations under the standard configuration of W7-X: input heating power across the innermost boundary surface (P_{in}), radiation power fraction (f_{rad}), electron density at the separatrix ($n_{e,sep}$), cross-field particle diffusivity $D_{\perp}$ and cross-field thermal diffusivity $\chi_{\perp}$. The ranges of these parameters, as shown in Table 6.1, were selected to cover typical operational window of W7-X in its standard configuration. As we known, at constant heating power, increasing the plasma density drives the edge from low to high recycling and eventually into detachment. Figure 6.1 illustrates this trend in EMC3-EIRENE by tracking the total recycling flux during a density ramp with constant carbon sputtering yield of 0.04. To ensure the surrogate captures this transition, the separatrix electron density was sampled more densely in the range 3-$6 \times 10^{19}\, m^{-3}$ as shown in Table 6.1. Conducting EMC3-EIRENE simulations with exhaustive sampling across the entire parameter space defined in Table 6.1 would be prohibitively expensive computationally and unnecessary in practice. We found that, beyond a certain dataset size, adding further samples yielded only marginal improvements in model performance. Thus, 463 input sets were chosen at random from this parameter space and each was run to numerical convergence with EMC3-EIRENE, yielding the converged-simulation dataset. The suitability of this sample size for neural-network training is assessed later.

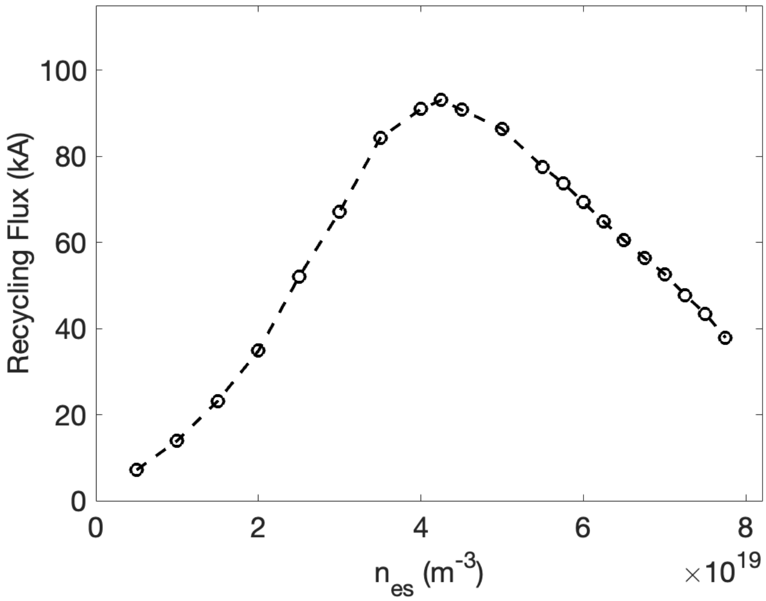

Figure 6.1: Total recycling flux versus separatrix electron density for simulations with $P_{in} = 10\text{MW}, D_{\perp} = 0.5\text{m}^2/\text{s}, \chi_{\perp} = 1.5\text{m}^2/\text{s}$, and a carbon sputtering yield of 0.04.

EMC3-EIRENE produces a wide array of quantities, including electron temperature, electron density, ion temperature, radiation distribution, divertor heat load and so on. In the present study, a separate neural network is trained for each individual output. The following pages focus on two key outputs, namely electron temperature and electron density, and evaluate the surrogate's predictive accuracy for each. The EMC3-EIRENE simulation dataset was constructed using a 3D numerical mesh with a resolution of 144 × 256 × 144 cells in the radial, poloidal, and toroidal directions, respectively. To facilitate interpretation, a toroidal slice was extracted, producing a 2D 144×256 (radial × poloidal) cross-section at the bean-shaped location on W7-X, which serves as the representative output discussed below. All inputs and outputs were scaled with independent min-max normalization to promote stable and rapid training: For every case, each of the five input parameters and every grid cell of outputs was transformed by subtracting its global minimum (over the 463 simulations) and dividing by the corresponding range (maximum minus minimum).

Table 6.1: Range of EMC3-EIRENE input parameters.

Parameters	Ranges								
P_{in} (MW)	4		6		8		10		
f_{rad}	0.2		0.4		0.6		0.8		
$n_{e,sep}$ (10^{19} m^{-3})	2	3	3.5	4	4.5	5	5.5	6	7
$D_\perp$ (m^2/s)	0.1	0.3	0.5	0.7	0.9	1.1	1.5	2	
$\chi_\perp$ (m^2/s)	0.2		0.6		0.9		1.2		1.5

6.1.2.2 Training Workflow

Figure 6.2 display the training workflow of the fine-tuned surrogate. After min-max normalization, the 463 converged EMC3-EIRENE cases are randomly divided into a training set (80%) and an independent test set (20%). The training set is used in all subsequent network-training stages, while the test subset is reserved for the final performance assessment. The surrogate network training includes three stages: (1) autoencoder pre-training, (2) inputs-to-latent mapping, and (3) joint fine-tuned of the combined networks. Hyperparameters are selected by searching the training set with five-fold cross-validation and the ASHA early-stopping scheduler, and the optimal hyperparameter settings were then applied to complete the training. The next subsections describe the network architecture and the hyperparameter optimization in detail.

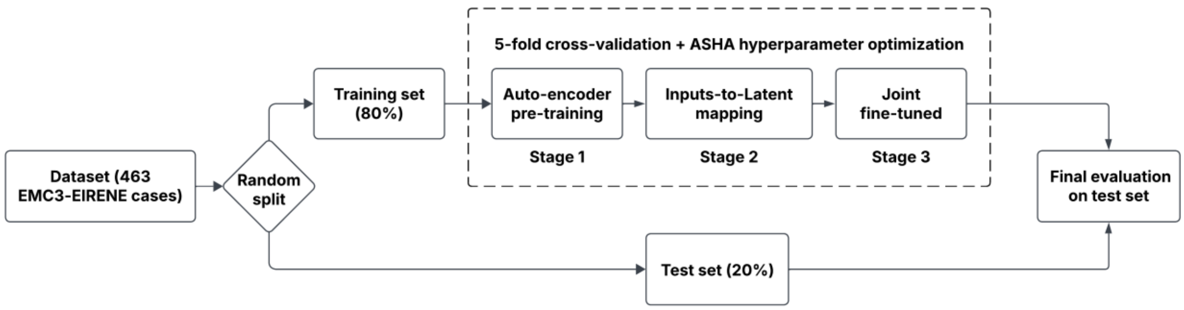

Figure 6.2: Training workflow for the autoencoder-based surrogate with cross-validation and ASHA optimization.

The neural network architecture for each phase is illustrated in Figure 6.3. The surrogate aims to efficiently map five input parameters of EMC3-EIRENE to the corresponding simulation outputs.

Stage 1. Autoencoder pre-training

The autoencoder consists of an encoder-decoder pair. The encoder compresses high-dimensional simulation outputs into a low-dimensional latent representation, whereas the decoder reconstructs these outputs from a compact latent representation. The aim is to learn an information-rich latent space that captures the dominant spatial structure of the outputs. With the learned representation, Stage 2 predicts a compact latent vector from the five inputs instead of full outputs, simplifying the complex mapping task. The loss function used for autoencoder pre-training is the mean squared error (MSE).

In this work, the encoder contains four down-sampling residual blocks. Each block applies a 3×3 convolution with stride 2 and padding 1, followed by batch normalization and a two-layer residual refinement before the shortcut is added. After the four blocks, the spatial resolution is reduced from 144×256 to 9×16, while the channel width increases from 1 to 128. The resulting 128×9×16 tensor is flattened and transformed into a compact latent vector by a fully-connected layer, whose size is optimized through hyperparameter tuning. This latent vector significantly compresses the data dimensions while preserving the essential spatial structure.

The decoder can be regarded as the inverse transformation of the encoder. It first reshapes the latent vector back to 128×9×16 with a fully connected layer, then uses four up-sampling residual blocks. In each block a transposed convolution doubles the spatial dimensions and decreases the channel width, restoring the resolution to 144×256 while reducing the channels from 128 to 1.

Stage 2. Input-to-latent mapping

The mapping network is a fully connected feed-forward neural network translating the five EMC3-EIRENE input parameters into the latent vector produced by the encoder. The encoder is frozen during this stage. A residual branch projects the original 5 inputs directly to the first hidden representation and is added to the activated output, retaining the linear part of the mapping. The subsequent hidden layers use identity skip connections, which help gradients flow and stabilize training. A final linear layer produces the latent vector. Key hyperparameters such as the number of hidden layers L, the number of neurons in each layer, and the choice of activation function are selected through the hyperparameter search described below. The loss function is a latent space MSE between the mapping output and the encoder-derived latent vector.

Stage 3. Joint fine-tuned

In the final stage, the inputs-to-latent mapping network and the decoder are chained to form an end-to-end surrogate that predicts the EMC3-EIRENE outputs from the 5 input parameters. Specifically, the mapping network generates latent vectors based on the 5 EMC3-EIRENE inputs, which subsequently serve as inputs to the decoder. The decoder then reconstructs the simulation outputs through its up-sampling structure. Training minimizes the MSE between the surrogate prediction and the reference simulation, updating both the mapping network and the decoder.

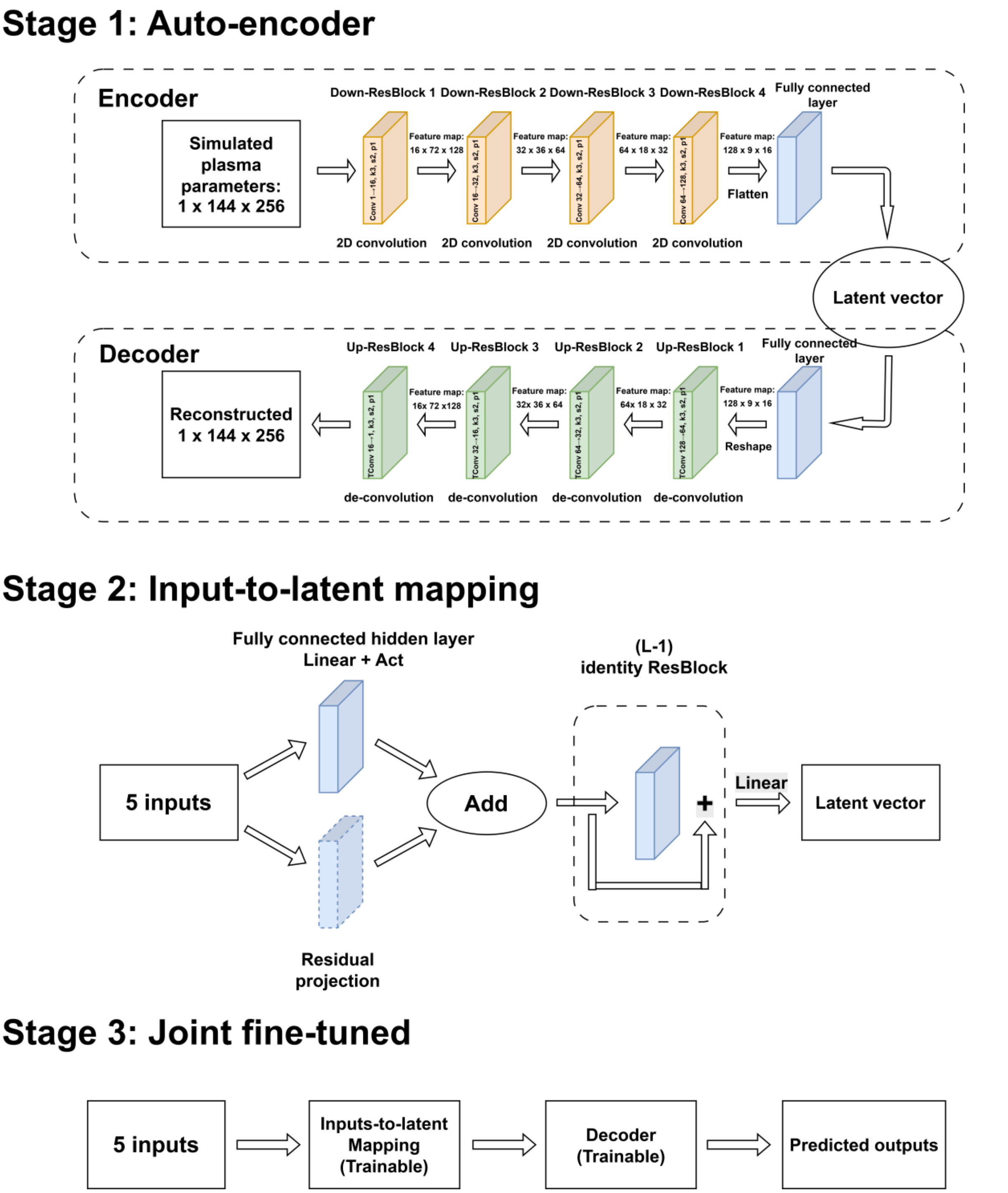

Figure 6.3: Architecture of the three-stage (a) autoencoder pre-training, (b) input-to-latent mapping, (c) joint fine-tuned.

For comparison, we replace the three stages with a single MLP and train it using the same procedure. The architecture of this reference MLP, illustrated in Figure 6.4, comprises an input layer, hidden layers, and an output layer. Every neuron in a layer connects to every neuron in the next, and nonlinear activation functions supply the required nonlinearity. The MLP's hyperparameters are also determined through the same optimization process.

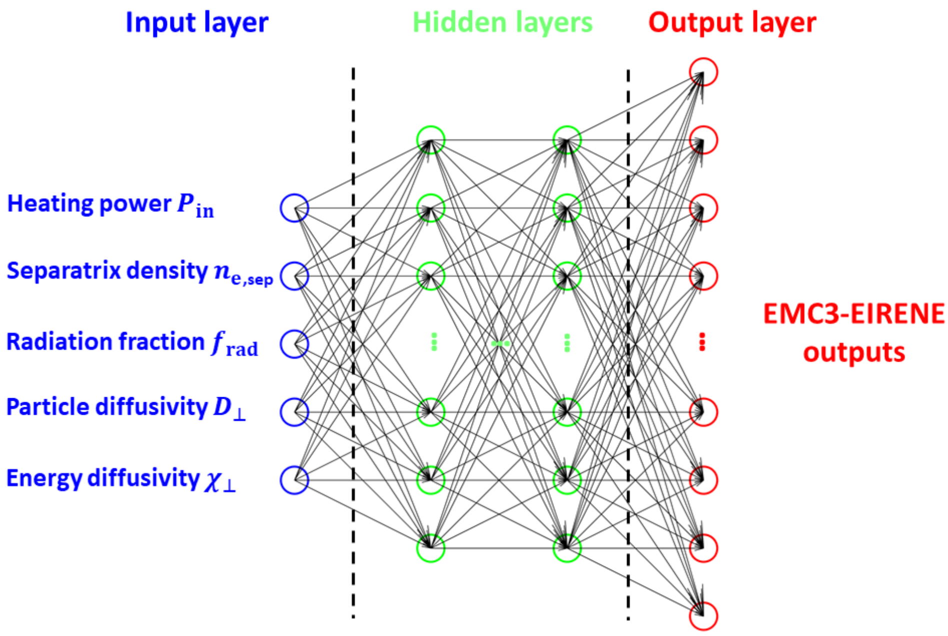

Figure 6.4: Schematic of the reference MLP used for direct input-output mapping.

This section illustrates the hyperparameter optimization strategy using plasma temperature prediction as an example. To maximize data utilization and minimize sampling errors, we employ 5-fold cross-validation. Specifically, the training dataset is divided into five equal subsets, with each four subsets used for training and the remaining subset reserved for validation. Table 6.2 summarizes the search intervals for the key hyperparameters: learning rate, number of hidden layers, number of units per hidden layer, latent vector dimension, batch size, L2 regularization strength, optimizer, activation function, scheduler, and the training epochs for the auto-encoder, mapping stage, and fine-tuned stage. During random search, the learning rate and L2 coefficient were sampled logarithmically, while all other hyper-parameters were drawn uniformly at random within their specified ranges. A total of 100 random hyperparameter sets were selected from this search space, and each set underwent 5-fold cross-validation to calculate the average training and validation loss. The set with the lowest average validation loss was selected as optimal. To efficiently carry out hyperparameter optimization and reduce computational costs associated with numerous trials, the asynchronous successive halving algorithm (ASHA) early stopping strategy [137] was employed.

Table 6.2: Search space of the key hyperparameters.

Hyperparameters	Search space
Learning rate	0.0001-0.01
Number of fully connected layers	2, 3, 4
Number of units per fully connected layer	64, 128, 256
Number of latent vector	32, 64, 128, 256
Batch size	8, 16, 32
L2 regularization	10^{-7} - 10^{-3}
Optimizer	Adam
Activation function	ELU, GELU, LeakyReLU
Scheduler	ASHA
Auto-encoder epoch	100, 200, 300, 400, 500, 600
Mapping epoch	1000, 1500, 2000, 2500, 3000
Fine tuning epoch	400, 600, 800, 1000

Figure 6.5 shows the average training and validation losses from the 5-fold cross-validation across these 100 hyperparameter sets. In panel (a) the blue curve traces the average training loss for the best set, while in panel (b) the red curve traces the corresponding average validation loss. Both panels illustrate that the early-stopping strategy clearly terminate poorly performing trials at an early stage and reallocates computational resources to more promising sets. With the optimal hyperparameter settings, the average training and validation losses during the joint fine-tuned stage converge to approximately 2.0×10^{-4} and 5.6×10^{-4}, respectively. These low converged losses indicate that the surrogate model has accurately captured the nonlinear mapping between EMC3-EIRENE input parameters and outputs.

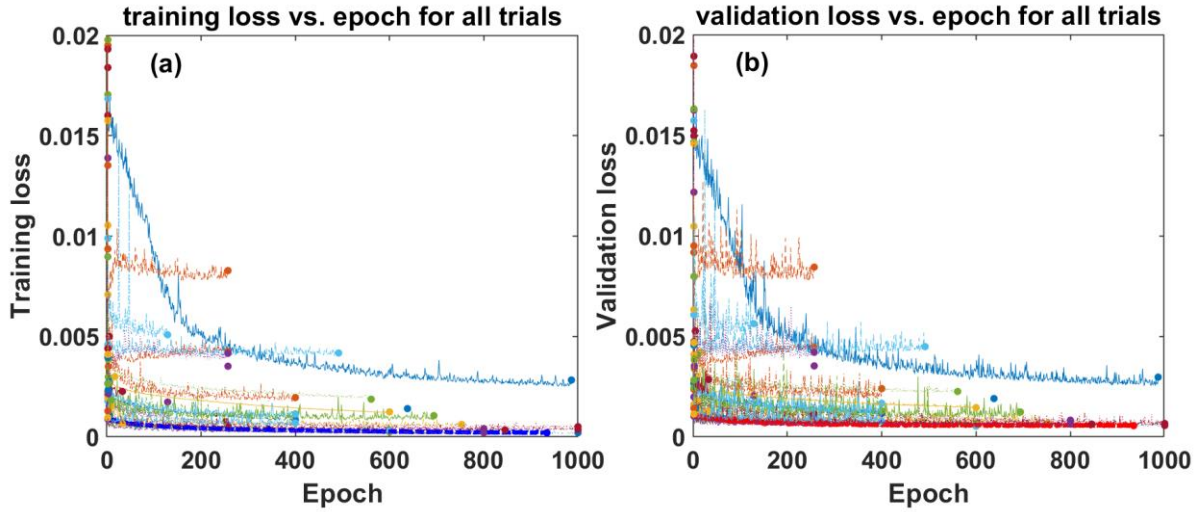

Figure 6.5: (a) Average training loss and (b) validation loss over 100 hyperparameter trials using 5-fold cross-validation.

Figure 6.6 summarizes the 100 randomly sampled hyperparameter settings explored across the search space. Across all training stages (autoencoder pre-training, mapping network training, and joint fine-tuned), the Adam optimizer was employed with a learning rate of 6.2×10^{-5}, batch size of 8, ELU activation functions [144], and L2 regularization [126] of 3.7×10^{-7}. The autoencoder compresses simulation outputs into a 256-dimensional latent representation. The mapping network consists of 3 hidden layers, each containing 256 neurons. Training involved 100 epochs for autoencoder pre-training, 2500 epochs for the mapping network, and 1000 epochs for the joint fine-tuned stage. In practice, due to the ASHA early stopping strategy, the actual number of epochs may be fewer than initially configured. We retrained the network on the full training set using these optimized hyperparameters and applied the same settings to all subsequent model training. Figure 7 shows the hyperparameter importance and their correlations with the validation loss. The importance metric [162] quantifies how each hyperparameter impacts the validation loss. Higher importance values means stronger impact. The correlation indicates the linear relationship between the value of hyperparameters and the validation loss. Positive correlation (green) suggests that lower values lead to smaller validation loss. As shown, the learning rate, batch size, and latent vector dimension are the most significant parameters. All of them demonstrate positive correlations with validation loss, indicating that within the explored parameter space, lower these three parameters generally enhance model performance.

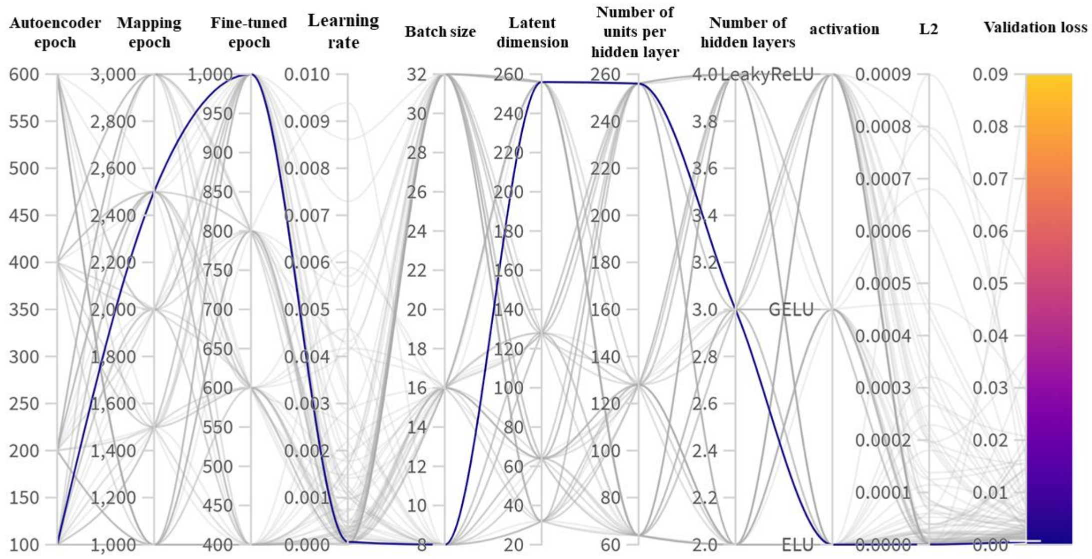

Figure 6.6: Parallel coordinates plot of hyperparameter settings for 100 trials. Each grey line represents a trial's hyperparameter combination and its corresponding validation loss, while the blue line highlights the optimal trial.

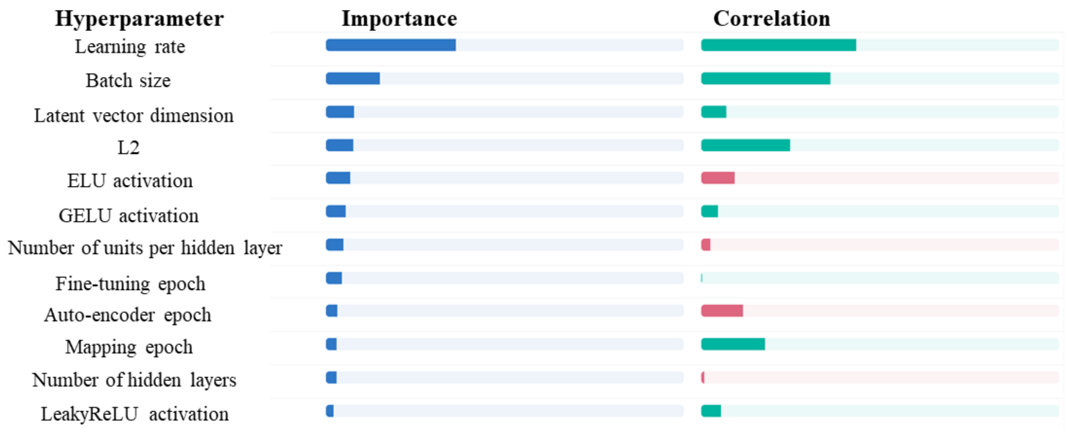

Figure 6.7: Hyperparameter importance and their correlations with validation loss. Left: importance quantifying each hyperparameter's contribution to the variability of the validation loss. Right: correlation with the validation loss.

The reference MLP model underwent a similar hyperparameter optimization procedure, resulting in the following optimal configuration: 200 epochs, three hidden layers with 256 neurons each, a learning rate of 9.6×10^{-4}, batch size of 16, ELU activation function, and L2 regularization of 2.3×10^{-7}, also employing the Adam optimizer. Compared to the joint fine-tuned surrogate, the MLP model exhibited higher training and validation losses, approximately 9.1×10^{-4} and 1.3×10^{-3}, respectively.

To ensure the dataset provides sufficient information for accurately learning the relationship between EMC3-EIRENE inputs and outputs, an assessment of dataset adequacy was performed. We varied the fraction of data allocated to training from 0.1 to 0.9, using the remaining cases as the test set. Throughout, we maintained the same network architecture and hyperparameters and recorded the resulting test loss. As shown in Figure 6.8, the test loss decreases as the trainset fraction increases, then remains nearly unchanged once the fraction exceeds 0.7. This behaviour indicates that the current dataset is already dense enough for accurate surrogate training, and adding more data would produce limited performance gains. This result also justifies using k=5 for k-fold cross-validation in hyperparameter tuning mentioned above. Each fold effectively trains on approximately 64 % of the full dataset, close to the performance plateau. Therefore, k=5 is deemed sufficient and increasing k further would mainly increase runtime with a negligible impact on hyperparameter selection.

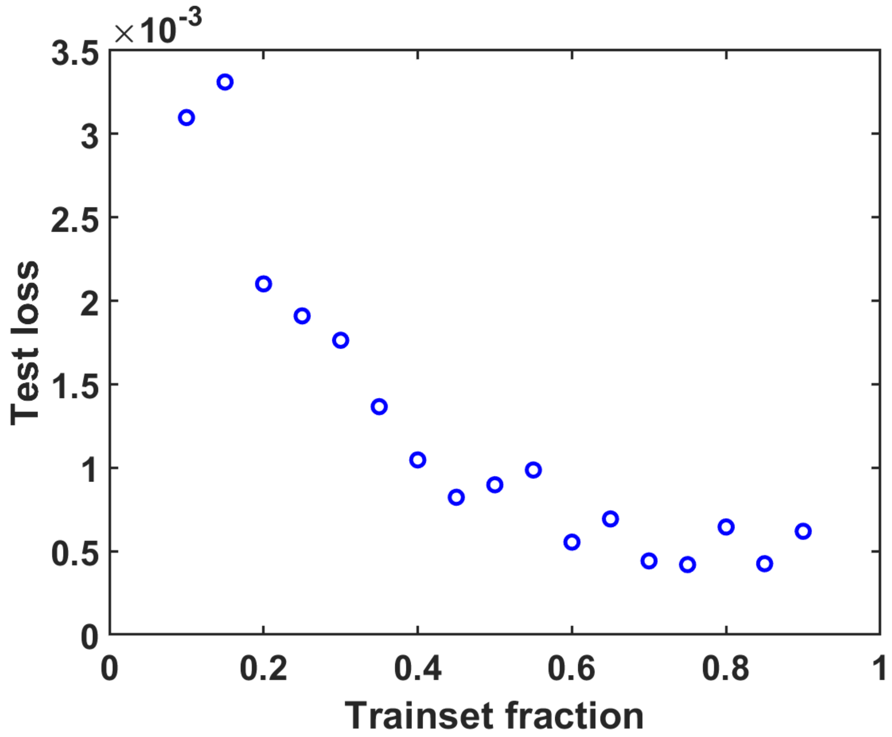

Figure 6.8: Test loss versus the fraction of the dataset used for training.

6.1.3 Performance evaluation of the fine-tuned method

6.1.3.1 Accuracy assessment

The accuracy of the joint fine-tuned surrogate is assessed by benchmarking its predictions of key plasma quantities, namely electron temperature and electron density profiles, against both a conventional multilayer perceptron (MLP) and the reference EMC3-EIRENE simulation. Figure 6.9 plots the probability-density functions of the cell-wise relative error between the surrogate-predicted electron temperature and the EMC3-EIRENE reference for the entire 10 % test set. Both the joint fine-tuned surrogate and the baseline MLP yield error distributions that are close to Gaussian, yet the joint fine-tuned curve is markedly narrower. Only about 7.25 % of the grid cells exhibit a relative error larger than 10 % with the joint fine-tuned model, whereas the MLP exceeds that threshold in 13.33 % of the cells. This tighter distribution confirms that the joint fine-tuned surrogate achieves significantly higher accuracy than the conventional MLP in temperature prediction.

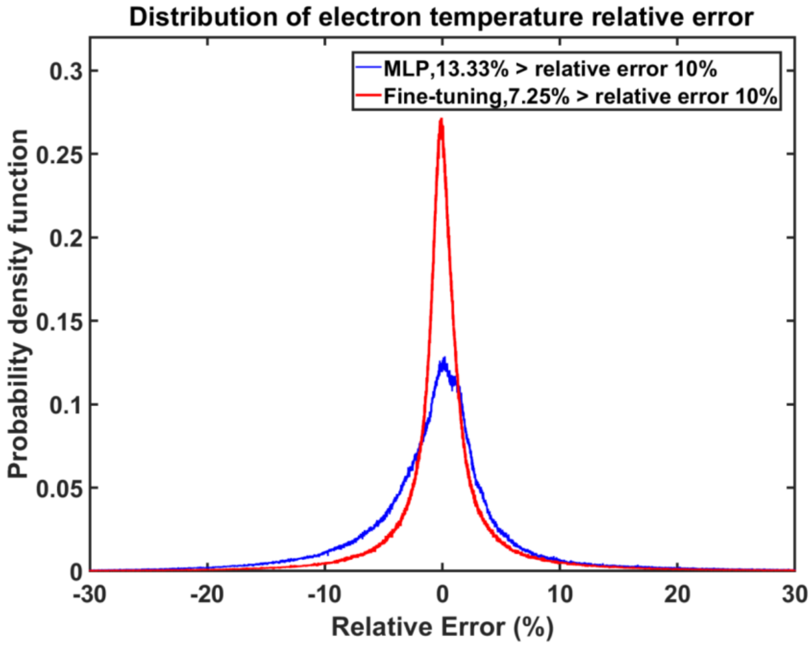

Figure 6.9: Probability density curves of electron temperature relative error: joint fine-tuned surrogate (red), baseline MLP (blue).

To visualize where the largest relative-error regions lie and to clarify the performance gap between the joint fine-tuned surrogate and the baseline MLP, we analyze two representative radiation cases. The first case uses a low radiation fraction ($f_{rad} = 0.2$), while the second a high radiation fraction ($f_{rad} = 0.8$), in which the plasma has accessed a detached scenario. Both cases are run with an input power of $P_{in} = 6\text{MW}$, cross-field particle diffusivity $D_{\perp} = 1.1\ \text{m}^2/\text{s}$, and cross-field thermal diffusivity $\chi_{\perp} = 1.5\ \text{m}^2/\text{s}$. The separatrix electron densities are $n_{e,sep} = 3.0 \times 10^{19}\ \text{m}^{-3}$ for the low-radiation case and $5.5 \times 10^{19}\ \text{m}^{-3}$ for the high-radiation case. Figure 6.10 compares the two models for the low radiation fraction case at the bean-shaped cross-section. Panel (a) shows the EMC3-EIRENE electron-temperature distribution. Panels (b) and (c) give the MLP's absolute and relative errors, while panels (d) and (e) show the same errors for the joint fine-tuned surrogate, all measured against the EMC3-EIRENE reference. In contrast, the joint surrogate is much more accurate. Its relative error remains low across most of the domain and rises only in very low-temperature areas, where the absolute deviation stays modest at roughly 2-4 eV. Figure 6.11 presents the corresponding results for the high radiation fraction case. Accuracy decreases for both surrogates, and the MLP's SOL relative error exceeds 10 %, making it unreliable for prediction. The joint surrogate's relative error approaches 10 % only in a few low-temperature regions, where the absolute error remains small at roughly 1 eV, while relative errors elsewhere stay modest. Overall, grid cells where the joint fine-tuned surrogate exceeds a 10 % temperature-relative-error threshold are confined to the far-edge, low-temperature region of the plasma.

Consequently, the surrogate remains well suited for accurate temperature predictions in both attached and detached operating regimes.

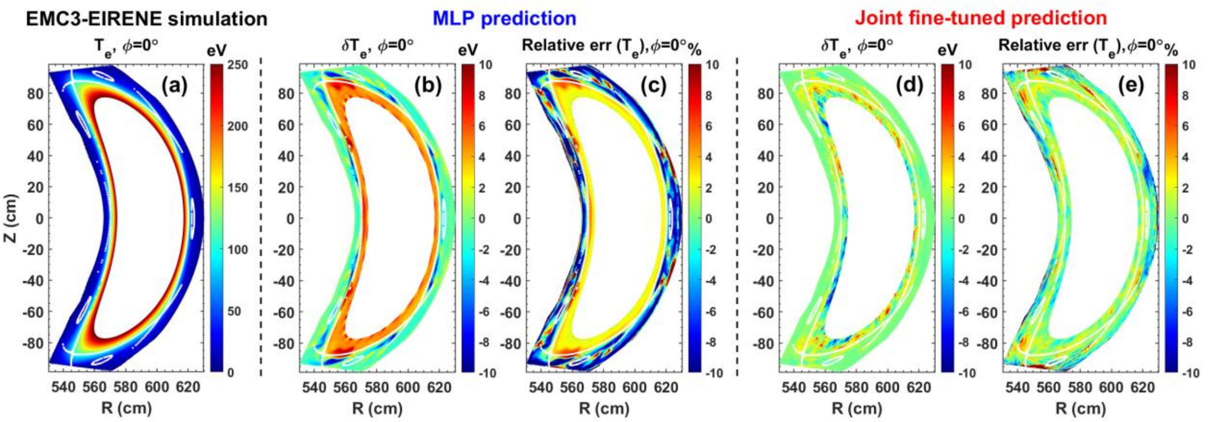

Figure 6.10: Comparison of electron temperature predictions for a low radiation fraction case. (a) EMC3-EIRENE reference, (b, c) MLP absolute and relative errors, (d, e) joint fine-tuned surrogate absolute and relative errors.

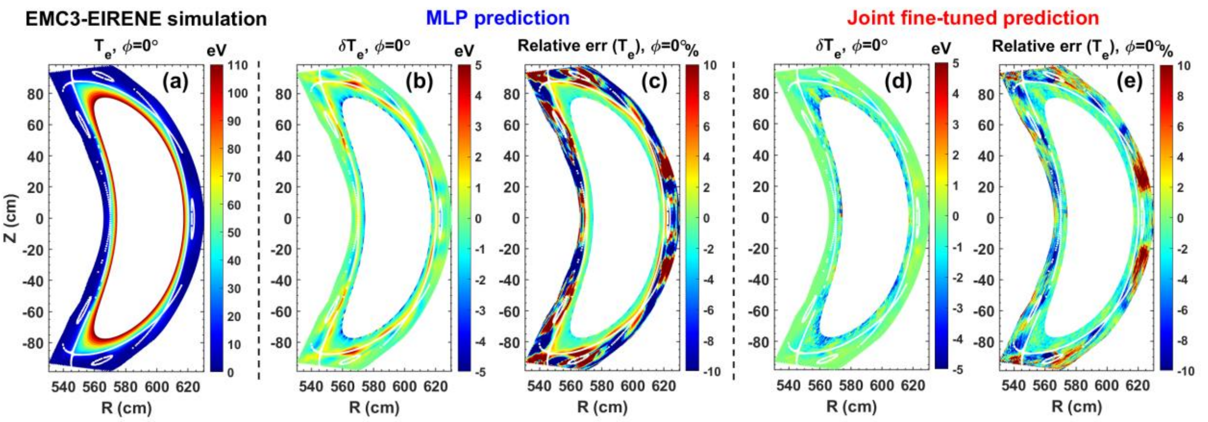

Figure 6.11: Comparison of electron temperature predictions for a high radiation fraction case. (a) EMC3-EIRENE reference, (b, c) MLP absolute and relative errors, (d, e) joint fine-tuned surrogate absolute and relative errors.

An identical assessment was carried out for electron-density predictions. Figure 6.12 displays the probability-density functions of the relative error on the test set for both the joint fine-tuned surrogate (red) and the baseline MLP (blue). As with temperature, each PDF is approximately Gaussian; however, the joint surrogate's curve is noticeably sharper and more concentrated around zero. Only about 4.92 % of grid cells show a density-relative error larger than 10 % for the joint model, whereas the MLP exceeds that threshold in roughly 8.33 % of cells. These results confirm that the joint fine-tuned surrogate also outperforms the MLP in density prediction accuracy.

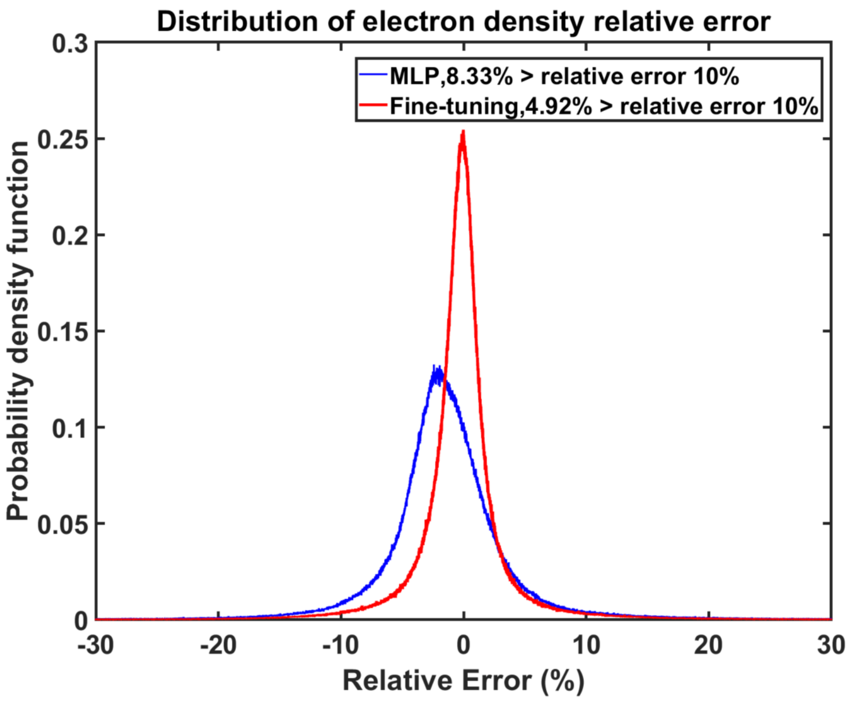

Figure 6.12: Probability-density curves of electron density relative error: joint fine-tuned surrogate (red), baseline MLP (blue)

We also selected the same high- and low-radiation cases used in the temperature comparison to examine density, aiming to identify where each surrogate's density-prediction error exceeds 10 %. Figure 6.13 compares the low-radiation case at the bean-shaped cross-section. Panel (a) gives the reference EMC3-EIRENE electron-density map, while panels (b)-(c) show the MLP's absolute and relative errors and panels (d)-(e) present the corresponding errors for the joint fine-tuned surrogate. The joint model is clearly more accurate: its relative error stays low across most of the domain and rises only at the extreme boundary. Figure 6.14 shows the high-radiation case. Accuracy drops for both surrogates, but the MLP's relative error in the SOL exceeds 10 %, making its prediction unreliable there. For the joint model, only the region near the magnetic island, where the absolute density is high, approaches the 10 % threshold, while the rest of the domain remains well within acceptable limits. Overall, grid cells with density-relative errors above 10 % for the joint fine-tuned surrogate are confined to edge zones or to the very high-density region near the island. The model therefore remains well suited for accurate density predictions in both attached and detached regimes.

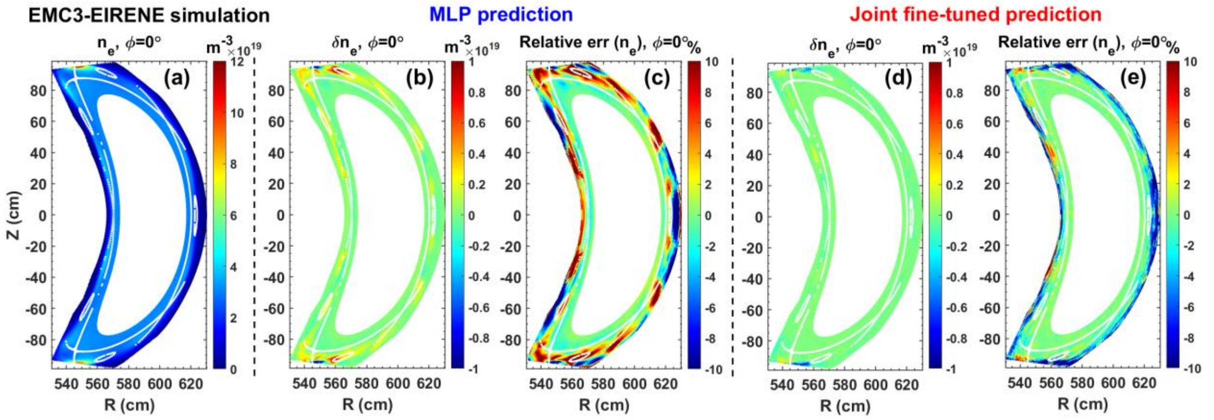

Figure 6.13: Comparison of electron density predictions for a low radiation fraction case. (a) EMC3-EIRENE reference, (b, c) MLP absolute and relative errors, (d, e) joint fine-tuned surrogate absolute and relative errors.

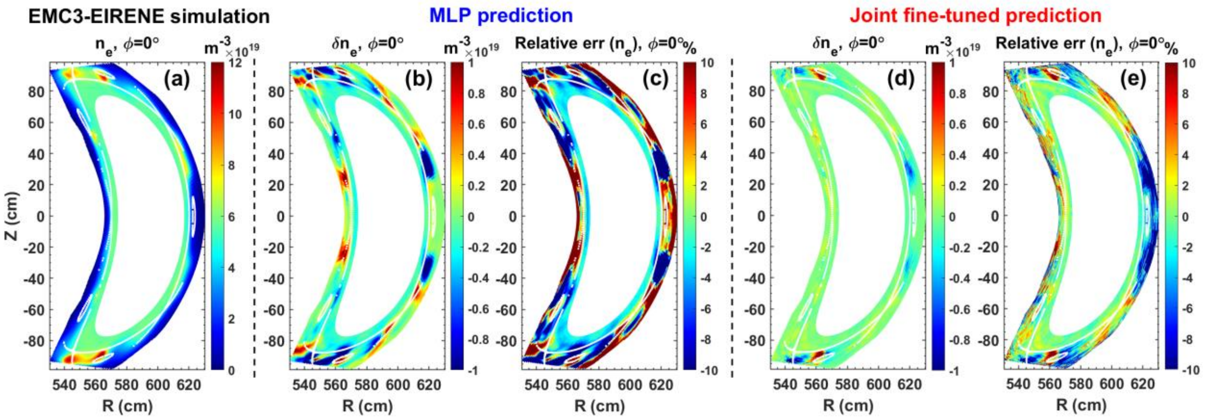

Figure 6.14: Comparison of electron density predictions for a low radiation fraction case. (a) EMC3-EIRENE reference, (b, c) MLP absolute and relative errors, (d, e) joint fine-tuned surrogate absolute and relative errors.

6.1.3.2 Interpolation and extrapolation capability assessment

Generalization was evaluated with a leave-one-value-out (LOVO) test. For each of the five input parameters, we removed every sample that shared one discrete value, trained on the remaining data, and then used the removed cases as an independent test set. Figure 6.15 shows the resulting test loss for each omitted value. The horizontal red dashed line indicates the average validation loss from 5-fold cross-validation obtained with the optimal hyperparameters. When the omitted value falls inside the training range, the test loss stays close to this baseline, demonstrating strong interpolation performance. If the omitted value lies outside the training range, the loss rises sharply, which reveals a limited ability to extrapolate. The results suggest that future dataset development should prioritize extending the range of input parameters over further refining the resolution within already covered regions. We also observe that parameters whose omission removes more than 20% of the cases (panels a, b and e) yield higher test losses, whereas those with fewer removed cases (panels c and d) show test losses comparable to the cross-validation baseline.

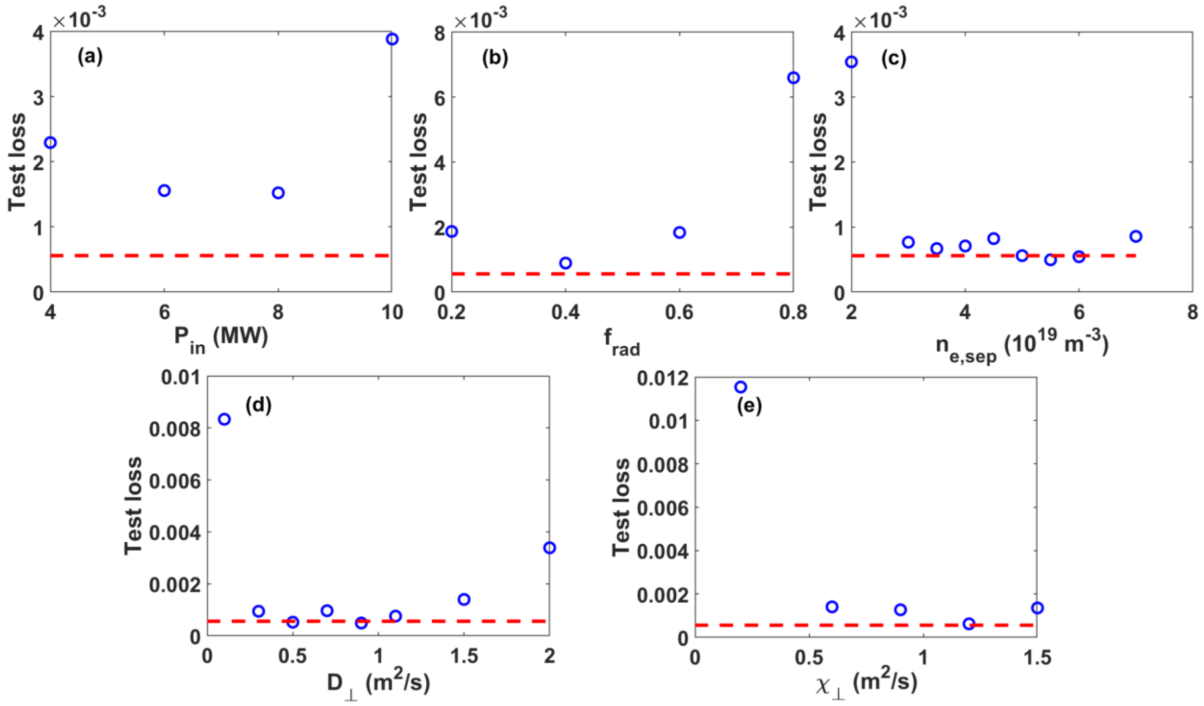

Figure 6.15: Leave-one-value-out test loss as a function of the omitted input value. The red dashed line indicates the average validation loss from 5-fold cross-validation with the optimal hyperparameters.

6.1.3.3 Sensitivity assessment of the surrogate's input-output mapping

To examine the input-output relations learned by the final joint fine-tuned surrogate, we performed a post-hoc sensitivity analysis of the three outputs T_e, n_e and P_{rad}, complemented by spatial maps that indicate where each input is most influential. We use partial dependence (PDP) to summarize how each spatial mean output varies with one input. The PDP is the average of the individual conditional expectation (ICE) curves, where each ICE curve is obtained by sweeping a single input while holding the other four fixed. A set of 128 background samples was randomly selected from the input parameter ranges listed in Table 6.1. For each input value, we computed the central 95% pointwise interval across the ICE curves, forming the shaded "95% ICE band" in Figure 6.16. The PDP for the mean T_e increases with heating power P_{in} and decreases with f_{rad}, $n_{e,sep}$, $D_\perp$, and $\chi_\perp$. The mean n_e grows almost linearly with $n_{e,sep}$, drops rapidly and then becomes nearly flat as $D_\perp$ increases, and shows only weak dependence on the remaining inputs when $n_{e,sep}$ is fixed. For the mean P_{rad}, the PDP rises with either P_{in} (at fixed f_{rad}) or f_{rad} (at fixed P_{in}), and is nearly flat with respect to the other inputs. Notably, these trends are consistent with physical expectations, further validating the reliability of the joint fine-tuned surrogate model. The width of the 95% ICE band reflects dependence on the other inputs: wider bands indicate stronger input-input interactions.

To provide a dimensionless global ranking of input importance for the mean outputs, normalized gradient sensitivities (elasticities) are used, defined as:

$$E_i = \left|\frac{\partial S}{\partial x_i}\frac{x_i}{S}\right| \quad 6.1.1$$

where S is the spatial mean of each output and x_i is an input. Gradients $\frac{\partial S}{\partial x_i}$ were obtained by automatic differentiation at 256 randomly selected input sets within the range of Table 1. Figure 6.17 plots the median E_i across those sets. For the mean T_e the dominant inputs are P_{in} and $n_{e,sep}$, followed by $\chi_\perp$, f_{rad}, and $D_\perp$. For the mean n_e, $n_{e,sep}$ dominates with smaller contributions from P_{in} and f_{rad}. For the mean P_{rad}, P_{in} and f_{rad} are most influential, with the other inputs playing minor roles.

Finally, non-dimensional sensitivity (elasticity) maps were constructed at each grid cell to localize sensitivity in space. The maps adopt the same elasticity as Equation 6.1.1 but are evaluated pointwise by replacing the global mean with the local output value. The resulting per-sample sensitivities are then averaged at each grid cell over 128 background samples. The resulting maps are shown in Figure 6.18. Bright regions mark locations where small fractional changes in an input lead to large fractional changes in the output. The maps indicate that P_{in}, $n_{e,sep}$, and $\chi_\perp$ affect T_e broadly across the domain, whereas f_{rad} and $D_\perp$ act mainly in the SOL. For n_e, $n_{e,sep}$ controls the profile almost everywhere while the other inputs mostly modulate the SOL. For P_{rad}, the cross-field particle transport $D_\perp$ can have a more global impact, because $D_\perp$ alters the penetration depth of impurity carbon and thereby reshapes the radiation distribution throughout. The remaining inputs primarily influence radiation in the SOL.

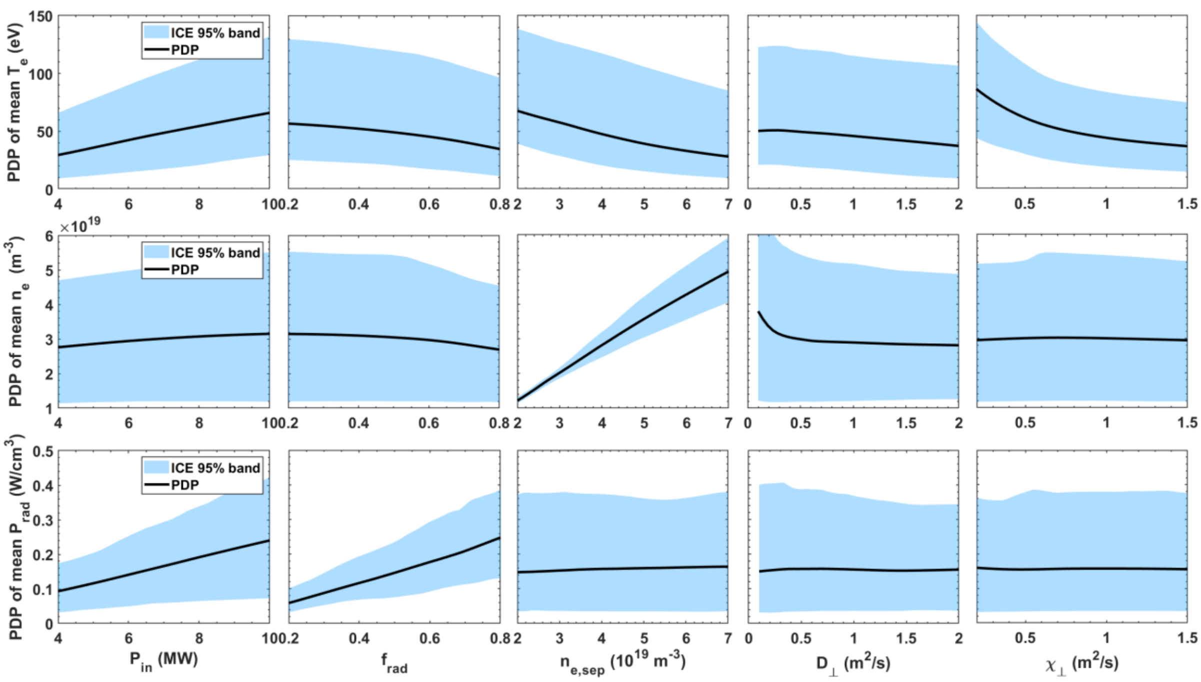

Figure 6.16: Partial dependence of mean outputs (T_e, n_e and P_{rad}) on each input with 95% ICE band. The black curve is the PDP, and the blue shaded region is the 95% ICE band.

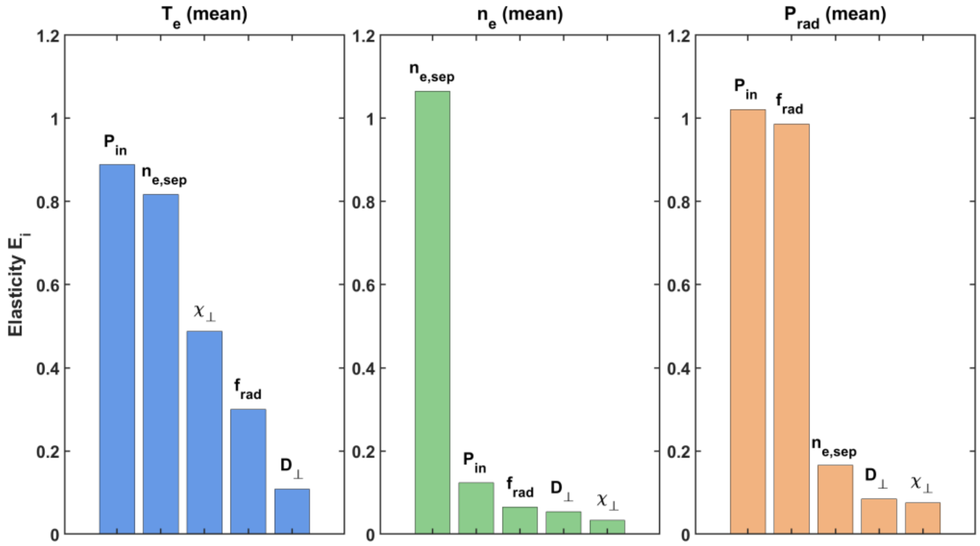

Figure 6.17: Global normalized gradient sensitivities (elasticities) of mean T_e, n_e and P_{rad} to the inputs. Bars show the median elasticities over 256 samples.

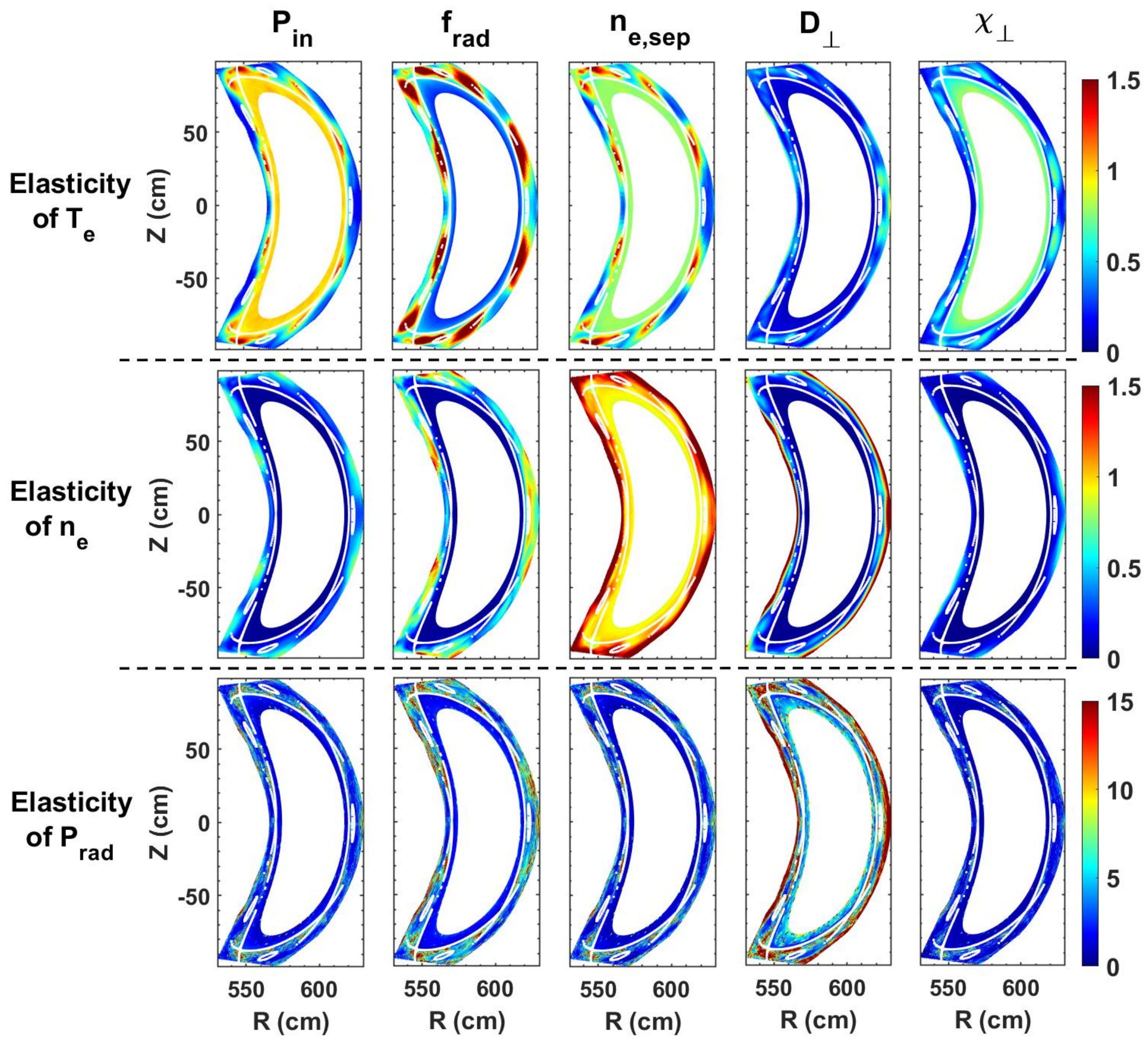

Figure 6.18: Spatial sensitivity maps of T_e, n_e and P_{rad} to each input, averaged over 128 background samples. Bright regions mark locations of large fractional response to inputs.

6.1.3.4 Computational efficiency of the fine-tuned surrogate model

An illustrative example is presented here to show the computational efficiency of the surrogate model. The surrogate requires only about 3.0×10^{-6} core · h for each prediction, whereas a full EMC3-EIRENE run on the present $144 \times 256 \times 144$ mesh on W7-X typically exceeds 1000 core · h, providing a speed-up of over 10^8. This performance makes the surrogate ideal for rapid studies. Tasks such as mapping divertor operating regimes against density, radiation fraction, or cross-field transport coefficients, which are prohibitively slow with the full EMC3-EIRENE code, can now be completed in milliseconds. Figure 6.19(a) plots the surrogate's recycling flux predictions versus separatrix electron density and f_{rad} at fixed $P_{in} = 6$MW, $D_{\perp} = 0.5$, and $\chi_{\perp} = 1.5\ \mathrm{m}^2/\mathrm{s}$. The plot includes about 6 000 cases, all evaluated on a standard office PC within a few seconds. The recycling flux is a key parameter. Its rollover is an indicator of the onset of divertor detachment. Figure 6.19 (b) shows the probability density of the relative error in recycling flux predictions on the test set. In the vast majority of cases, the errors are quite small, further demonstrating that surrogate model is well-suited for fast large-scale data analysis in the future.

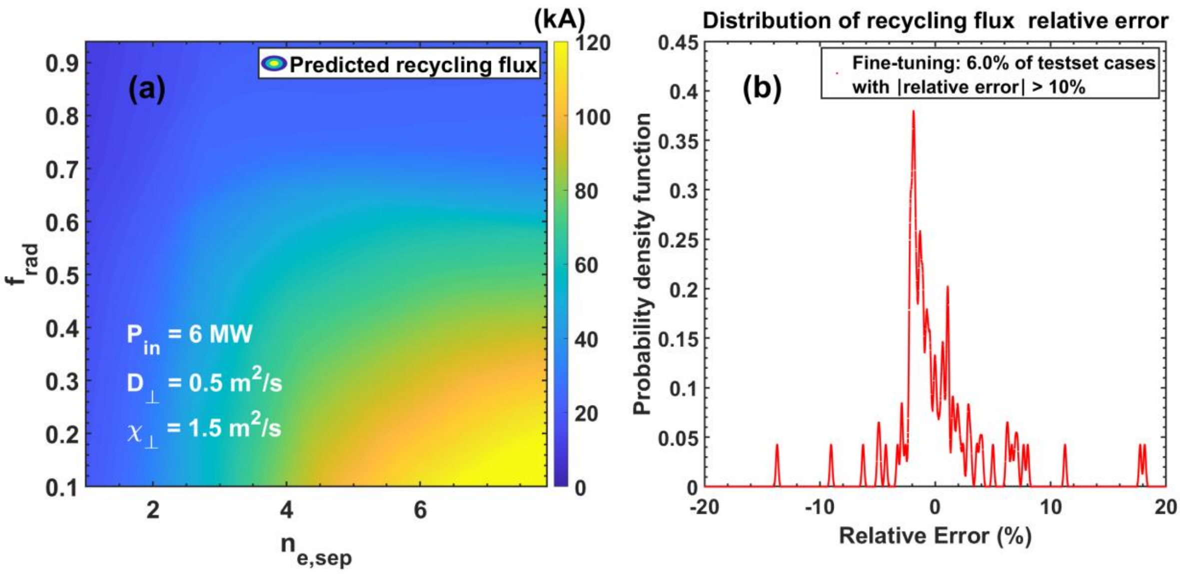

Figure 6.19: (a) Surrogate-predicted recycling flux as a function of separatrix electron density and radiation fraction at $P_{in} = 6$ MW, $D_{\perp} = 0.5$, and $\chi_{\perp} = 1.5\ \mathrm{m}^2/\mathrm{s}$. (b) probability density of relative error in recycling flux predictions on the test set.

In short, the joint fine-tuned surrogate outperforms the simple MLP, especially for detachment cases prediction, and reproduces EMC3-EIRENE results with satisfactory accuracy. In generalization tests, the surrogate exhibits excellent interpolation capability but weaker extrapolation performance, providing guidance for future dataset construction and expansion. Compared to full EMC3-EIRENE runs, the surrogate achieves at least an 8-order-of-magnitude speed-up, making it well suited for rapid predictions and extensive parameter scans, and paving the way for real-time application of 3D edge transport simulations in the control room to optimize ongoing experiments.

6.2 Real-time predictions of edge plasma via EMC3-EIRENE-AI

In this section we combine the results of Sections 5.3 and 6.1 to create an end-to-end neural surrogate, EMC3-EIRENE-AI, that predicts full 3D plasma profiles consistent with the EMC3-EIRENE model directly from routine edge-diagnostic measurements and their uncertainties, delivering results on an office desktop in just seconds to a few minutes. The accuracy of EMC3-EIRENE-AI follows from the benchmarks in the previous sections: fewer than 10 % of all grid cells show relative errors above 10 %, and those larger errors are almost confined to the edge plasma region. Within the overall workflow, the Bayesian inference stage is the slowest step, requiring anywhere from a few seconds to several minutes on a 2-core CPU. Compared with the traditional trial-and-error adjustment of EMC3-EIRENE inputs to match diagnostics, where each run demanding thousands of CPU-core-hours, this approach cuts both human effort and computing time while providing results that are statistically consistent with measurement uncertainties. Fast surrogates are likely to become essential for large, integrated modelling tasks. Because EMC3-EIRENE-AI can achieve fast predictions, it could eventually be used directly in the control room to guide ongoing experiments, greatly improving future experimental efficiency.

6.2.1 Workflow for EMC3-EIRENE-AI

In order for EMC3-EIRENE-AI to predict full 3D plasma profiles that remain faithful to the EMC3-EIRENE physics model using only routine edge-diagnostic measurements and their uncertainties, we introduce an intermediate representation, the EMC3-EIRENE input parameters. As shown in Figure 6.20, the workflow proceeds in two steps:

1. Inference of EMC3-EIRENE input parameters from diagnostics
   The neural-surrogate-assisted Bayesian framework developed in Section 5.3 rapidly infers the posterior distributions of the EMC3-EIRENE input parameters from the experimental data. Within this framework, a feed-forward neural-network surrogate maps candidate inputs to synthetic diagnostic data, and Dynamic Nested Sampling, working in concert with the surrogate, systematically explores the parameter space to find the input parameters that best reproduce the measurements within their stated uncertainties.
2. EMC3-EIRENE output prediction
   The joint fine-tuned surrogate introduced in Section 6.1 then converts these inferred inputs into complete EMC3-EIRENE outputs (electron density, ion temperature, radiation distribution, heat load, etc.) in milliseconds.

The data-preparation pipeline for EMC3-EIRENE-AI therefore merges the preprocessing steps from Sections 5.3 and 6.1. Both surrogates draw on the same EMC3-EIRENE simulation dataset; they differ only in their mapping direction. Because EMC3-EIRENE-AI is designed to guide live experiments and enable rapid analysis of their data, its preprocessing pipeline handles large volumes of diagnostic measurements together with their uncertainties. It can produce predictions at every time step in a discharge, and its time resolution is only limited by the slowest diagnostic (interpolation can be applied if needed). All diagnostic inputs are normalized exactly as in the EMC3-EIRENE training dataset, and each measurement must carry an uncertainty estimate and a prior distribution for the corresponding EMC3-EIRENE input parameter.

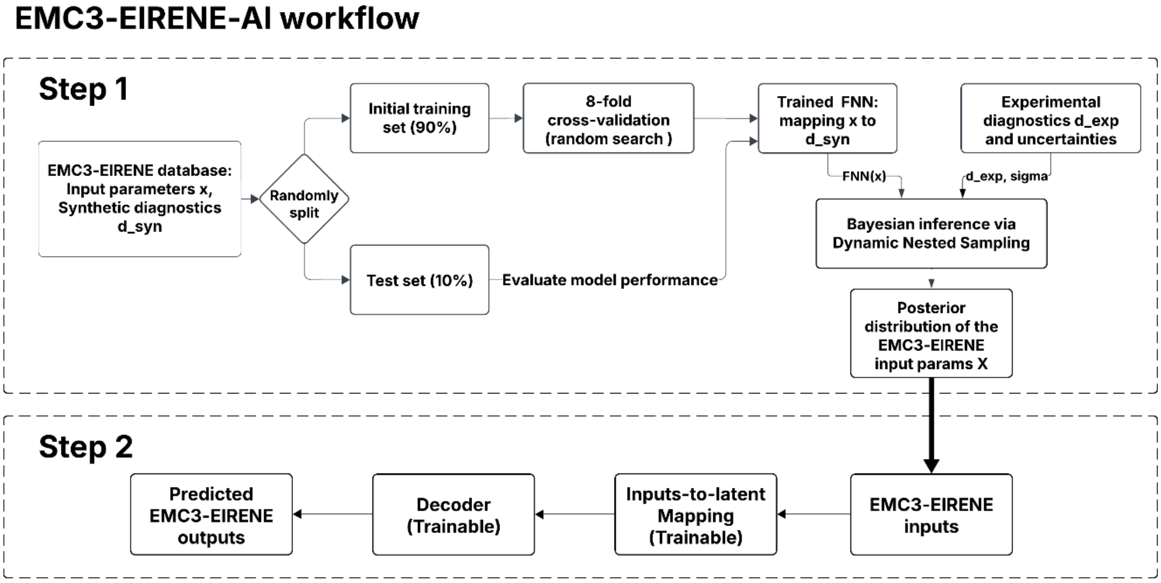

Figure 6.20: Two-step EMC3-EIRENE-AI workflow for Bayesian input inference and fast surrogate output prediction.

6.2.2 EMC3-EIRENE-AI: runtime-accuracy trade-off

In the full EMC3-EIRENE-AI prediction pipeline, the most time-consuming step is the Bayesian inference in Step 1, since its cost depends on parameters such as the number of live points and the stopping threshold *d*logz. The live points parameter specifies how many "active" samples are initially maintained in the parameter space: more live points yield more uniform coverage of that space and therefore more accurate estimates of both the posterior distribution and the model evidence. The stopping threshold *d*logz dictates that sampling ceases once the remaining uncertainty in the log-evidence falls below the preset value; a smaller *d*logz thus corresponds to a more precise evidence estimate. EMC3-EIRENE-AI's two steps differ dramatically in speed: inferring the input parameters takes seconds to minutes, whereas predicting the full outputs from those inputs via the surrogate model takes only milliseconds. Here, we examine the trade-off between runtime and accuracy in EMC3-EIRENE-AI. For experimental data at $t = 5.9$ s in discharge #20181010.028, we used the Bayesian framework to infer the EMC3-EIRENE input parameters while scanning both the number of live points and the value of *d*logz. Figure 6.21 shows, at fixed *d*logz = 0.1, how both the total inference time and the resulting MAP estimates vary as the number of live points is increased. Although the computation time grows with more live points, the MAP values change very little. This demonstrates that for our simple, unimodal, low-dimensional problem, a few dozen live points suffice, and increasing them further yields diminishing returns. Moreover, reducing the number of live points does not drastically shorten the runtime: even with as few as 20 live points, inference still requires on the order of 400 s.

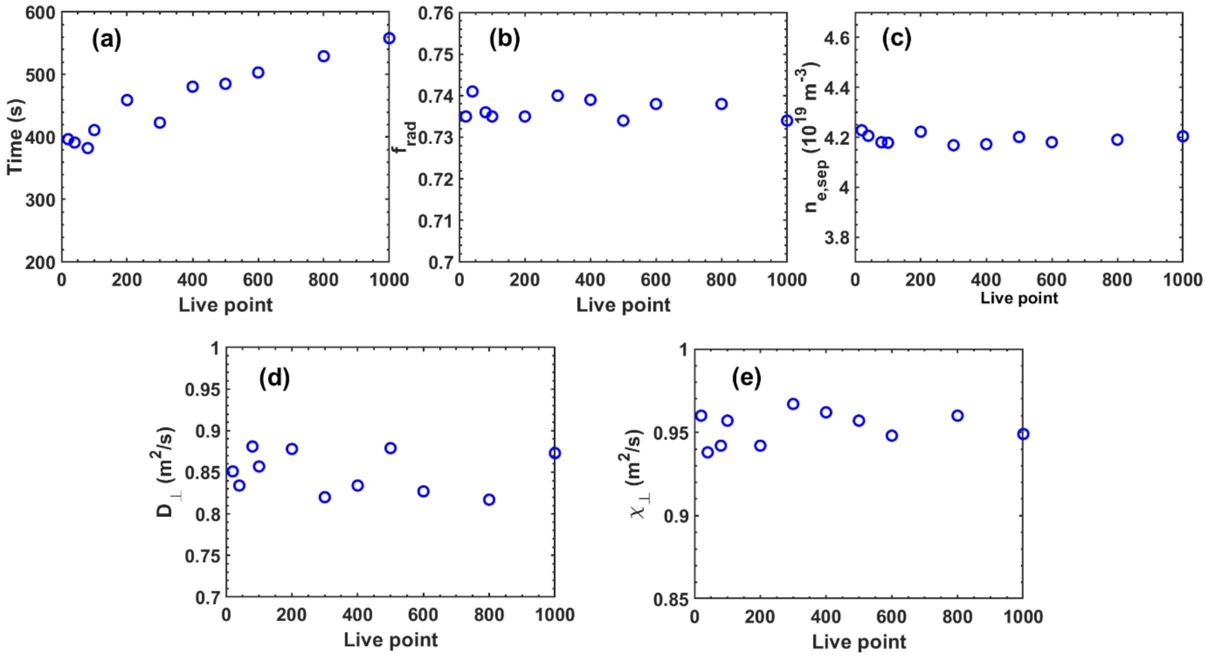

Figure 6.21: Influence of the number of live points on computational cost and MAP estimates at a fixed convergence threshold $d$logz = 0.1. (a) wall-clock time required for the Bayesian inference versus live-point count; (b) inferred radiation fraction f_{rad}; (c) separatrix electron density $n_{e,sep}$; (d) cross-field particle diffusivity $D_\perp$; and (e) cross-field thermal diffusivity $\chi_\perp$, all plotted as functions of the live-point count.

Figure 6.22 investigates the runtime-accuracy compromise obtained by varying dlogz while keeping the number of live points fixed at 600. Panel (a) demonstrates that relaxing the convergence threshold ($d$logz) sharply shortens the computational time. Panels (b)-(e) track the MAP estimates of f_{rad}, $n_{e,sep}$, $D_\perp$ and $\chi_\perp$. As $d$logz decreases (i.e. as the allowed runtime increases) the MAP values gradually stabilise. When $d$logz is set to 20 the run finishes in roughly 32 s, and the MAP estimates are already very close to their asymptotic values. Thus, if one is interested only in a reliable MAP and can tolerate a loose evidence tolerance, half a minute per inference is usually sufficient. Conversely, achieving a tighter evidence criterion ($d$logz $\lesssim$ 5) requires substantially more samples and pushes the runtime to ~2 min or beyond. A convergence threshold of $d$logz $\approx$ 0.1-1 is often adopted. Reaching that level of accuracy enlarges the sample set and typically stretches the runtime to about 3-6 minutes per inference.

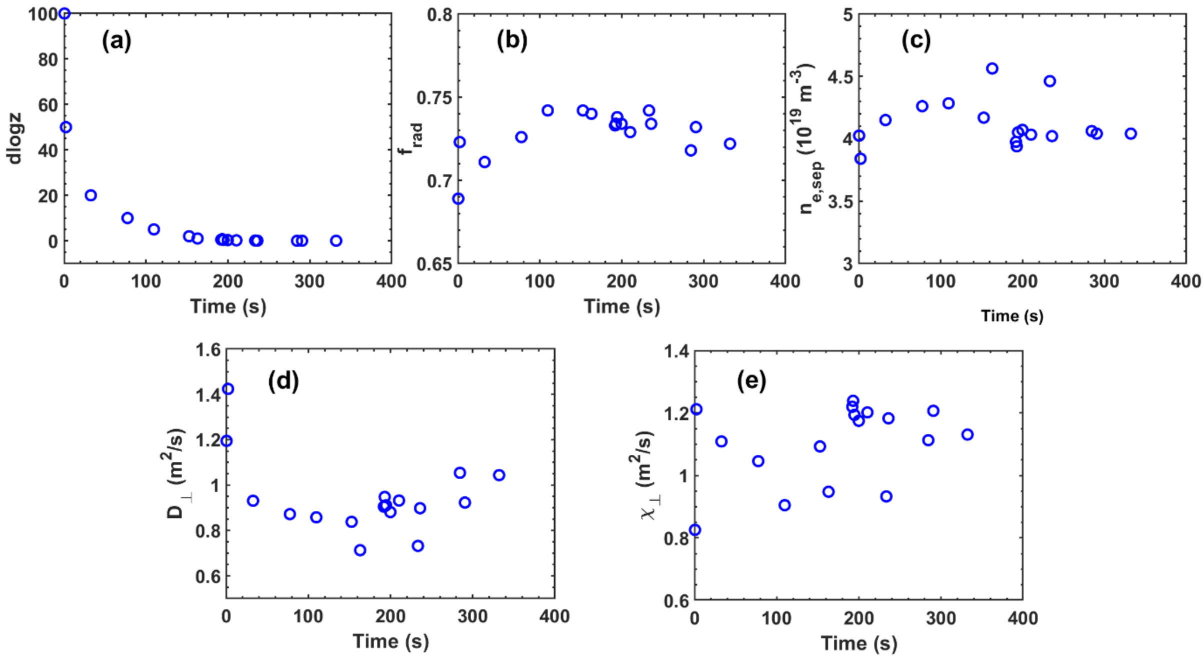

Figure 6.22: Runtime-accuracy trade-off obtained by varying the convergence threshold $d$logz (with live-point count fixed at 600). (a) runtime versus $d$logz; (b) corresponding MAP estimate of f_{rad}; (c) $n_{e,sep}$; (d) $D_\perp$; and (e) $\chi_\perp$ as functions of the total inference time.

6.2.3 Application of the EMC3-EIRENE-AI

EMC3-EIRENE-AI takes routine diagnostic signals together with their error bars and, inside a Bayesian framework,

- infers the EMC3-EIRENE input parameters, including the cross-field transport coefficients, that best match the data, and
- instantly predicts the full set of outputs required by the EMC3-EIRENE physics model.

To demonstrate its practical use, we revisit the density-ramp discharge #20181010.028 on W7-X already analyzed in Section 5.3. Figure 6.23 again shows the principal plasma traces for this shot, which undergoes a characteristic ramp to divertor detachment. The three marked time points correspond to different radiation fractions: $f_{rad}\ (t1) = 0.2$, $f_{rad}\ (t2) = 0.5$, and $f_{rad}\ (t3) = 0.8$. For each time slice, and to capture both upstream and divertor conditions, we feed the following edge-diagnostic set into the framework:

- divertor C II line emission from divertor spectroscopy,
- Thomson-scattering (TS) radial profiles of electron density and temperature,
- ion-saturation currents from divertor Langmuir probes (LPs), and
- divertor heat-flux profiles from divertor thermography.

From these inputs EMC3-EIRENE-AI returns self-consistent 3D plasma solutions, opening up several concrete application routes that we outline in the next subsections.

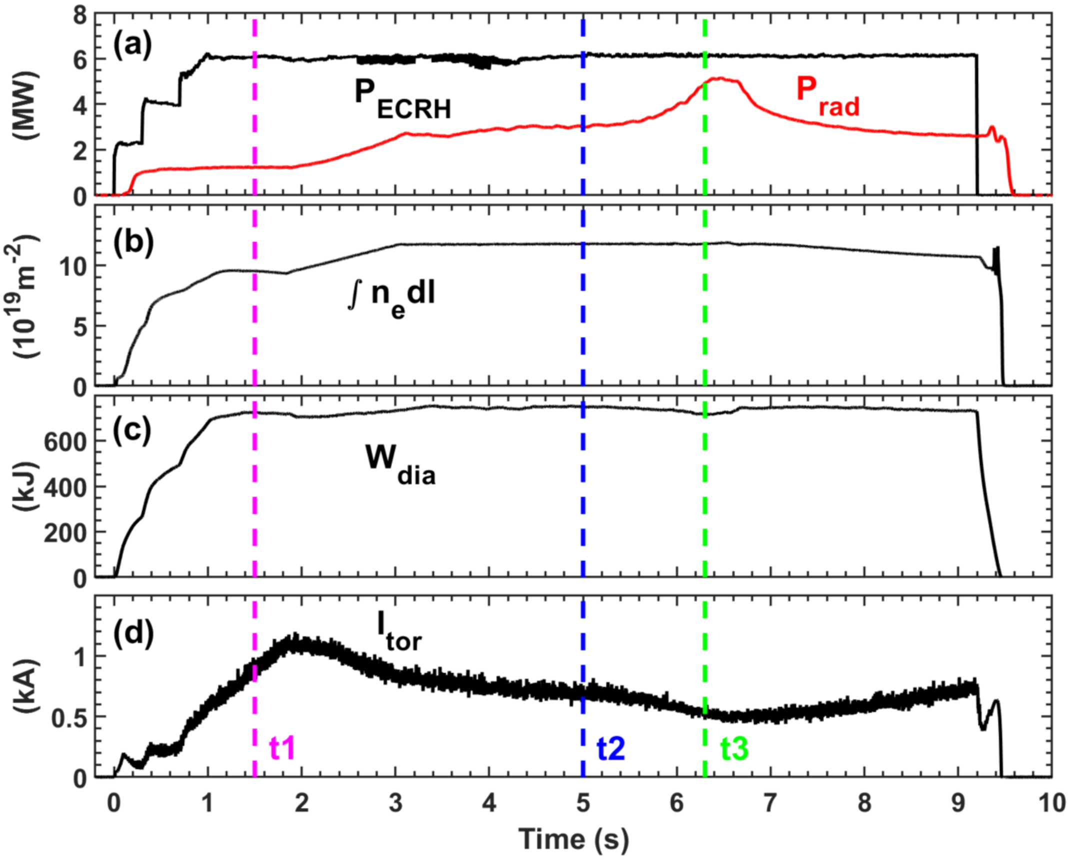

Figure 6.23: Temporal evolution of key signals for the density ramp-up discharge #20181010.028 on W7-X. The three vertical dashed lines indicate the moments selected for detailed analysis.

Fast transport analysis

Figure 6.24 tracks the input parameters returned by the Bayesian step of EMC3-EIRENE-AI, the cross-field particle diffusivity $D_{\perp}$ (blue) and thermal diffusivity $\chi_{\perp}$ (red), together with the experimentally determined radiation fraction f_{rad} (cyan). When building the EMC3-EIRENE dataset we assumed that

- $D_{\perp}$ is spatially uniform and identical for impurity ions and background hydrogen ions, and
- $\chi_{\perp}$ is also spatially uniform and the same for electrons and ions.

Accordingly, the curves in Figure 6.24 represent volume-averaged transport coefficients for the entire computational domain. Overall, as the radiation fraction increases and the plasma moves from low to high recycling regime, both the cross-field particle diffusivity $D_{\perp}$ and the thermal diffusivity $\chi_{\perp}$ rise. $\chi_{\perp}$ suddenly drops at about $t \approx 3$ s, while $D_{\perp}$ shows a sharp decrease around $t \approx 5.0$ s. As the plasma then transitions from high recycling toward detachment, both $D_{\perp}$ and $\chi_{\perp}$ climb slightly. After the divertor gas puffing is terminated, the plasma density falls, the radiation level decreases, and $D_{\perp}$ and $\chi_{\perp}$ drop again. The inferred parameter ranges yield direct gains in physical insight, model calibration, and the planning of experiments.

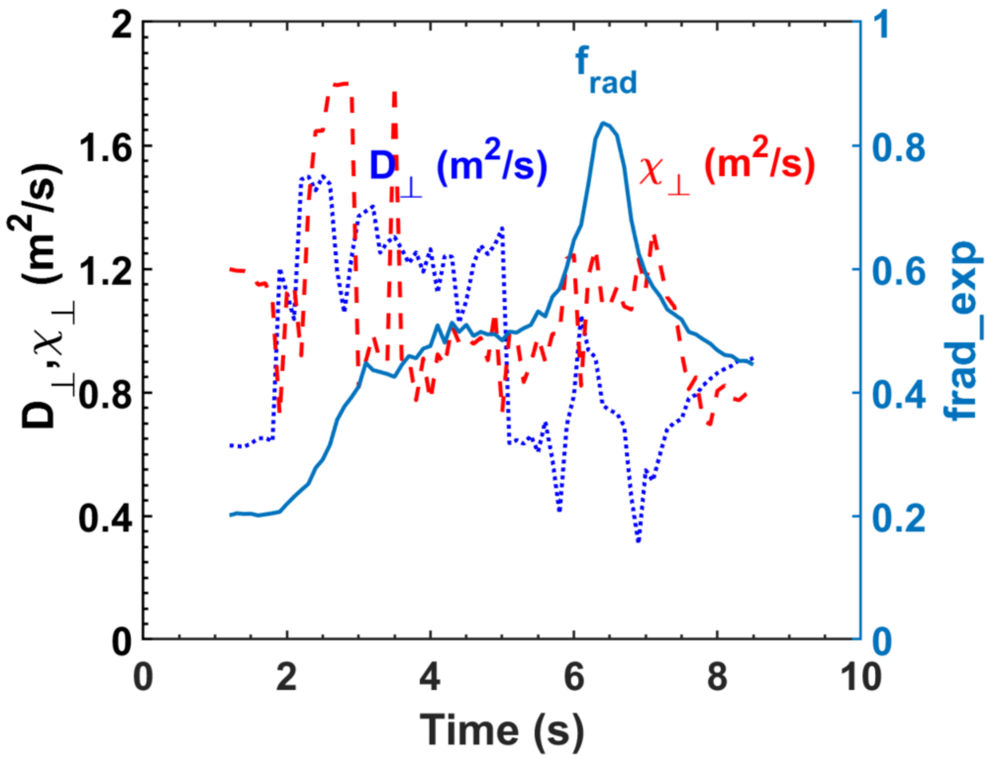

Figure 6.24: Variation of the inferred cross-field transport coefficients with time. The blue curve represents cross-field particle diffusivity, and the red curve shows thermal diffusivity.

Boundary-regime evaluation

Because the EMC3-EIRENE outputs include the recycling flux, EMC3-EIRENE-AI can track its evolution directly from the experimental diagnostics, providing a real-time indication of which divertor regime the plasma occupies. Under the simplifying assumption of fixed cross-field transport coefficients, the qualitative dependence of the recycling flux on separatrix electron density $n_{e,sep}$ or on the radiated-power fraction f_{rad} is sketched in Figure 6.25: in the low-recycling regime the flux rises almost linearly; in the high-recycling regime it increases non-linearly (roughly quadratically); and at the onset of detachment the flux reaches a roll-over and begins to decline.

Figure 6.26 shows the recycling-flux time trace predicted by EMC3-EIRENE-AI for the density-ramp discharge. The orange curve gives the recycling flux, while the cyan curve shows the measured radiation fraction $f_{rad,exp}$. As $f_{rad,exp}$ climbs, the recycling flux initially grows, but around $f_{rad,exp} = 0.45$ it rolls over, signalling the transition from high recycling into detachment. With further radiation rise the flux continues to drop, and when the plasma enters deep detachment (sharp radiation increase) the recycling flux falls rapidly. Finally, when divertor gas puffing is switched off, $f_{rad,exp}$ decreases and the recycling flux rebounds. For this 6 MW ECRH discharge, the threshold between high recycling and detachment is therefore $f_{rad,exp} = 0.45$, a useful reference point for planning and controlling detachment experiments.

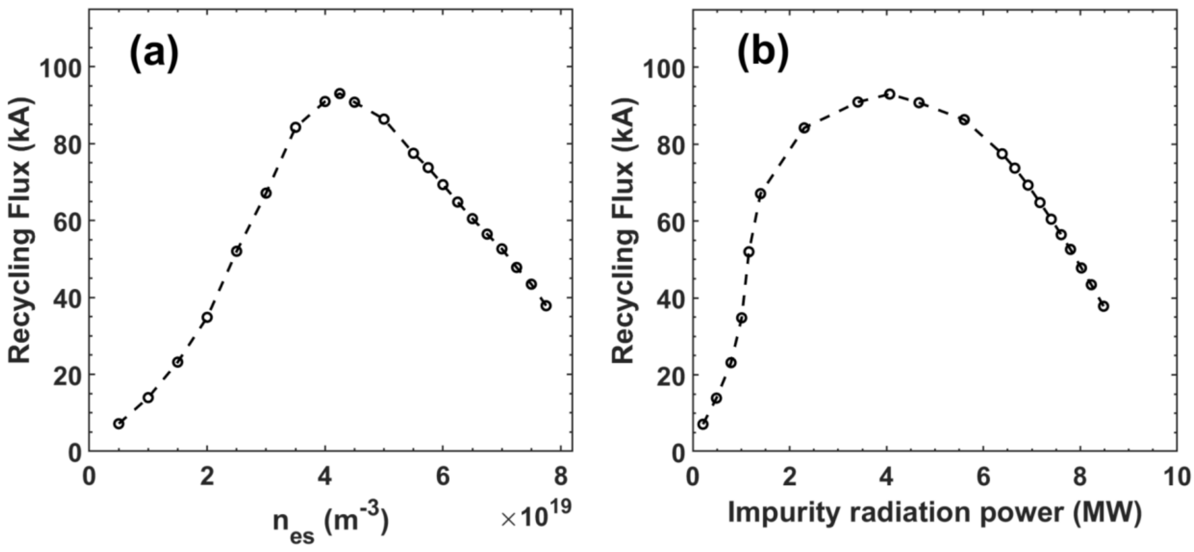

Figure 6.25: The recycling flux as a function of (a) the separatrix density and (b) radiation power. Imaged adapted from [47].

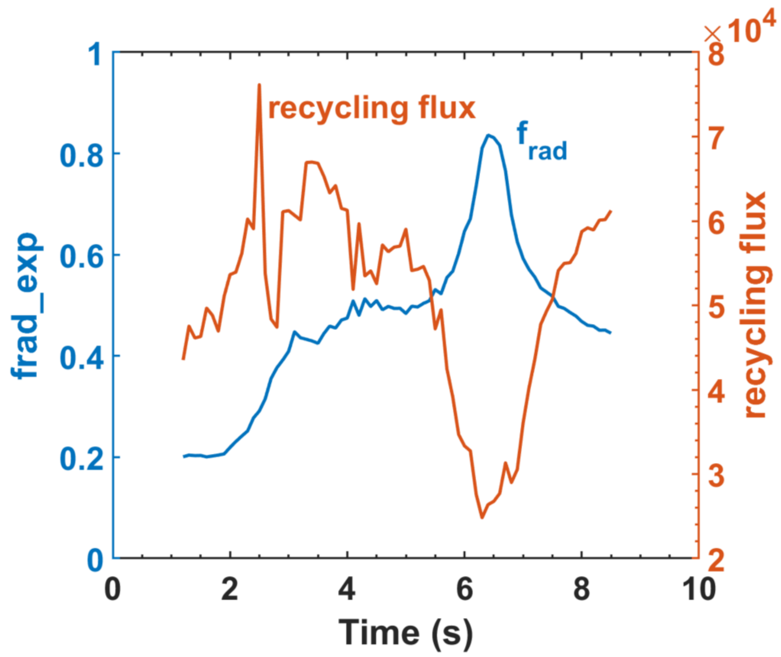

Figure 6.26: Evolution of the recycling flux during the radiation ramp. The orange curve gives the recycling flux, whereas the cyan curve shows the measured radiation fraction.

Real-time mapping

One of EMC3-EIRENE-AI's key capabilities is its ability to turn diagnostic measurements (together with their uncertainties) into full plasma parameter maps across the entire simulation domain. This provides an immediate, intuitive picture of how the plasma profiles consistent with both the measurements and the EMC3-EIRENE physics model evolve in real time. As an illustration, we apply EMC3-EIRENE-AI to the three moments marked in Figure 6.23 and compare the resulting profiles at different radiated-power fractions. Figures 6.27-6.29 display the electron temperature, electron density, and impurity radiation distributions for those three instants. With increasing total radiation fraction, the scrape-off layer (SOL) progressively cools; the temperature drop is particularly pronounced around the magnetic island, while the local density rises correspondingly. These trends are governed by the evolving impurity distribution. At low radiation ($f_{rad} \approx 0.20$) the plasma remains relatively hot, so impurities are ionized close to the divertor strike points quickly and the most intense radiation is confined there. As the radiation fraction rises to about 0.50, impurities penetrate further upstream; radiation in the SOL grows stronger, although its peak intensity is still anchored near the strike points. When the fraction reaches roughly 0.80 and detachment takes hold, impurities penetrate even deeper toward the core and the main radiation front relocates into the SOL. In a fully detached phase the process continues, and impurities and their associated radiation accumulate near the separatrix itself, producing a pronounced peak at that location. These real-time maps make it easy to see how the temperature, density, and radiation fronts shift as the plasma progresses from low-recycling, through high-recycling, and into detachment.

Of course, the value of real-time mapping goes well beyond displaying 2D or 3D plasma parameters on demand. As the EMC3-EIRENE-AI dataset grows and the tool is coupled to feedback control systems, it can close the loop between measurement, EMC3-EIRENE physics modeling, and actuator response. A few feasible applications are outlined below:

1. **Cross-validation of diagnostics and models.**
   Whenever a diagnostic signal deviates markedly from the EMC3-EIRENE-AI prediction, the discrepancy points to either an experimental inaccuracy or missing physics in the simulation, providing an integrated consistency check.
2. **Shot guidance in control room.**
   If a given divertor regime want to be maintained, the tool can display the current radiation front and recommend actuator actions, such as impurity puffing or controlled density ramp-ups, to keep the front in the desired location.
3. **Predictive inputs for feedback control.**
   Because EMC3-EIRENE-AI maps plasma conditions from inputs (heating power, radiated fraction, magnetic configuration, etc.), it can feed controllers directly. For example, if the model foresees an impending hot spot, the system could automatically call for power reduction or additional gas fuelling.
4. **Rapid analysis of large discharge sets.**
   The model's speed enables users to scan and analyze large volumes of discharge data quickly.
5. **Coupling with core codes.**
   Integrated with a core-transport solver, EMC3-EIRENE-AI would enable fast core-edge, whole-plasma predictions suitable for scenario development and real-time supervision.

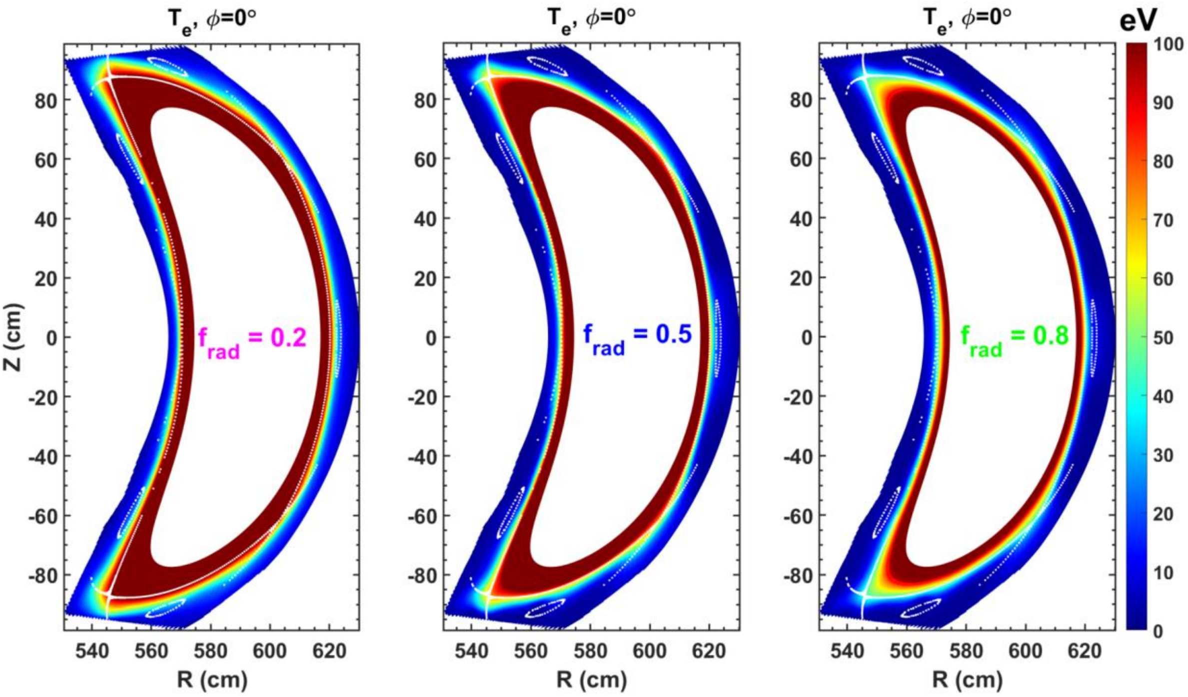

Figure 6.27: Electron temperature contours in the bean-shaped cross-section predicted by EMC3-EIRENE-AI using the measured diagnostics and their uncertainties for three radiation fractions.

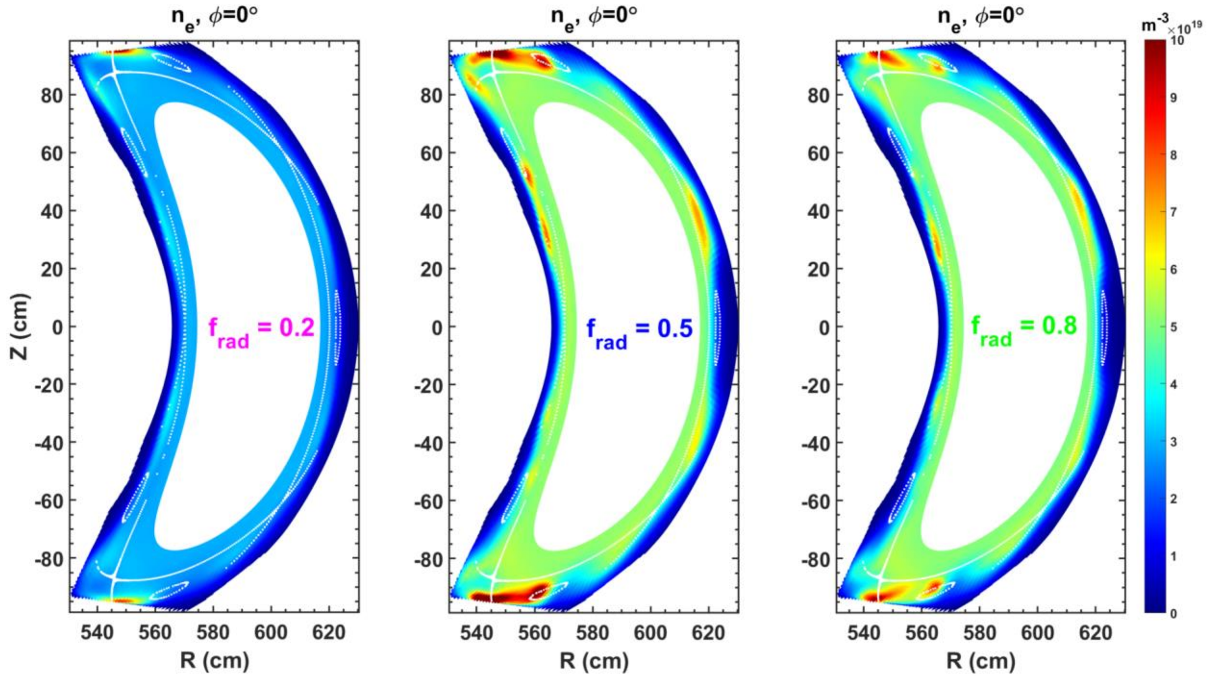

Figure 6.28: Electron density contours in the bean-shaped cross-section predicted by EMC3-EIRENE-AI using the measured diagnostics and their uncertainties for three radiation fractions.

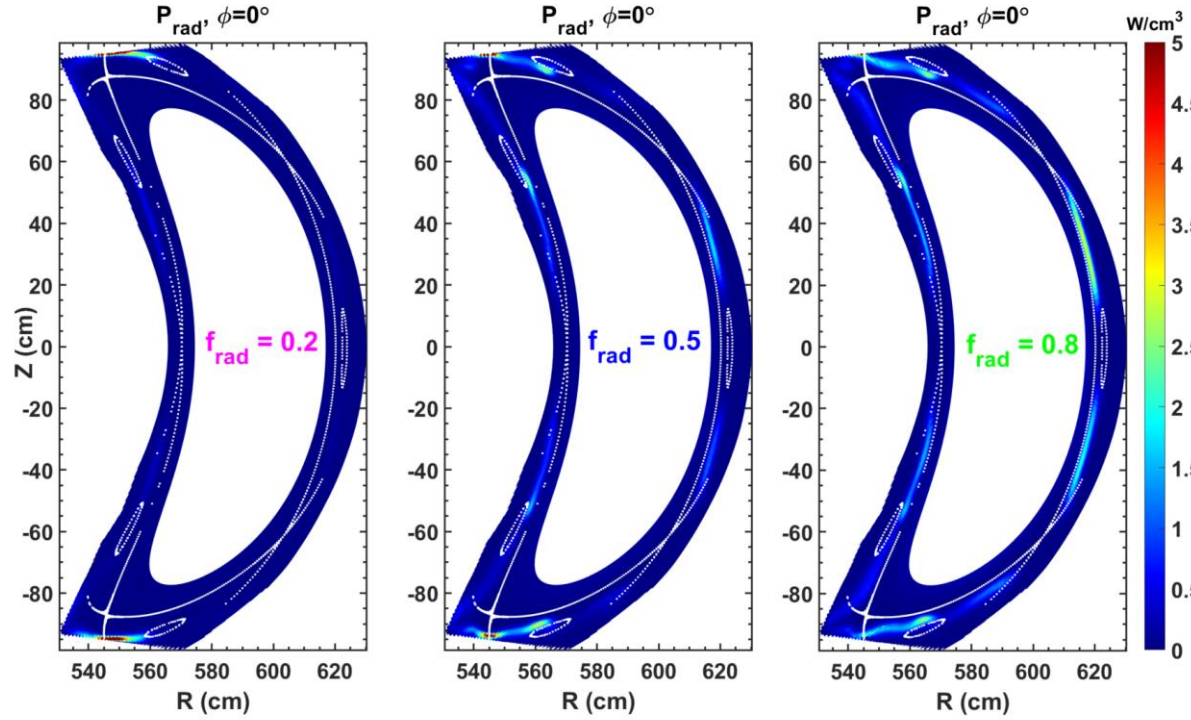

Figure 6.29: Radiation power contours in the bean-shaped cross-section predicted by EMC3-EIRENE-AI using the measured diagnostics and their uncertainties for three radiation fractions.

6.3 Summary

To enable rapid and accurate impurity transport analysis with EMC3-EIRENE using diagnostic data, this chapter has developed EMC3-EIRENE-AI. From routine edge-diagnostic measurements and their uncertainties, EMC3-EIRENE-AI directly predicts plasma parameter profiles that are fully consistent with the underlying EMC3-EIRENE physics model. The framework is built on two key components:

- Bayesian input inference.
  A Bayesian neural-surrogate module quickly infers the EMC3-EIRENE input parameters that best reproduce the experimental diagnostics within their error bars.
- Fast input-to-output mapping.
  A separate surrogate model then maps the inferred inputs to complete EMC3-EIRENE outputs, providing full 3-D profiles in milliseconds. Parameter-sensitivity analysis of this model further clarifies how edge-physics parameters affect impurity transport.

Beyond delivering near-real-time impurity transport analysis and 3D plasma maps, our ultimate goal is to close the loop among edge diagnostics, EMC3-EIRENE physics modeling, and feedback control. Such integration would furnish W7-X with a visual, quantitative, and continuously self-improving decision-support platform for long-pulse, high-power, and safe, stable operation.

7 Conclusion and outlook

7.1 Conclusion

The primary aim of this thesis is to investigate edge-plasma behavior, especially impurity transport, by constraining EMC3-EIRENE with multiple W7-X edge diagnostics.

To this end, comprehensive synthetic reconstructions were implemented in the EMC3-EIRENE post-processing for both upstream and divertor diagnostics. One major contribution is the synthetic reconstruction of spectroscopic line-emission signals. Starting from the radiance formula commonly used for experimental calibration, an integral form was derived and implemented in EMC3-EIRENE post-processing. The overlap between each line of sight (LOS) and the EMC3-EIRENE mesh is computed using a Monte Carlo algorithm. Because line radiation is highly sensitive to local plasma conditions, the new synthetic channel places tighter constraints on EMC3-EIRENE input parameters, especially on the cross-field transport coefficients. The same method is directly applicable to the 2D endoscope cameras.

By performing a systematic manual scan of EMC3-EIRENE input parameters, simulation cases were found that match the suite of edge diagnostics to a satisfactory level. In particular, simulations that employ spatially non-uniform cross-field transport coefficients reproduce both upstream and divertor measurements more accurately than simulations that assume uniform coefficients. Coupling the HINT equilibrium solver to EMC3-EIRENE allowed us to assess equilibrium effects on impurity transport, most notably those of on-axis $\beta_{ax} = 2\ \%$ and toroidal plasma current. In W7-X, finite β and a toroidal plasma current both modify the edge magnetic topology, though via different mechanisms. Finite β enhances edge field-line stochastization and shortens open-field-line connection lengths in the SOL, thereby altering the impurity source and edge transport, and ultimately producing a pronounced redistribution of temperature, density, and impurity distributions. In cases with a toroidal plasma current, the current-driven increase in iota leads to only minor changes in these profiles. During the density-ramp discharge an experimentally observed radiation front moves from the strike points toward the X-points and finally into the separatrix on standard configuration. The simulations reproduce this main path but reveal a second branch: sputtered impurities migrate from the strike points toward the O-points and then enter the last-closed flux surface near the O-points. Extensive scans over transport coefficients, sputtering yields, initial energies and so on cannot account for this discrepancy, suggesting that error fields or drift effects, not yet implemented in the present EMC3-EIRENE version, may be responsible.

Manual scans of EMC3-EIRENE inputs are time-consuming, computationally costly, and seldom yield a truly optimal match to experiment. To replace that trial and error procedure, this thesis introduces a two-stage machine learning pipeline, dubbed EMC3-EIRENE-AI, that delivers EMC3 EIRENE consistent edge plasma parameters directly from diagnostic data and their uncertainties. In first stage, a feed-forward network is trained to reproduce the synthetic diagnostic signals that EMC3-EIRENE generates for a given input set. Embedding this network in a Bayesian framework with Dynamic Nested Sampling produces the posterior distribution of input parameters that best matches the measured signals within their error bars. In second stage, a jointly fine-tuned surrogate rapidly maps those inferred inputs to full EMC3-EIRENE outputs: an auto-encoder compresses each high-dimensional output into a compact latent vector, a mapper projects the inputs into this latent space, and joint tuning of mapper and decoder enables millisecond-level reconstruction of the complete edge-plasma fields. Chaining the two stages yields EMC3-EIRENE-AI, which turns raw diagnostics (with uncertainties) into real-time, three-dimensional plasma profiles consistent with EMC3-EIRENE physics model. The framework supports a wide range of applications, such as fast transport analysis, boundary-regime identification, and real-time mapping, paving the way for to massive parameter scans, rapid discharge planning, and closed-loop control during W7-X operations.

Key original contributions of this work are as follows:

- Synthetic spectroscopy in EMC3-EIRENE post-processing. Dedicated synthetic diagnostics were implemented in the EMC3-EIRENE post-processing for the edge and divertor spectroscopy as well as endoscope systems, reproducing the measured radiance signals. The additional comparison channels tighten the constraints on EMC3-EIRENE input parameters and improve impurity transport simulations. (chapter 3, 4)
- Bayesian neural inference of EMC3-EIRENE inputs. A Bayesian-neural framework was devised that exploits the experimental signals with their uncertainties to infer the most probable EMC3-EIRENE input parameters. This corrects the measurement of the total impurity radiated power and rapidly provides EMC3-EIRENE input parameters consistent with the measurements. (chapter 5)
- Fast surrogate predicting EMC3-EIRENE outputs from inputs. A fine-tuned neural-network surrogate was trained to replace the full EMC3-EIRENE solver, delivering instant predictions of the code outputs for any given input set. This enables rapid, large-scale parameter-sensitivity analysis and deepens understanding of impurity-transport behavior across diverse plasma scenarios. (chapter 6)
- Closed real-time loop. By coupling the Bayesian input inference with the surrogate output predictor, the workflow closes the loop and provides real-time reconstructions that stay consistent with both the measurements and the underlying EMC3-EIRENE physics model. (chapter 6)

7.2 Outlook

During impurity transport studies based on diagnostic constraints, one of the most difficult tasks is to specify the free input, the cross-field transport coefficients, in EMC3-EIRENE. In the code, anomalous transport is modelled by diffusion, which introduces a cross-field particle diffusivity and a thermal diffusivity. Although EMC3-EIRENE allows these coefficients to be spatially non-uniform, the actual physical profile of anomalous transport is unknown, so there is no obvious recipe for prescribing them. Typical choices include a single uniform value over the whole domain; piece-wise constants, e.g. distinct values inside the last closed flux surface, inside the magnetic island, and in the open-field region; profiles that scale with turbulence drivers such as pressure gradient scale lengths, magnetic curvature, and collisionality. Our scans show that giving each region its own coefficient improves the match between simulation and measurements. Because no robust general method is yet available, the present EMC3-EIRENE-AI dataset still uses uniform coefficients for simplicity. A key goal for future work is to parameterize the spatial dependence of cross-field transport coefficients in the dataset, so that the AI can adjust them self-consistently to the diagnostic data.

This study demonstrates that EMC3-EIRENE-AI can rapidly produce edge-plasma parameter distributions that are consistent with the EMC3-EIRENE code while taking the diagnostic measurements and their uncertainties into account. Compared with the conventional trial-and-error tuning of inputs, the AI workflow is far faster and more objective. Because the surrogate is trained entirely on EMC3-EIRENE data, however, its validity is bounded by the code's capability (for example, the present EMC3-EIRENE version neglects drift effects). For EMC3-EIRENE-AI, the main challenge is to minimize the number of full EMC3-EIRENE runs needed to construct the training dataset while still preserving the required predictive accuracy. Potential strategies include:

1. One promising strategy is to exploit the data that are already generated on the way to a converged solution. During every EMC3-EIRENE run, many intermediate (still-unconverged) iterations are produced before the final, fully converged state is reached. Although these snapshots do not satisfy the strict convergence criterion, they nevertheless contain valuable information on the underlying physical trends. By tagging each snapshot with simple metadata, such as the iteration number or the relative residual between successive steps, both converged and unconverged states can be fed into the surrogate training pipeline. Incorporating these by-product samples means that far fewer fully converged simulations are required to reach a given prediction accuracy, thereby cutting the overall computational cost of building the dataset.
2. Another practical route is multi-fidelity learning, in which inexpensive, lower-physics runs are combined with a limited set of high-fidelity simulations. For example, extensive EMC3-Lite calculations (ignoring impurity transport and neutral interactions) can pre-train the network and map the coarse input-output relations. A much smaller collection of full-physics EMC3-EIRENE cases then fine-tunes the model, injecting the missing details and correcting residual biases. The same idea applies to mesh resolution: many

coarse-grid cases capture global trends quickly, while a handful of fine-grid simulations provide the necessary accuracy. By blending fidelities in this way, the surrogate retains high predictive quality while drastically reducing the number of computationally expensive, fully resolved runs needed to build the training set.

3. A long-term challenge lies in transfer learning across magnetic topologies and even across devices. Parameters such as connection length, field-line closure, magnetic island location, strike point position, and other topology-dependent quantities vary from one configuration to another, and each of these features can reshape EMC3-EIRENE results. Retraining an entire data set for every magnetic configuration is clearly unrealistic. A more practical roadmap is to encode the magnetic geometry with a handful of concise, quantitative descriptors (e.g., 1D radial shear profiles, 2D Poincaré cuts, and selected 3D grid metrics) and append these descriptors to the input vector passed to the network. Equipped with such topology tags, the surrogate becomes conditional on the magnetic state. The remaining gap is data. Here an active-learning loop could keep the cost in check: the surrogate flags regions of high predictive uncertainty, and only there are new, high-fidelity EMC3-EIRENE simulations launched. In principle, combining topology parameterization with conditional surrogates and targeted active learning could transfer EMC3-EIRENE-AI from the W7-X standard configuration to other configurations, and even to different devices, using only a modest number of additional simulations instead of rebuilding the data base from scratch.

Acknowledgments

For me, the pursuit of a doctoral degree has been a journey of building confidence.

First and foremost, I am profoundly grateful to my advisor, Prof. Dr. Yunfeng Liang. He gave me the opportunity to embark on this doctoral journey and provided a clear direction for my research. Our many fruitful discussions kept this project moving forward. Professor Liang's unique scientific perspective and passion for research have left a deep impression on me and will continue to guide my future career. I also wish to thank my second supervisor, Prof. Dr. Georg Pretzler. His advice and support were essential to the completion of this thesis.

I gratefully acknowledge Prof. Dr. Sebastijan Brezinsek for supporting my research project, and Prof. Dr. Detlev Reiter for several highly valuable comments on my work. I thank Dr. Dirk Reise for helpful discussions during theory group meetings, and Dr. Derek Harting for assistance with applications for computing time on the JURECA supercomputer.

I am grateful to Dr. Yuhe Feng for many discussions and for numerous constructive suggestions on this work. I also thank the IPP Greifswald divertor spectroscopy group, including Dr. Maciej Krychowiak, Dr. Dorothea Gradic, and colleagues, for help in obtaining the divertor spectroscopy data and for their helpful comments. I further thank Dr. Daihong Zhang for providing bolometer data and for advice on optimizing the EMC3-EIRENE input parameters.

My sincere thanks go to Dr. Shuai Xu, who patiently taught me how to use the EMC3-EIRENE code and offered invaluable assistance. Whenever I felt lost, he encouraged me and helped resolve concrete difficulties. I am grateful to Dr. Erhui Wang for his tremendous help. He imparted diagnostic expertise and led me to participate in the W7-X experiments. Special thanks to Dr. Jianqing Cai for his suggestion on my AI-related work and the manuscript. I also appreciate Dr. Alexander Knieps for his support, especially the constructive feedback he provided during each of my progress reports. I also thank Dr. Philipp Drews for his help with my work and for assistance with computer allpication process.

I would like to thank Dr. Haoming Xiang and Dr. Liang Liao—with their car, I was able to explore neighboring countries as a happy passenger. My gratitude also goes to Dr. Jie Huang, Dr. Jianwen Liu, Dr. Jia Huang, Dr. Jie Yang, Dr. Xingting Yan, Dr. Pengcheng Xie, Yichen Gao, Jiankun Hua, Pei Ren, Yutong Yang and Muzhi Tan for their help in both work and everyday life.

Finally, I am deeply grateful to my family, especially my wife, Dr. Qiqi Wang, who stood by me even when she needed support herself.

The project this report was supported with funds from the German Federal Ministry of Education and Research in the Förderprogramm "Fusion 2040-Forschung auf dem Weg zum Fusionskraftwerk", contract number 13F1012C. The sole responsibility for the report's contents lies with the authors. This work has also been carried out within the framework of the

EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No. 101052200—EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

The authors gratefully acknowledge the computing time granted by the JARA-HPC Vergabegremium and VSR commission on the supercomputer JURECA at Forschungszentrum Jülich.

Hardware and software

All EMC3-EIRENE simulations were executed on the batch nodes of the JURECA supercomputer [163]. To account for equilibrium effects, EMC3-EIRENE [54, 91] was coupled to the HINT equilibrium solver [57].

All ML-related data analysis and surrogate-model development (design, implementation, and training) were performed in Python using PyTorch 1.13.1 [164], PyTorch Geometric [165], and torch-sparse [166]; NumPy [167] and SciPy [168] for numerical routines; scikit-learn [169] for preprocessing and baselines; Ray Tune [170] for hyperparameter optimization; Weights & Biases [162] for experiment tracking; dynesty [146] for Bayesian inference; and Matplotlib [171] and corner [170] for visualization. Several graph libraries were built against the PyTorch 1.13.1 toolchain, , and a subset of figures was generated in MATLAB.

Bibliography

[1] S. F. Lincoln, "Fossil fuels in the 21st century", Ambio 34, 621-627 (2005). DOI: 10.1579/0044-7447-34.8.621.

[2] N. Abas, et al., "Review of fossil fuels and future energy technologies", Futures 69, 31-49 (2015). DOI: 10.1016/j.futures.2015.03.003.

[3] I. Dincer, et al., "A review on clean energy solutions for better sustainability", International Journal of Energy Research 39, 585-606 (2015). DOI: 10.1002/er.3329.

[4] V. Saenko, et al., "The Chernobyl Accident and its Consequences", Clinical Oncology 23, 234-243 (2011). DOI: 10.1016/j.clon.2011.01.502.

[5] F. N. von Hippel, "The radiological and psychological consequences of the Fukushima Daiichi accident", Bulletin of the Atomic Scientists 67, 27-36 (2011). DOI: 10.1177/0096340211421588.

[6] A. Pastore, et al., "From nuclei to neutron stars: simple binding energy computer modelling in the classroom (part 1)", Physics Education 56, 035012 (2021). DOI: 10.1088/1361-6552/abe09d.

[7] T. Tanabe, Tritium: Fuel of Fusion Reactors. Springer Japan, 2016.

[8] J. D. Lawson, "Some Criteria for a Power Producing Thermonuclear Reactor", Proceedings of the Physical Society of London Section B 70, 6-10 (1957). DOI: 10.1088/0370-1301/70/1/303.

[9] D. Clery, "ENERGY Laser-powered fusion effort nears 'ignition'", Science 373, 841-841 (2021). DOI: 10.1126/science.373.6557.841.

[10] N. Holtkamp, "The status of the ITER design", Fusion Engineering and Design 84, 98-105 (2009). DOI: 10.1016/j.fusengdes.2008.12.119.

[11] J. Wesson, et al., Tokamaks. Oxford university press, 2011.

[12] G. S. Xu, et al., "A promising approach to steady-state fusion: High-temperature superconducting strong-field stellarator with precise omnigenity", Innovation 5, 100537 (2024). DOI: 10.1016/j.xinn.2023.100537.

[13] S. Sudo, et al., "Scalings of Energy Confinement and Density Limit in Stellarator Heliotron Devices", Nuclear Fusion 30, 11-21 (1990). DOI: 10.1088/0029-5515/30/1/002.

[14] M. Greenwald, et al., "A New Look at Density Limits in Tokamaks", Nuclear Fusion 28, 2199-2207 (1988). DOI: 10.1088/0029-5515/28/12/009.

[15] Y. Xu, "A general comparison between tokamak and stellarator plasmas", Matter and Radiation at Extremes 1, 192-200 (2016). DOI: 10.1016/j.mre.2016.07.001.

[16] L. Spitzer Jr "A proposed stellarator": Princeton Univ., NJ Forrestal Research Center 1951.

[17] L. Spitzer, "The Stellarator Concept", Physics of Fluids 1, 253-264 (1958). DOI: 10.1063/1.1705883.

[18] T. Coor, et al., "Experiments on the Ohmic Heating and Confinement of Plasma in a Stellarator", Physics of Fluids 1, 411-420 (1958). DOI: 10.1063/1.1724358.

[19] B. A. Carreras, et al., "Progress in Stellarator Heliotron Research - 1981-1986", Nuclear Fusion 28, 1613-1694 (1988). DOI: 10.1088/0029-5515/28/9/011.

[20] M. Fujiwara, et al., "Large Helical Device (LHD) program", Journal of Fusion Energy 15, 7-153 (1996). DOI: 10.1007/Bf02266926.

[21] C. Alejaldre, et al., "TJ-II Project - a Flexible Heliac Stellarator", Fusion Technology 17, 131-139 (1990). DOI: 10.13182/Fst17-131-139.

[22] T. S. Pedersen, et al., "Confirmation of the topology of the Wendelstein 7-X magnetic field to better than 1:100,000", Nature Communications 7, 13493 (2016). DOI: 10.1038/ncomms13493.

[23] J. Geiger, et al., "Physics in the magnetic configuration space of W7-X", Plasma Physics and Controlled Fusion 57, 014004 (2015). DOI: 10.1088/0741-3335/57/1/014004.

[24] F. Sardei, et al., "Island divertor studies on W7-AS", Journal of Nuclear Materials 241, 135-148 (1997). DOI: 10.1016/S0022-3115(97)80036-0.

[25] P. Grigull, et al., "First island divertor experiments on the W7-AS stellarator", Plasma Physics and Controlled Fusion 43, A175-A193 (2001). DOI: 10.1088/0741-3335/43/12a/313.

[26] F. Sardei, et al., "3D modelling of the island divertor for W7-AS", Contributions to Plasma Physics 36, 166-170 (1996). DOI: 10.1002/ctpp.2150360213.

[27] T. S. Pedersen, et al., "First divertor physics studies in Wendelstein 7-X", Nuclear Fusion 59, 096014 (2019). DOI: 10.1088/1741-4326/ab280f.

[28] D. Sharma, et al., "A 3D Monte-Carlo study of the W7-X island divertor transport for different magnetic configurations", Nuclear Fusion 46, S127-S138 (2006). DOI: 10.1088/0029-5515/46/4/S03.

[29] Y. Liang, et al., "Towards advanced divertor configurations on the J-TEXT tokamak", Plasma Science & Technology 24, 124021 (2022). DOI: 10.1088/2058-6272/acaa8d.

[30] S. Zhou, et al., "First application of the island divertor configuration in the J-TEXT tokamak", Nuclear Fusion 65, 016020 (2025). DOI: 10.1088/1741-4326/ad90f1.

[31] Y. T. Yang, et al., "Characteristics of divertor heat flux distribution with an island divertor configuration on the J-TEXT tokamak", Plasma Science & Technology 26, 125102 (2024). DOI: 10.1088/2058-6272/ad6816.

[32] Y. Feng, et al., "Physics of island divertors as highlighted by the example of W7-AS", Nuclear Fusion 46, 807-819 (2006). DOI: 10.1088/0029-5515/46/8/006.

[33] R. A. Pitts, et al., "Physics basis for the first ITER tungsten divertor", Nuclear Materials and Energy 20, 100696 (2019). DOI: 10.1016/j.nme.2019.100696.

[34] T. Eich, et al., "Scaling of the tokamak near the scrape-off layer H-mode power width and implications for ITER", Nuclear Fusion 53, 093031 (2013). DOI: 10.1088/0029-5515/53/9/093031.

[35] P. C. Stangeby, The plasma boundary of magnetic fusion devices. Institute of Physics Pub. Philadelphia, Pennsylvania, 2000.

[36] B. Lipschultz, "Review of Marfe Phenomena in Tokamaks", Journal of Nuclear Materials 145, 15-25 (1987). DOI: 10.1016/0022-3115(87)90306-0.

[37] Y. Luo, et al., "Investigation of multifaceted asymmetric radiation from the edge (MARFE) with impurity injection from the upper divertor on the Experimental Advanced Superconducting Tokamak", Plasma Physics and Controlled Fusion 62, 075005 (2020). DOI: 10.1088/1361-6587/ab8a62.

[38] U. Wenzel, et al., "Observation of Marfes in the Wendelstein 7-X stellarator with inboard limiters", Nuclear Fusion 58, 096025 (2018). DOI: 10.1088/1741-4326/aacb86.

[39] F. C. Schuller, "Disruptions in tokamaks", Plasma Physics and Controlled Fusion 37, A135-A162 (1995). DOI: 10.1088/0741-3335/37/11a/009.

[40] S. Wiesen, et al., "The new SOLPS-ITER code package", Journal of Nuclear Materials 463, 480-484 (2015). DOI: 10.1016/j.jnucmat.2014.10.012.

[41] Y. Feng, et al., "A 3D Monte Carlo code for plasma transport in island divertors", Journal of Nuclear Materials 241, 930-934 (1997). DOI: 10.1016/S0022-3115(97)80168-7.

[42] V. R. Winters, et al., "Impurity leakage mechanisms in the Wendelstein 7-X island divertor under friction-dominated conditions", Nuclear Fusion 64, 056042 (2024). DOI: 10.1088/1741-4326/ad3768.

[43] D. Bold, et al., "Impact of spatially varying transport coefficients in EMC3-Eirene simulations of W7-X and assessment of drifts", Nuclear Fusion 64, 126055 (2024). DOI: 10.1088/1741-4326/ad845c.

[44] F. Effenberg, et al., "Numerical investigation of plasma edge transport and limiter heat fluxes in Wendelstein 7-X startup plasmas with EMC3-EIRENE", Nuclear Fusion 57, 036021 (2017). DOI: 10.1088/1741-4326/aa4f83.

[45] V. R. Winters, et al., "First experimental confirmation of island SOL geometry effects in a high radiation regime on W7-X", Nuclear Fusion 64, 126047 (2024). DOI: 10.1088/1741-4326/ad820e.

[46] F. Effenberg, et al., "First demonstration of radiative power exhaust with impurity seeding in the island divertor at Wendelstein 7-X", Nuclear Fusion 59, 106020 (2019). DOI: 10.1088/1741-4326/ab32c4.

[47] S. Xu, et al., "Modeling of plasma beta effects on the island divertor transport in the standard configuration of W7-X", Nuclear Fusion 63, 066005 (2023). DOI: 10.1088/1741-4326/acc7b8.

[48] Y. Feng, et al., "Conditions and benefits of X-point radiation for the island divertor", Nuclear Fusion 64, 086027 (2024). DOI: 10.1088/1741-4326/ad5606.

[49] Y. Feng, et al., "Understanding detachment of the W7-X island divertor", Nuclear Fusion 61, 086012 (2021). DOI: 10.1088/1741-4326/ac0772.

[50] L. Stephey, et al., “Impact of magnetic islands in the plasma edge on particle fueling and exhaust in the HSX and W7-X stellarators”, Physics of Plasmas 25, 062501 (2018). DOI: 10.1063/1.5026324.

[51] D. Boeyaert, et al., “Analysis of the neutral fluxes in the divertor region of Wendelstein 7-X under attached and detached conditions using EMC3-EIRENE”, Plasma Physics and Controlled Fusion 66, 015005 (2024). DOI: 10.1088/1361-6587/ad0e22.

[52] S. Xu, et al., “First three-dimensional edge plasma transport simulations with magnetic perturbations induced by lower hybrid waves on EAST”, Nuclear Fusion 58, 106008 (2018). DOI: 10.1088/1741-4326/aad296.

[53] S. Xu, et al., “Mechanism of the active divertor flux control by the supersonic-molecular-beam-injection with lower hybrid wave-induced magnetic perturbations on the EAST tokamak”, Nuclear Fusion 60, 056006 (2020). DOI: 10.1088/1741-4326/ab796a.

[54] Y. Feng, et al., “3D edge modeling and island divertor physics”, Contributions to Plasma Physics 44, 57-69 (2004). DOI: 10.1002/ctpp.200410009.

[55] M. Rack, et al., “A fluid-kinetic approach for 3D plasma edge transport in He plasma”, Nuclear Fusion 57, 056011 (2017). DOI: 10.1088/1741-4326/aa60e4.

[56] H. P. Summers, et al., “Atomic Data and Modelling for Fusion: the ADAS Project”, 7th International Conference on Atomic and Molecular Data and Their Applications Icamdata-2010 1344, 179-187 (2011). DOI: 10.1063/1.3585817.

[57] Y. Suzuki, et al., “Development and application of HINT2 to helical system plasmas”, Nuclear Fusion 46, L19-L24 (2006). DOI: 10.1088/0029-5515/46/11/L01.

[58] Y. Suzuki, “HINT modeling of three-dimensional tokamaks with resonant magnetic perturbation”, Plasma Physics and Controlled Fusion 59, 054008 (2017). DOI: 10.1088/1361-6587/aa5adc.

[59] Y. Feng, et al., “Physics of island divertor plasmas in stellarators”, Contributions to Plasma Physics 46, 504-514 (2006). DOI: 10.1002/ctpp.200610037.

[60] M. Kobayashi, et al., “Edge impurity transport study in the stochastic layer of LHD and the scrape-off layer of HL-2A”, Nuclear Fusion 53, 033011 (2013). DOI: 10.1088/0029-5515/53/3/033011.

[61] S. Y. Dai, et al., “EMC3-EIRENE modelling of edge impurity transport in the stochastic layer of the large helical device compared with extreme ultraviolet emission measurements”, Nuclear Fusion 56, 066005 (2016). DOI: 10.1088/0029-5515/56/6/066005.

[62] T. Xie, et al., “EMC3-EIRENE modelling of edge plasma and impurity emissions compared with the liquid lithium limiter experiment on EAST”, Nuclear Fusion 58, 106017 (2018). DOI: 10.1088/1741-4326/aad42f.

[63] T. Xie, et al., “Investigation of edge impurity transport and divertor fluxes by toroidally localized lithium injection on EAST with EMC3-EIRENE”, Plasma Physics and Controlled Fusion 61, 115006 (2019). DOI: 10.1088/1361-6587/ab434a.

[64] T. Lunt, et al., “First EMC3-Eirene simulations of the TCV snowflake divertor”, Plasma Physics and Controlled Fusion 56, 035009 (2014). DOI: 10.1088/0741-3335/56/3/035009.

[65] T. Lunt, et al., "Numerical study of potential heat flux mitigation effects in the TCV snowflake divertor", Plasma Physics and Controlled Fusion 58, 045027 (2016). DOI: 10.1088/0741-3335/58/4/045027.

[66] J. D. Lore, et al., "EMC3-EIRENE modeling of toroidally-localized divertor gas injection experiments on Alcator C-Mod", Journal of Nuclear Materials 463, 515-518 (2015). DOI: 10.1016/j.jnucmat.2014.09.053.

[67] K. C. Hammond, et al., "Drift effects on W7-X divertor heat and particle fluxes", Plasma Physics and Controlled Fusion 61, 125001 (2019). DOI: 10.1088/1361-6587/ab4825.

[68] D. M. Kriete, et al., "Effects of drifts on scrape-off layer transport in W7-X", Nuclear Fusion 63, 026022 (2023). DOI: 10.1088/1741-4326/acab75.

[69] O. Neubauer, et al., "Diagnostic setup for investigation of plasma wall interactions at Wendelstein 7-X", Fusion Engineering and Design 96-97, 891-894 (2015). DOI: 10.1016/j.fusengdes.2015.06.102.

[70] H.-J. Kunze, Introduction to plasma spectroscopy. Springer Science & Business Media, 2009.

[71] M. Krychowiak, et al. "Gaussian Process Tomography of carbon radiation in the transition to detached plasmas in the Wendelstein 7-X stellarator" *47th EPS Conference on Plasma Physics*: European Physical Society 2021.

[72] E. Wang, et al., "Impurity sources and fluxes in W7-X: from the plasma-facing components to the edge layer", Physica Scripta T171, 014040 (2020). DOI: 10.1088/1402-4896/ab4c04.

[73] E. Wang, et al., "Radiation characteristics of detached divertor plasmas in W7-X", Nuclear Materials and Energy 33, 101283 (2022). DOI: 10.1016/j.nme.2022.101283.

[74] P. Drews, et al., "Operation of probe heads on the Multi-Purpose-Manipulator at W7-X in OP 1.2a", Fusion Engineering and Design 146, 2353-2355 (2019). DOI: 10.1016/j.fusengdes.2019.03.188.

[75] P. Drews, et al., "Measurement of the plasma edge profiles using the combined probe on W7-X", Nuclear Fusion 57, 126020 (2017). DOI: 10.1088/1741-4326/aa8385.

[76] S. A. Bozhenkov, et al., "The Thomson scattering diagnostic at Wendelstein 7-X and its performance in the first operation phase", Journal of Instrumentation 12, P10004 (2017). DOI: 10.1088/1748-0221/12/10/P10004.

[77] E. Pasch, et al., "The Thomson scattering system at Wendelstein 7-X", Review of Scientific Instruments 87, 11E729 (2016). DOI: 10.1063/1.4962248.

[78] G. Anda, et al., "Measurement of edge plasma parameters at W7-X using alkali beam emission spectroscopy", Fusion Engineering and Design 146, 1814-1819 (2019). DOI: 10.1016/j.fusengdes.2019.03.042.

[79] M. Jakubowski, et al., "Infrared imaging systems for wall protection in the W7-X stellarator (invited)", Review of Scientific Instruments 89, 10e116 (2018). DOI: 10.1063/1.5038634.

[80] A. Pandey, et al., "Pop-up Langmuir probe diagnostic in the water cooled divertor of Wendelstein 7-X", Review of Scientific Instruments 95, 043503 (2024). DOI: 10.1063/5.0188738.

[81] R. Laube, et al., "Designs of Langmuir probes for W7-X", Fusion Engineering and Design 86, 1133-1136 (2011). DOI: 10.1016/j.fusengdes.2010.12.059.

[82] H. P. Summers, et al., "Ionization state, excited populations and emission of impurities in dynamic finite density plasmas: I. The generalized collisional-radiative model for light elements", Plasma Physics and Controlled Fusion 48, 263-293 (2006). DOI: 10.1088/0741-3335/48/2/007.

[83] G. Federici, et al., "Plasma-material interactions in current tokamaks and their implications for next step fusion reactors", Nuclear Fusion 41, 1967-2137 (2001). DOI: 10.1088/0029-5515/41/12/218.

[84] A. Loarte, et al., "Power and particle control", Nuclear Fusion 47, S203 (2007). DOI: 10.1088/0029-5515/47/6/S04.

[85] J. Rapp, et al., "Reduction of divertor heat load in JET ELMy H-modes using impurity seeding techniques", Nuclear Fusion 44, 312-319 (2004). DOI: 10.1088/0029-5515/44/2/013.

[86] I. Ivanova-Stanik, et al., "Mitigation of the divertor heat load in DEMO reactor by impurity seeding", Journal of Nuclear Materials 463, 596-600 (2015). DOI: 10.1016/j.jnucmat.2014.11.105.

[87] P. Monier-Garbet, et al., "High radiation from intrinsic and injected impurities in Tore Supra ergodic divertor plasmas", Journal of Nuclear Materials 290, 925-929 (2001). DOI: 10.1016/S0022-3115(00)00459-1.

[88] U. Stroth, et al., "Model for access and stability of the X-point radiator and the threshold for marfes in tokamak plasmas", Nuclear Fusion 62, 076008 (2022). DOI: 10.1088/1741-4326/ac613a.

[89] S. Putvinski, et al., "Impurity fueling to terminate Tokamak discharges", Journal of Nuclear Materials 241, 316-321 (1997). DOI: 10.1016/S0022-3115(97)80056-6.

[90] P. Shi, et al., "First time observation of local current shrinkage during the MARFE behavior on the J-TEXT tokamak", Nuclear Fusion 57, 116052 (2017). DOI: 10.1088/1741-4326/aa80a6.

[91] D. Reiter, et al., "The EIRENE and B2-EIRENE codes", Fusion Science and Technology 47, 172-186 (2005). DOI: 10.13182/Fst47-172.

[92] M. Kobayashi, et al., "3D numerical transport study of the edge ergodized plasma in TEXTOR-DED", Nuclear Fusion 44, S64-S73 (2004). DOI: 10.1088/0029-5515/44/6/S07.

[93] K. Schmid, et al., "3D global impurity transport modeling with WallDYN and EMC3-Eirene", Nuclear Materials and Energy 17, 200-205 (2018). DOI: 10.1016/j.nme.2018.11.005.

[94] Y. Feng, et al., "Transport in island divertors: physics, 3D modelling and comparison to first experiments on W7-AS", Plasma Physics and Controlled Fusion 44, 611-625 (2002). DOI: 10.1088/0741-3335/44/5/308.

[95] M. Kobayashi, et al., "Model prediction of impurity retention in stochastic magnetic boundary and comparison with edge carbon emission in LHD", Journal of Nuclear Materials 390-91, 325-329 (2009). DOI: 10.1016/j.jnucmat.2009.01.147.

[96] D. Sharma, et al., “3D studies of the island divertor transport in W7-X”, Contributions to Plasma Physics 46, 534-538 (2006). DOI: 10.1002/ctpp.200610040.

[97] J. Cosfeld, et al., “Impact of impurities on the effective charge state distribution in the limiter plasmas of Wendelstein 7-X”, Nuclear Materials and Energy 18, 307-311 (2019). DOI: 10.1016/j.nme.2019.01.007.

[98] Y. Liang, et al., “Diagnostic set-up and modelling for investigation of synergy between 3D edge physics and plasma-wall interactions on Wendelstein 7-X”, Nuclear Fusion 57, 066049 (2017). DOI: 10.1088/1741-4326/aa6cde.

[99] J. Nührenberg, “Development of quasi-isodynamic stellarators”, Plasma Physics and Controlled Fusion 52, 124003 (2010). DOI: 10.1088/0741-3335/52/12/124003.

[100] R. C. Wolf, et al., “Major results from the first plasma campaign of the Wendelstein 7-X stellarator”, Nuclear Fusion 57, 102020 (2017). DOI: 10.1088/1741-4326/aa770d.

[101] T. S. Pedersen, et al., “Experimental confirmation of efficient island divertor operation and successful neoclassical transport optimization in Wendelstein 7-X”, Nuclear Fusion 62, 042022 (2022). DOI: 10.1088/1741-4326/ac2cf5.

[102] H. Renner, et al., “Divertor concept for the W7-X stellarator and mode of operation”, Plasma Physics and Controlled Fusion 44, 1005-1019 (2002). DOI: 10.1088/0741-3335/44/6/325.

[103] R. König, et al., “The divertor program in stellarators”, Plasma Physics and Controlled Fusion 44, 2365-2422 (2002). DOI: 10.1088/0741-3335/44/11/306.

[104] D. Zhang, et al., “Design criteria of the bolometer diagnostic for steady-state operation of the W7-X stellarator”, Review of Scientific Instruments 81, 10E134 (2010). DOI: 10.1063/1.3483194.

[105] S. Zhou, et al., “Equilibrium effects on the structure of island divertor and its impact on the divertor heat flux distribution in Wendelstein 7-X”, Nuclear Fusion 62, 106002 (2022). DOI: 10.1088/1741-4326/ac8439.

[106] A. Knieps, et al., “Plasma beta effects on the edge magnetic field structure and divertor heat loads in Wendelstein 7-X high-performance scenarios”, Nuclear Fusion 62, 026011 (2022). DOI: 10.1088/1741-4326/ac3a18.

[107] Y. Gao, et al., “Effects of toroidal plasma current on divertor power depositions on Wendelstein 7-X”, Nuclear Fusion 59, 106015 (2019). DOI: 10.1088/1741-4326/ab32c2.

[108] J. Seo, et al., “Feedforward beta control in the KSTAR tokamak by deep reinforcement learning”, Nuclear Fusion 61, 106010 (2021). DOI: 10.1088/1741-4326/ac121b.

[109] J. Degrave, et al., “Magnetic control of tokamak plasmas through deep reinforcement learning”, Nature 602, 414-+ (2022). DOI: 10.1038/s41586-021-04301-9.

[110] J. Seo, et al., “Avoiding fusion plasma tearing instability with deep reinforcement learning”, Nature 626, 746-751 (2024). DOI: 10.1038/s41586-024-07024-9.

[111] J. Kates-Harbeck, et al., “Predicting disruptive instabilities in controlled fusion plasmas through deep learning”, Nature 568, 526-+ (2019). DOI: 10.1038/s41586-019-1116-4.

[112] Y. C. Fu, et al., “Machine learning control for disruption and tearing mode avoidance”, Physics of Plasmas 27, 022501 (2020). DOI: 10.1063/1.5125581.

[113] J. Q. Cai, et al., “Improved training framework in a neural network model for disruption prediction and its application on EXL-50”, Plasma Science & Technology 26, 055102 (2024). DOI: 10.1088/2058-6272/ad1571.

[114] O. Meneghini, et al., “Self-consistent core-pedestal transport simulations with neural network accelerated models”, Nuclear Fusion 57, 086034 (2017). DOI: 10.1088/1741-4326/aa7776.

[115] K. L. van de Plassche, et al., “Fast modeling of turbulent transport in fusion plasmas using neural networks”, Physics of Plasmas 27, 022310 (2020). DOI: 10.1063/1.5134126.

[116] S. Dasbach, et al., “Towards fast surrogate models for interpolation of tokamak edge plasmas”, Nuclear Materials and Energy 34, 101396 (2023). DOI: 10.1016/j.nme.2023.101396.

[117] E. Higson, et al., “Dynamic nested sampling: an improved algorithm for parameter estimation and evidence calculation”, Statistics and Computing 29, 891-913 (2019). DOI: 10.1007/s11222-018-9844-0.

[118] A. Menon, et al., “Characterization of a class of sigmoid functions with applications to neural networks”, Neural Networks 9, 819-835 (1996). DOI: 10.1016/0893-6080(95)00107-7.

[119] A. H. Namin, et al., “Efficient Hardware Implementation of the Hyperbolic Tangent Sigmoid Function”, Iscas: 2009 Ieee International Symposium on Circuits and Systems, Vols 1-5, 2117-2120 (2009). DOI: 10.1109/Iscas.2009.5118213.

[120] X. Glorot, et al. "Deep Sparse Rectifier Neural Networks" *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research 2011*

[121] A. Krizhevsky, et al., “ImageNet Classification with Deep Convolutional Neural Networks”, Communications of the Acm 60, 84-90 (2017). DOI: 10.1145/3065386.

[122] D. E. Rumelhart, et al., “Learning Representations by Back-Propagating Errors”, Nature 323, 533-536 (1986). DOI: 10.1038/323533a0.

[123] L. Bottou "Stochastic gradient descent tricks" *Neural Networks: Tricks of the Trade: Second Edition*: Springer 2012:421-436.

[124] D. P. Kingma, “Adam: A method for stochastic optimization”, arXiv preprint arXiv:1412.6980, (2014). DOI: 10.48550/arXiv.1412.6980.

[125] G. Hinton, et al. "Lecture 6a overview of mini–batch gradient descent" *Coursera Lecture slides https://class. coursera. org/neuralnets-2012-001/lecture,[Online]* 2012.

[126] A. Krogh, et al., “A simple weight decay can improve generalization”, Advances in neural information processing systems 4, (1991). DOI: 10.5555/2986916.2987033.

[127] K. O'shea, et al., “An introduction to convolutional neural networks”, arXiv preprint arXiv:1511.08458, (2015). DOI: 10.48550/arXiv.1511.08458.

[128] Z. Li, et al., “A survey of convolutional neural networks: analysis, applications, and prospects”, IEEE transactions on neural networks and learning systems 33, 6999-7019 (2021). DOI: 10.1109/TNNLS.2021.3084827.

[129] D. Scherer, et al. "Evaluation of pooling operations in convolutional architectures for object recognition" *International conference on artificial neural networks*: Springer 2010:92-101.

[130] L. Alzubaidi, et al., “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions”, Journal of big Data 8, 53 (2021). DOI: 10.1186/s40537-021-00444-8.

[131] N. Murray, et al. "Generalized max pooling" *Proceedings of the IEEE conference on computer vision and pattern recognition* 2014:2473-2480.

[132] F. Bieder, et al., “Comparison of methods generalizing max-and average-pooling”, arXiv preprint arXiv:2103.01746, (2021). DOI: 10.48550/arXiv.2103.01746.

[133] J. Bergstra, et al., “Random search for hyper-parameter optimization”, The journal of machine learning research 13, 281-305 (2012). DOI: 10.5555/2188385.2188395.

[134] L. Li, et al., “A system for massively parallel hyperparameter tuning”, Proceedings of machine learning and systems 2, 230-246 (2020). DOI: 10.48550/arXiv.1810.05934.

[135] S. Falkner, et al. "BOHB: Robust and efficient hyperparameter optimization at scale" *International conference on machine learning*: PMLR 2018:1437-1446.

[136] L. Bottou "Stochastic gradient learning in neural networks" *Proceedings of Neuro-Nımes* 1991:12.

[137] L. Li, et al., “Hyperband: A novel bandit-based approach to hyperparameter optimization”, Journal of Machine Learning Research 18, 1-52 (2018). DOI: 10.48550/arXiv.1603.06560.

[138] M. Brenzke, et al., “Divertor power load predictions based on machine learning”, Nuclear Fusion 61, 046023 (2021). DOI: 10.1088/1741-4326/abdb94.

[139] D. Orozco, et al., “Neural Network-Based Confinement Mode Prediction for Real-Time Disruption Avoidance”, Ieee Transactions on Plasma Science 50, 4157-4164 (2022). DOI: 10.1109/Tps.2022.3198596.

[140] J. C. Hou, et al., “Deep Learning-Based Plasma Optical Boundary Recognition and Reconstruction on EAST Tokamak”, Fusion Science and Technology 81, 425-436 (2025). DOI: 10.1080/15361055.2024.2431783.

[141] J. E. Lee, et al., “A machine learning approach to identify the universality of solitary perturbations accompanying boundary bursts in magnetized toroidal plasmas”, Scientific Reports 11, 3662 (2021). DOI: 10.1038/s41598-021-83192-2.

[142] H. Leung, et al., “The Complex Backpropagation Algorithm”, Ieee Transactions on Signal Processing 39, 2101-2104 (1991). DOI: 10.1109/78.134446.

[143] N. Benvenuto, et al., “On the Complex Backpropagation Algorithm”, Ieee Transactions on Signal Processing 40, 967-968 (1992). DOI: 10.1109/78.127967.

[144] D.-A. Clevert, et al., “Fast and accurate deep network learning by exponential linear units (elus)”, arXiv preprint arXiv:1511.07289, (2015). DOI: 10.48550/arXiv.1511.07289.

[145] J. S. Speagle, “DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences”, Monthly Notices of the Royal Astronomical Society 493, 3132-3158 (2020). DOI: 10.1093/mnras/staa278.

[146] S. Koposov, et al., “joshspeagle/dynesty: v2. 1.4”, Zenodo, (2023). DOI: 10.5281/zenodo.12537467.svg.

[147] J. Skilling, “Nested sampling”, Bayesian Inference and Maximum Entropy Methods in Science and Engineering 735, 395-405 (2004). DOI: 10.1063/1.1835238.

[148] J. Skilling, “Nested Sampling for General Bayesian Computation”, Bayesian Analysis 1, 833-859 (2006). DOI: 10.1214/06-Ba127.

[149] Y. Luo, et al. "3D edge impurity transport analysis using the EMC3-EIRENE code by taking into account multiple edge plasma measurements on W7-X" *49th EPS Conference on Plasma Physics*: European Physical Society 2023.

[150] D. Zhang, et al., “Bolometer tomography on Wendelstein 7-X for study of radiation asymmetry”, Nuclear Fusion 61, 116043 (2021). DOI: 10.1088/1741-4326/ac2778.

[151] S. A. Bozhenkov, et al., “Service oriented architecture for scientific analysis at W7-X. An example of a field line tracer”, Fusion Engineering and Design 88, 2997-3006 (2013). DOI: 10.1016/j.fusengdes.2013.07.003.

[152] Y. Feng, et al., “Review of magnetic islands from the divertor perspective and a simplified heat transport model for the island divertor”, Plasma Physics and Controlled Fusion 64, 125012 (2022). DOI: 10.1088/1361-6587/ac9ed9.

[153] G. L. Derks, et al., “Benchmark of a self-consistent dynamic 1D divertor model DIV1D using the 2D SOLPS-ITER code”, Plasma Physics and Controlled Fusion 64, 125013 (2022). DOI: 10.1088/1361-6587/ac9dbd.

[154] Y. Poels, et al., “Fast dynamic 1D simulation of divertor plasmas with neural PDE surrogates”, Nuclear Fusion 63, 126012 (2023). DOI: 10.1088/1741-4326/acf70d.

[155] T. Rognlien, et al., Users manual for the UEDGE edge-plasma transport code. 2000.

[156] B. Zhu, et al., “Data-driven model for divertor plasma detachment prediction”, Journal of Plasma Physics 88, 895880504 (2022). DOI: 10.1017/S002237782200085x.

[157] S. Wiesen, et al., “Data-driven models in fusion exhaust: AI methods and perspectives”, Nuclear Fusion 64, 086046 (2024). DOI: 10.1088/1741-4326/ad5a1d.

[158] F. Felici, et al., “Real-time-capable prediction of temperature and density profiles in a tokamak using RAPTOR and a first-principle-based transport model”, Nuclear Fusion 58, 096006 (2018). DOI: 10.1088/1741-4326/aac8f0.

[159] O. Meneghini, et al., “Neural-network accelerated coupled core-pedestal simulations with self-consistent transport of impurities and compatible with ITER IMAS”, Nuclear Fusion 61, 026006 (2021). DOI: 10.1088/1741-4326/abb918.

[160] T. S. Pedersen, et al., “First results from divertor operation in Wendelstein 7-X”, Plasma Physics and Controlled Fusion 61, 014035 (2019). DOI: 10.1088/1361-6587/aaec25.

[161] R. Anirudh, et al., “Improved surrogates in inertial confinement fusion with manifold and cycle consistencies”, Proceedings of the National Academy of Sciences 117, 9741-9746 (2020). DOI: 10.1073/pnas.1916634117.

[162] “Weights & Biases”, (2017–2025). https://docs.wandb.ai/guides/app/features/panels/parameter-importance/.

[163] P. Thörnig, "JURECA: Data centric and booster modules implementing the modular supercomputing architecture at Jülich supercomputing centre", Journal of large-scale research facilities JLSRF 7, A182 (2021). DOI: 10.17815/jlsrf-7-182.

[164] A. Paszke, et al., "PyTorch: An Imperative Style, High-Performance Deep Learning Library", Advances in Neural Information Processing Systems 32 (Nips 2019) 32, (2019).

[165] M. Fey, et al., "Fast graph representation learning with PyTorch Geometric", arXiv preprint arXiv:1903.02428, (2019). DOI: 10.48550/arXiv.1903.02428.

[166] M. Fey, "torch-sparse", (2023). https://pypi.org/project/torch-sparse/.

[167] C. R. Harris, et al., "Array programming with NumPy", Nature 585, 357-362 (2020). DOI: 10.1038/s41586-020-2649-2.

[168] P. Virtanen, et al., "SciPy 1.0: fundamental algorithms for scientific computing in Python", Nature Methods 17, 261-272 (2020). DOI: 10.1038/s41592-019-0686-2.

[169] F. Pedregosa, et al., "Scikit-learn: Machine Learning in Python", Journal of Machine Learning Research 12, 2825-2830 (2011).

[170] R. Liaw, et al., "Tune: A research platform for distributed model selection and training", arXiv preprint arXiv:1807.05118, (2018). DOI: 10.48550/arXiv.1807.05118.

[171] J. D. Hunter, "Matplotlib: A 2D graphics environment", Computing in Science & Engineering 9, 90-95 (2007). DOI: 10.1109/Mcse.2007.55.