

Extreme Learning Machine hálók alkalmazása a nyalábemissziós spektroszkópiában

Tudományos Diákköri Dolgozat

Készítette:

Karácsonyi Máté

(ELTE Fizika MSc, HUN-REN Energiatudományi Kutatóközpont)
(karacsonyi.mate@ek.hun-ren.hu)

Témavezető:

Dr. Asztalos Örs

(BME Nukleáris Technika Intézet, HUN-REN Energiatudományi
Kutatóközpont)
(asztalos.ors@ek.hun-ren.hu)

2025. január 8.

Tartalomjegyzék

1. Bevezető	4
1.1. Fúzió	4
1.1.1. Fúziós berendezések	5
2. Plazma diagnosztika	6
2.1. Nyaláb emissziós spektroszkópia (BES)	7
2.1.1. A háttérben lévő fizika, a ráta egyenletek	9
2.1.2. RENATE-OD	11
3. Neurális hálók	11
3.1. A neuron	12
3.2. Perceptron hálók	13
3.3. Extreme Learning machines, avagy ELM-ek	14
4. A probléma	16
5. Eredmények	17
5.1. Adathalmazok	17
5.1.1. Tangens hiperbolikus adathalmazok generálása	17
5.1.2. Fluktuációs tangens hiperbolikus adathalmazok generálása	21
5.1.3. HESEL adathalmazok	22
5.1.4. W7-X adathalmaz	24
5.2. Adattisztítás	25
5.2.1. Szintetikus adathalmazok	25
5.2.2. W7-X adathalmaz	25
5.3. Teljesítmény metrikák	26
5.4. Modellek tesztelése statikus profilokra	26
5.5. ELM-ek teljesítménye a különböző adatszetekre	30
5.5.1. Modell eredmények a tangens hiperbolikus adathalmazokon	31
5.5.2. Modell eredmények a fluktuációs tangens hiperbolikus adathalmazokon	33
5.5.3. Modell eredmények a HESEL adathalmazokon	35
5.5.4. Modell eredmények a W7-X adathalmazon	37
5.6. Modell kimenetek "folytonossága"	38
5.6.1. Fluktuációs tangens hiperbolikus adatkészlet göröngyössége	39
5.6.2. HESEL adatkészletek göröngyössége	40
5.7. Felbontásra való érzékenység	42
5.8. Modell gyorsaság	43

6. Konklúzió	44
7. Köszönetnyilvánítás	45
8. Bibliográfia	45
A. Saját eredmények	51

Kivonat

A nyalábemissziós spektroszkópia(BES) egy széles körben elterjedt fúziós diagnosztikai módszer a plazma sűrűség és annak fluktuációinak mérésére [1]. Alkalmazása során a reaktorplazmába egy ismert anyagi összetételű, keresztmetszetű, nyalábáramú és energiájú semlegesített alkáli vagy hidrogén atomnyalábot lőnek. A nyaláb atomok a plazmába hatolva fokozatosan gerjesztődnek, majd ionizálódnak és a folyamat közben a nyaláb tengelye mentén fotonokat bocsátanak ki. A diagnosztika ezt a foton emissziót egy dedikált megfigyelőrendszerrel vizsgálja és az emisszió térbeli eloszlásából következtet a nyaláb menti plazmasűrűsége.

A RENATE-OD [2] szintetikus diagnosztika segítségével sűrűség profil alapján a nyaláb menti emisszió numerikusan meghatározható. A számítás a rátaegyenletek megoldásával történik, ami a semleges nyalábatomok vegyérték elektronjainak eloszlását határozza meg a releváns atomi szinteken, ahol a figyelembe vett reakciós ráták a nyaláb atom ütközéses fel és le gerjesztése, ionizációja, töltéscsere reakció és a spontán emisszió. Ezen számolások egy infinitezimálisan vékony egydimenziós nyaláb mentén történnek és erőforrás igényesek. Ez erősen korlátozza a szintetikus diagnosztika alkalmazhatóságát kiterjedt három dimenziós nyalábok esetén, ahol több ezer vagy több tízezer ilyen számolás is szükséges.

Ezen dolgozat célja RENATE-OD által végzett számolásokra használható gépi tanulás algoritmusok vizsgálata volt, illetve egy gépi tanulás modell kifejlesztése amivel a sűrűség profil alapján végzett emissziószámítás gyorsabban elvégezhető.

A kutatás során elsősorban Extreme Learning Machine típusú hálókkal foglalkoztam (a PyRCN [3] csomag használatával). A munka elsődleges célja egy működő modell kifejlesztése a sűrűségprofilokhoz tartozó emisszió számítására. Ezt több lépcsőben vizsgáltam, első körben a sűrűség perturbációk elhanyagolásával, majd perturbációk elhanyagolása nélkül.

A modellek tanításához és validálásához több tízezer realisztikus plazma paraméterekkel rendelkező sűrűségprofilt, illetve a hozzájuk RENATE-OD-vel számított emissziós profilt hoztam létre. Emellett a kifejlesztett modelleket teszteltem a

HESEL[4] modell által létrehozott turbulens adatszettekre és valós mérési adatokra. A munka során számos fontos technikai paramétert is vizsgáltam, mint az eredmények folytonossága, illetve a felbontásra való érzékenység.

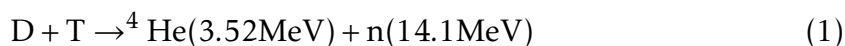
1. Bevezető

1.1. Fúzió

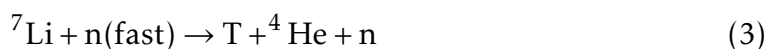
Amikor modern, iparosodott társadalmunk alapvető szükségleteiről beszélünk, nem sok olyan elemi és tagadhatatlan dolog van, mint az energiaszükségletünk. Ennek az igénynek a kielégítésére a tudósok folyamatosan új energiaforrások után kutatnak, és ez a kutatás a nem megújuló energiaforrások fokozatos kivonása érdekében ma aktuálisabb mint valaha.

A magfúzió segítségével villamos energiát termelő reaktor létrehozását már az 1950-es évek óta fontolgatják, [5], de ennek eléréséhez számos nehézséget kell megoldani.

A különböző fúziós folyamatokhoz szükséges feltételek közelebbi vizsgálata alapján a deutérium-trícium fúzió tűnik az egyik legkönnyebben megvalósíthatónak, mivel az összes releváns reakció közül ez rendelkezik a legnagyobb hatáskeresztmetszettel a legalacsonyabb impakt energián. Ezért számos fúziós kísérletben ezt az üzemanyagkeveréket használják. A reakció során két hidrogénizotóp, a deutérium és a trícium atommagok fuzionálnak, és az (1) egyenletben leírt folyamat során egy ${}^4\text{He}$ atomot és egy szabad neutron hoznak létre, miközben $17,62\text{MeV}$ mozgási energia szabadul fel (ami körülbelül négyszerese az urán hasadása során felszabaduló energiának egy tömegegységre vetítve[6]).



A deutérium begyűjtése a jövőbeli reaktorok számára nem jelent majd problémát, mivel az óceánok vizében található hidrogénből minden 6000 atomból körülbelül egy deutérium, ezért könnyen begyűjthető a vízből. A trícium viszont egy radioaktív izotóp, amelynek felezési ideje 12,33 év, és a természetben nem található meg. Megoldásként a termikus (2) és a gyors (3) neutronok lítiumon történő magreakcióit lehet felhasználni a trícium előállítására [7] a fúziós berendezéseken elhelyezett, célzottan épített lítiumtakarók segítségével.



A termonukleáris fúziós folyamatok további vizsgálatából arra lehet következtetni, hogy a gyújtási hőmérséklet 10^4 eV tartományban van, a D-T fúzió hatáskeresztmetszetének ütközési energia függése miatt. Ebben a hőmérséklettartományban az üzemanyagkeverék teljesen ionizált plazmaállapotban van [8].

A fúzióhoz szükséges magas hőmérsékletű plazmák előállítására kétféle megközelítés létezik, az inerciális és a mágneses összetartást. Az inerciális összetartás esetében a szükséges plazma hőmérsékletet a plazmasűrűség extrém növelésével érik el, míg a mágneses összetartás esetében a cél a plazma hosszabb időre történő összetartása, eközben pedig a jó szigetelés és a kiegészítő fűtés biztosítása, amíg a plazma el nem éri a gyújtási pontot.

Most nézzük meg, milyen feltételek (hőmérséklet, energia összetartási idő stb.) szükségesek ahhoz, hogy egy mágnesesen összetartott fúziós berendezésben elérjük a gyújtási pontot, ahol a fúziós égés önfenntartóvá válik. A fúziós hármastermék [9] írja le ezeket a paramétereket. Félig deutérium félig trícium plazma esetén ez az összefüggés az (4) egyenletben leírt alakot veszi fel.

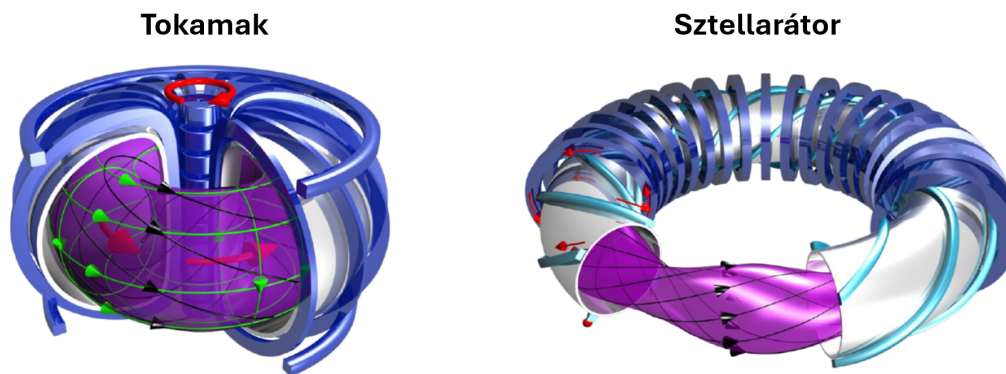
$$n\tau_e T_i \geq 10^{21} \text{keVsm}^{-3} \quad (4)$$

Ahol n a plazmasűrűség, τ_e a energia összetartási idő és T_i a plazma ionpopulációjának hőmérséklete.

1.1.1. Fúziós berendezések

A mágneses összetartás a legelterjedtebb fúziós megközelítés, ahol a forró plazmát toroid alakban zárják be. Ezt úgy érik el, hogy a mágneses erővonalakat spirálisan egy tórusz köré tekerik (fekete vonalak az 1 ábrán). A mágneses erővonalak spirális tekercselését két mágneses térkomponens szuperpozíciójával érik el: egy toroidális térkomponens, amely az eszköz középpontja körül halad, és egy poloidális komponens, amely merőleges az előzőre.

A tokamakok[10] a mágneses összetartású fúzió jelenleg legszélesebb körben alkalmazott eszközei (lásd a 1 bal oldali ábra), míg a sztellarátorok[11] (lásd a 1 jobb oldali ábra) egyre népszerűbbek. A kettő közötti alapvető különbség a mágneses tér létrehozásának módjában rejlik. A tokamakok kvázi 2D mágneses topológiával rendelkeznek, amely forgási szimmetriával rendelkezik, ahol a toroidális mezőt külső tekercsek, a poloidális mezőt pedig egy, a plazmában indukcióval generált erős áram hozza létre. A sztellarátorok ezzel szemben komplex 3D mágneses geometriával rendelkeznek, ahol a teljes mágneses teret csak külső tekercsek hozták létre. Ez a különbség a tokamakot egy impulzusüzemű, instabilitásra hajlamos, de könnyebben tervezhető és érthető eszközzé teszi, míg a sztellarátor egy folyamatos üzemű, stabilabb, de sokkal nehezebben tervezhető és megépíthető eszköz.



1. ábra. Egy tokamak (bal oldalon) és egy sztellarátor (jobb oldalon) belső szerkezete. Láthatjuk, hogy a plazmát(lila) teljesen körülveszik a tekercsek(sötét és világoskék), amelyek zárt mágneses felületeket hoznak létre, és nincsenek a gépekből kilépő erővonalak. (Forrás: R. Wolf "Fusionsforschung: Auf dem Weg zu einer neuen Primärenergiequelle" [12])

2. Plazma diagnosztika

Ahhoz, hogy egy tokamak berendezésben jól szabályozott és tartós reakciót érjünk el, a plazma különböző paramétereinek pontos mérésére szolgáló módszerekre van szükség, hogy jól megértsük a plazma viselkedését, összetételét és teljesítményét. Itt jön a képbe a fúziós diagnosztika. A fúziós eszközökben uralkodó szélsőséges körülmények miatt a mindennapi mérési módszerek itt használhatatlanná válnak, így új specifikus diagnosztikák kifejlesztésére van szükség. A diagnosztikai módszerek két kategóriába sorolhatók:

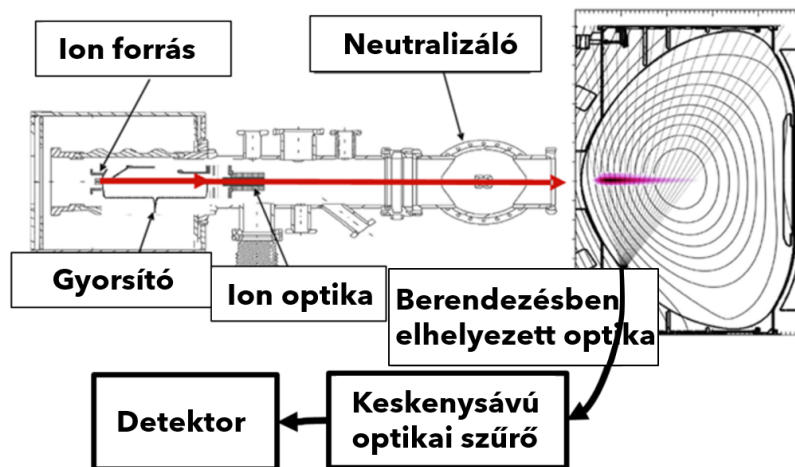
- A passzív rendszerek az elektromágneses tér változásainak mérésével vagy a plazma által kibocsátott részecskék és fotonok gyűjtésével határoznak meg információt, ami általános, de nem jól lokalizált képet ad a plazma paramétereiről.
- Az aktív rendszerek olyan módszereket használnak, amelyekkel lokálisan szondázzák a plazmát abból választ indukálva, és az így kapott elektromágneses vagy részecskeinformációk gyűjtésével nagyon jól lokalizált információkat kapnak.

A Nyalábemissziós spektroszkópia (BES) egy aktív diagnosztikai módszer, amelyet a plazmasűrűség és a sűrűség fluktuációk mérésére használnak, a diagnosztika tér és időfelbontása centiméter illetve mikroszekundum nagyságrendbe esik.

2.1. Nyaláb emissziós spektroszkópia (BES)

Mint fentebb említettük, a BES[13] egy aktív diagnosztikai módszer, amelyet a plazmasűrűség lokalizált mérésére használnak. A mérés vázlata a 2 ábrán látható, a módszer öt lépésre bontható:

1. Először egy ionforrást használnak töltött részecskék előállítására, amelyeket felgyorsítanak és keresztül vezetnek egy semlegesítőn. Ezzel az eljárással fókuszált semleges atomnyalábot állítanak elő, amelynek energiája pontosan beállítható. A mérés többféle nyalábbal végezhető, nyalábanyagként az egy vegyértékelektronnal rendelkező atomok a preferáltak. Ezek közül a lítium a történelmileg legtöbbet alkalmazott.
2. Ezt az atomnyalábot ezután a plazmába injektálják, itt az atomok semlegessége nagyon fontos, mivel egy ionos nyaláb nem tudna átjutni a plazma bezárásához használt nagy mágneses térerősségen.
3. Amikor a nyaláb atomjai elérik a plazmát, a különböző plazmarészecskékkel ütköznek. Ezek az ütközések azt eredményezik, hogy a nyalábatom vegyértékelektronja egy másik atomi szintre gerjesztődik, vagy a vegyérték-elektron közvetlen ionizáció vagy töltéscsere-ütközés révén elveszik és az adott nyalábatom ionizálódik.
4. A gerjesztett vegyérték-elektronok idővel spontán alacsonyabb atomi szintekre lépnek át, és közben fotonokat bocsátanak ki.
5. Ezek után a folyamat során keletkező vonalas emissziót egy független megfigyelőrendszer vizsgálja, amely a nyalábemissziót és a háttéremissziót is összegyűjti, egy a fúziós berendezésen belüli optikai rendszer segítségével. A teljes emissziót egy keskeny sávú optikai szűrőn keresztül vezetik át, hogy csak a megfigyelt atomi átmenetnek megfelelő vonalas sugárzást válasszák ki, amelyet egy detektorrendszerre vetítenek.



2. ábra. A BES-rendszer főbb összetevőit bemutató egyszerűsített ábra. A piros vonal az ionforrás által generált atomnyalábot jelöli, amelyet fókuszálnak és semlegesítenek, mielőtt befecskendezik a készülékbe (a jobb oldalon). A lila vonal a sugár befecskendezése által generált foton-emissziót jelképezi, amelyet az optika fókuszál és a detektorrendszer érzékel, miután áthaladt néhány optikai szűrőn, hogy kiszűrje a nem kívánatos spektrumokat. (Forrás: Asztalos, Ö. (2022). "Modell-aided design and interpretation of beam emission spectroscopy measurements on fusion devices." [14])

A nyaláb mentén történő emisszió a nyaláb mentén a magasabb elektronhéjakon lévő vegyérték-elektron populációból származik. A megfigyelt elektronátmenet felső szintjének vegyérték-elektron populációs sűrűsége, valamint a megfelelő Einstein-együttható ismeretében egy lineáris összefüggés segítségével az adott elektronátmenet fotonemissziója kiszámítható.

A vegyérték-elektronok eloszlása a magasabb elektronhéjakon a 2.1.1 szakaszban leírt ráta egyenletrendszer megoldásából származik, amely explicit sűrűség- és implicit hőmérsékletfüggéssel rendelkezik.

A BES-hez használt különböző nyalábatomok közé tartozik a hidrogén és az alkálifémek, mint például a nátrium (Na) és a lítium (Li).

A hidrogénnyalábos BES-rendszerek[15] detektorait úgy szerelik fel, hogy a fűtőnyalábokat megfigyelhessék, és úgy hangolják, hogy a $H(n = 3 \text{ to } n = 2)$ elektronátmenetet figyeljék meg, amelyet Balmer alfa vonalnak is neveznek. Ezeket a plazma teljes területén használják, ám csak a plazmaingadozás vizsgálatára.

Az alkáli BES-rendszerek Li[16] vagy Na[17] atomokat használnak nyalábatomként, és a megfigyelt elektronátmenetek a Li esetében $2p \rightarrow 2s$, a Na esetében pedig

$3p \rightarrow 3s$. Ez a diagnosztikai rendszer kis plazma behatolással rendelkezik, és alkalmas a határréteg plazma és a szélplazma sűrűségének és fluktuációinak mérésére.

2.1.1. A háttérben lévő fizika, a ráta egyenletek

A ráta egyenletek[2] [18] a plazmába injektált nyalábatomok elektronhéjainak vegyérték-elektron populációját írják le. Az i -dik szint populációját az (5) egyenletben leírtak szerint adjuk meg,

$$\frac{dN_i}{dx} = n_e(x) \cdot C_{ei}(T_e(x), N_i) + \sum_I n_I(x) \cdot C_{Ii}(T_I(x), N_i) + C_{Si}(N_i) \quad (5)$$

ahol N_i a nyalábatom i -dik elektronszintjének elektronpopulációja, $n_e(x)$ a plazma elektronsűrűsége, $C_{ei}(T_e(x))$ az összes nyalábatom elektronokkal való ütközéseket leíró operátor. n_I a plazma I -dik ion-sűrűsége, $C_{Ii}(T_I(x))$ pedig az adott ionnak a nyalábatomokkal való összes ütközéseket leíró operátor. Az egyenlet utolsó része C_{Si} a i -dik elektronszintre vonatkozó spontán emissziós tagok kontribúciója. Az egyenletrendszer megoldása megadja a vegyérték-elektron eloszlást az összes releváns atomi szinten. Ebből a nyaláb emissziója a $n_e(x)$, $n_I(x)$, $T_e(x)$ és $T_i(x)$ függvényében vezethető le, amelyek a plazma elektron és ionsűrűségei valamint a hőmérsékletei az x által jelölt nyaláb mentén, ahol $x = 0$ a nyaláb kezdete.

Az (5) egyenletben szereplő ütközési együtthatók általában a következőképpen írhatók le:

$$\begin{aligned} C_{Zi} = & \frac{-N_i}{v_b} \left(\sum_{j=i+1}^m R_{Z,(i \rightarrow j)}^{ex}(T_Z(x)) + \sum_{j=1}^{i-1} R_{Z,(i \rightarrow j)}^{dex}(T_Z(z)) \right. \\ & \left. + R_{Z,(i \rightarrow +)}^{ion}(T_Z(x)) + R_{Z,(i \rightarrow +)}^{CX}(T_Z(x)) \right) \\ & + \frac{1}{v_b} \left(\sum_{j=1}^{i-1} N_j R_X^{ex}(j \rightarrow i) + \sum_{j=i+1}^m N_j R_X^{dex}(j \rightarrow i) \right) \end{aligned} \quad (6)$$

Ahol C_{Zi} tartalmazza a i -dik elektronszinthez kapcsolódó összes ütközéses nyereség- és veszteség tagot a Z -dik plazmakomponenssel való nyalábatom ütközések tekintetében. A jobb oldalon N_i az i -dik szint populációja, v_b pedig a nyaláb sebessége. Az R tagok a reakciós ráták, amelyek a nyaláb atomjainak reakcióvalószínűségeit írják le egy termikus közegben. Az első négy R tag az i -dik elektronszintről történő elektronvesztéseket írja le, míg az utolsó kettő az alacsonyabb és magasabb elektronszintekről az i -dik elektronszintre történő nyereségeket.

$R_{Z,(i \rightarrow j)}^{ex}(T_Z(x))$ a vegyérték-elektron i -dik elektronszintről egy magasabb j -dik elektronszintre történő ütközéses gerjesztése által okozott elektronvesztés leíró ráta. $R_{Z,(i \rightarrow j)}^{dex}(T_Z(z))$ a vegyérték-elektron i -dik elektronszintről egy alacsonyabb j -dik elektronszintre történő ütközéses legerjesztése által okozott elektronvesztés leíró ráta. $R_{Z,(i \rightarrow +)}^{ion}(T_Z(x))$ a vegyérték-elektron i -dik elektronszintről történő ütközéses ionizációja által okozott elektronvesztés leíró ráta. $R_{Z,(i \rightarrow +)}^{CX}(T_Z(x))$ a plazmaionokkal való töltéscsere-ütközések okozta vegyérték-elektron vesztés leíró ráta. Az előző három reakciótól eltérően az elektronütközéseknél nem lépnek fel töltéscsere-reakciók.

Mint már említettük, az utolsó két komponens a j -dik elektronszintről a i -dik elektronszintre történő elektronnyereségért felelős. $R_X^{ex}(j \rightarrow i)$ a vegyérték-elektron j -dik elektronszintről egy magasabb i -dik elektronszintre történő ütközéses gerjesztése által okozott elektronnyereséget leíró ráta. $R_X^{dex}(j \rightarrow i)$ a vegyérték-elektron j -dik elektronszintről egy alacsonyabb i -dik elektronszintre történő ütközéses de-gerjesztése által okozott elektrongyarapodást leíró ráta.

Ezeket a rátákat úgy származtatjuk, hogy feltételezzük hogy a plazmát alkotó részecskék Maxwelli sebességpopulációval rendelkeznek, míg a nyaláb monoenergetikusnak van feltételezve, ahogy az alábbi egyenletben látható:

$$R = \langle \sigma v \rangle = \int f(T, \mathbf{v}) |\mathbf{v} - \mathbf{v}_b| \sigma(|\mathbf{v} - \mathbf{v}_b|) d\mathbf{v} \quad (7)$$

, ahol σ a hatáskeresztmetszetet, $\mathbf{v}_b$ a nyaláb sebességét, $\mathbf{v}$ pedig a sebességet, valamint $f(T, \mathbf{v})$ a nyalábbal kölcsönhatásba lépő részecsketípusok Maxwelli sebességeloszlását adja.

Mint korábban említettük, az (5) egyenlet utolsó része C_{Si} a spontán emissziós rátát veszi figyelembe, és az 8 egyenletben látható módon fejezhető ki.

$$C_{Si} = \frac{1}{v_b} \left(\sum_{j=i+1}^m N_j A(j \rightarrow i) - N_i \sum_{j=1}^{i-1} A(i \rightarrow j) \right) \quad (8)$$

Ahol v_b a nyalábsebesség az előzőek szerint, N_j a j -dik elektronszint vegyérték-elektron populációja, $A(j \rightarrow i)$ a j -dik elektronszintről történő nyereséget figyelembe vevő Einstein együttható, N_i az i -dik elektronszint vegyérték-elektron populációja, $A(i \rightarrow j)$ a j -dik elektronszintre történő vesztés leíró Einstein együttható.

2.1.2. RENATE-OD

A RENATE-Open Diagnostics a RENATE nyalábemissziós spektroszkópiai szintetikus diagnosztika újonnan átírt Python változata. A RENATE egy teljesen 3D-s nyalábmodellező kód, amely egy 3D-s kiterjesztett nyalábnak a plazmában való nyalábfejlődését modellezi, 3D-s megfigyelési geometriát, és az EFIT-ből [2] származó realisztikus mágneses geometriákat alkalmazva.

A RENATE-OD egy továbbfejlesztett ütközéses sugárzási modellt tartalmaz egy kiterjesztett ráta egyenletrendszerrel.

RENATE-OD jelenlegi változata a ráta egyenleteket klasszikus numerikus módszerekkel oldja meg, ami megbízható és pontos megoldást ad, de erősen számításigényes. A kód ezeket a számításokat egy dimenzióban végzi, és bár így egy számítás lefutása nem vesz igénybe túl sok időt, a fúziós plazmában terjedő 3D-s semleges nyaláb modellezése esetén több tíz- vagy százezer ilyen egydimenziós számításra is szükség lehet.

A diagnosztikai nyalábok átmérője nagyjából 2 – 4 cm, míg a fűtőnyalábok sokkal nagyobbak, 10–20 cm szélességűek és még ennél is magasabbak. Ha a plazmafluktuációk szintjén akarunk modellezni, akkor a térbeli emissziós profilhoz egy 1D nyaláb /mm numerikus felbontásra van szükség. Ez azt jelenti, hogy fűtőnyaláb modellezés esetén egy pillanatfelvételhez 10^4 nagyságrendű számításra van szükség. A szükséges számítások száma tovább nő, ha a diagnosztikával nem csak egy pillanatképet, hanem videót szeretnénk készíteni, mivel több pillanatképből álló sorozat időfeldolgozása lehetővé teszi a turbulens plazmamodellek kísérletekkel való összehasonlítását. Emiatt a numerikus megoldás kiszámításához szükséges számítási teljesítmény erősen korlátozza a RENATE-OD jelenlegi verziójának alkalmazhatóságát.

3. Neurális hálók

A neurális hálózatok olyan gépi tanulási modellek, amelyeket az emberi agy működésének mintája alapján fejlesztettek ki[19]. Ezen hálók egymáshoz kapcsolódó neuronokból állnak, melyek szerkezete viszonylag egyszerű (lásd 3.1. fejezet).

A neurális hálózatok használata napjainkban rohamosan terjed, mivel képesek nagyon hatékonyan megoldani különböző problémákat, ha nagy mennyiségű tiszta képzési adat áll rendelkezésre.

A kivonatban felvázolt probléma egyértelműen meghatározott bemeneti és kimeneti változókkal rendelkezik, a bemenet egy plazmasűrűségből, míg a kimenet egy emissziós profilból áll, illetve ezen paraméterek mindegyike egy kísérletre állandó numerikus felbontással rendelkezik. A be- és kimenet közötti kapcsolatot egy

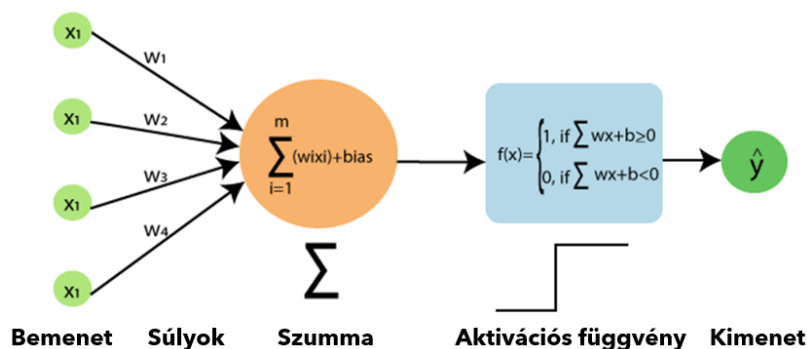
jól definiált differenciálegyenlet-rendszer, a ráta egyenletek és az Einstein-együtthatók írják le. Ennek eredményeképpen a probléma megoldására nagyon jól alkalmazhatók viszonylag egyszerű gépi tanulási módszerek (egyszerűek a ma létező összetettebb modellekhez képest, amikről gyakran olvashatunk cikkeket[20]), mivel a modelleknek fix struktúrája lehet, és csak egy determinisztikus egyenlet megoldását kell megtanulnia. Ráadásul a modell teljesítmény könnyen ellenőrizhető a RENATE-OD által adott pontos megoldásokkal való összehasonlítással.

3.1. A neuron

A neurális hálók megértéséhez először fő alkotóelemüket, a neuront kell megértenünk.

A neuron szerkezete a 3. ábrán látható. A működésének lépései a következők:

1. Zöld körökkel vannak jelölve a neuron x_i bemeneti paraméterei. Ezeket megszorozza a hozzájuk tartozó w_i súlyokkal, majd ezen súlyozott bemeneteket összegzi.
2. Az előbbi pontban kapott összeghez hozzáad egy bias paraméternek nevezett skalárt (ez az ábrán az összegzéssel együtt a narancssárga körben látható).
3. A 2. lépésben kapott összegre haddat egy nemlineáris aktivációs függvényt (az ábrán kékkel jelölt lépcsőfüggvény). Ennek a függvénynek a kimenete lesz a neuron kimenete, $\hat{y}$.



3. ábra. Egy neuron felépítése, amely a bemeneti kapcsolatokon végzett lineáris regresszióból[21] és egy nemlineáris aktivációs függvényből áll. (ebben az esetben lépcső függvény) (Forrás: towardsai.net, "Perceptron: A Basic Neural Network Model for Deep Learning"[22].)

A kékkel jelölt aktivációs függvény igazából bármilyen nemlineáris függvény lehet. Az aktivációs függvény miatt képes a neuron (és végső soron a belőle felépülő

neurális hálók) nemlineáris kapcsolatok leírására.

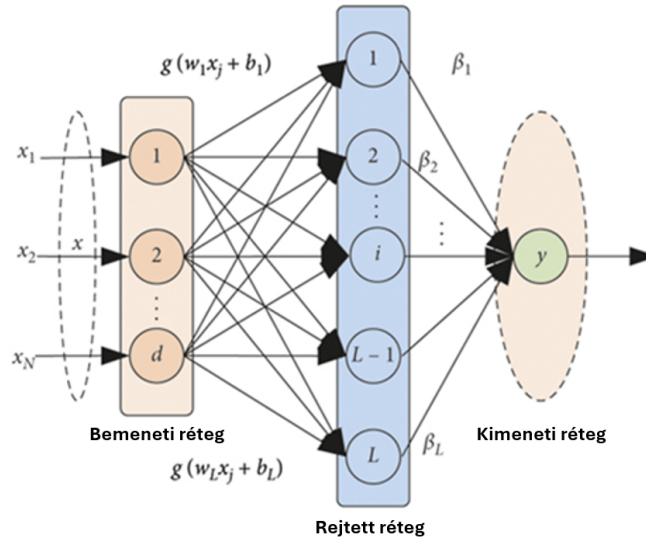
A neuron és a neurális hálók tanítása során a w_i súlyokat, illetve a bias paramétert optimalizáljuk olyan módon, hogy adott x_i bemenet esetén a neuron $\hat{y}$ kimenete minnél közelebb legyen az elvárt kimenethez.

3.2. Perceptron hálók

A perceptron-hálózatok[23] a neurális hálózatok egy altípusát képezik. Szerkezetileg egy bemeneti rétegből, változó számú rejtett rétegből és egy kimeneti rétegből állnak, ahogy az a 4. ábrán látható. A perceptron hálózatokban a rétegek a helyzetük alapján kategorizálhatók. Az első réteg, amely közvetlenül a modell bemeneti paramétereire kapcsolódik, a bemeneti réteg, hasonlóképpen az utolsó réteget kimeneti rétegnek nevezzük. E két réteg különlegessége, hogy a dimenziójukat előre meghatározza az adott probléma, amit a hálózattal meg kívánunk oldani. Ha például 20 bemeneti paraméter van, akkor a bemeneti rétegnek 20 neuronból kell állnia.

A bemenet és a kimenet között vannak a rejtett rétegek, ezek száma és mérete testreszabható, a 4. ábrán egy egyrétegű konfiguráció látható.

A perceptron hálózatok előrecsatolt (az angolban feed forward) hálózatok, ami azt jelenti, hogy a neuronok között nincsenek rekurrens kapcsolatok. Emellett teljesen összekötöttek, ami azt jelenti, hogy a hálózat minden rejtett rétegbeli neuronja kapcsolatban áll az előző és a következő réteg összes neuronjával. A tanítás során ezek a hálózatok egy backpropagation-nek nevezett folyamat segítségével optimalizálják minden neuron súlyait valamint a bias paramétereket [24].

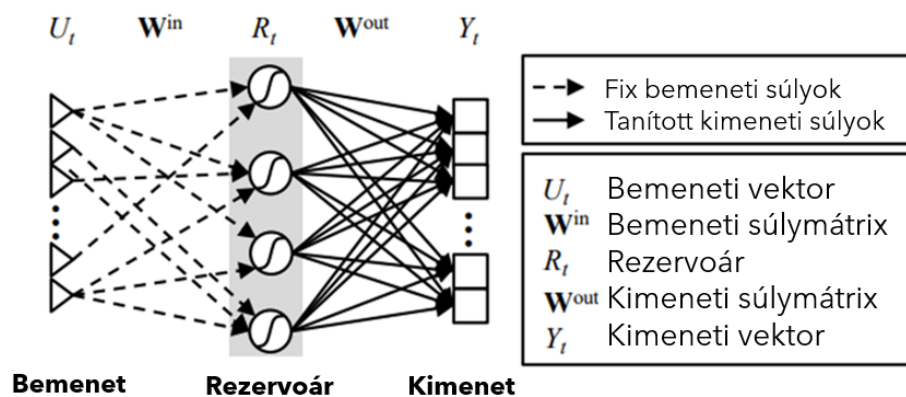


4. ábra. Perceptron hálózat egyetlen rejtett réteggel, N dimenziós bemenettel és 1 dimenziós kimenettel. A rejtett rétegben minden kör egy-egy neuront jelöl, amelyet a 3.1 szakaszban írtunk le. A perceptronnálzatok esetében a neuronok súlyait és a bias paramétereket a tanítás során optimalizálják a backpropagation nevű folyamat során. (Forrás: Zang, Shao-fei & Wang, Xuesong & Yan, Yongyi. (2021). "Transfer Extreme Learning Machine with Output Weight Alignment."[25])

3.3. Extreme Learning machines, avagy ELM-ek

Az Extreme Learning Machine-ek [26][27] vagy röviden ELM-ek a neurális hálózatok egy alcsoportja. A 5. ábrán látható az ELM típusú hálók szerkezete. Az ábrán U_t jelöli a háló bemeneti vektorát, Y_t a kimeneti vektort, R_t pedig az úgynevezett rezervoár, ez a háló neuron rétegeinek együttes neve.

ELM hálók esetén a rezervoárban tetszőleges számú egymás után sorba kapcsolt neuron réteg helyezkedik el. Az egyes neuron rétegek mérete (hogy hány neuront tartalmaznak) szintén tetszőlegesen megszabható, akár csak a neuronok által használt aktivációs függvény.



5. ábra. Egy rejtett réteggel rendelkező ELM illusztrációja. (Forrás: A. Jalvand, O. Asztalos, M. Karacsonyi and G. I. Pokol, "Application of neural networks in beam emission spectroscopy modelling" [28])

Az ELM modellek teljesen összekötött hálók, ami azt jelenti:

1. A rezervoár első rétegének minden neuronja minden bemeneti paraméterrel össze van kötve. Ezekhez az összeköttetésekhez tartoznak a W^{in} -el jelölt súlyok.
2. A rezervoáron belül a szomszédos rétegek minden neuronja össze van kötve egymással.
3. A rezervoár utolsó rétegének minden neuronja minden kimeneti paraméterrel össze van kötve. Ezekhez az összeköttetésekhez tartoznak a W^{out} -al jelölt súlyok.

ELM-ek esetén rekurzív összeköttetések nincsenek.

Az ELM hálók különlegessége, hogy a tanítás során csupán a W^{out} paramétereket optimalizáljuk, a W^{in} súlyok, valamint a rezervoáron belüli súlyok tanítás előtt le vannak rögzítve. A lerögzített súlyok értékeit megadhatja a felhasználó, illetve lehet őket random inicializálni vagy valamilyen eloszlásból mintavételezni, amihez az egyik módszer a batch intrinsic plasticity[29][30][31].

Ez a sajátosság miatt az ELM hálónál nincs szükség a backpropagation-nek[24] nevezett illesztési folyamatra, a modell betanítása gyakorlatilag a kimeneti súlyok illesztése egy egyszerű lineáris regresszióval.

4. A probléma

A 2.1. fejezetben bemutattam egy aktuális problémát a plazmadiagnosztika területén, nevezetesen azt, hogy nyalábemissziós spektroszkópia modellezési folyamatának az a része, amely a plazmasűrűségprofil alapján a nyaláb mentén emissziós profilt generál rendkívül számításigényes. Ez két felhasználási területen is problémás.

1. 3D nyalábmodellezés esetén egyetlen pillanatképhez is 10^4 1D nyaláb kiszámítására lehet szükség, turbulencia vizsgálata során pedig 10^4 – 10^6 pillanatkép szükséges. Az ebből eredő 10^6 – 10^{10} 1D nyaláb modellezése rendkívül időigényes.
2. A valós idejű plazma kontroll felhasználások esetén a BES diagnosztikák eredményeit 1ms alatt 10 – 100-szor kell kiértékelni, amihez a numerikus megoldók túl lassúak.

A 3.3. fejezetben bemutattam két neurális háló típust, a perceptronokat és az ELM-eket, amelyek alkalmazhatók az előbb említett számítás kiváltására, mivel a probléma jól definiált, illetve a bemeneti és kimeneti paramétervektorok dimenziója egy adott BES készülék esetén fix.

A TDK dolgozatom célja egy, a BES sűrűségmérés, modellezés, illetve méréskiértékelés során alkalmazható gyors neurális háló kifejlesztése a sűrűség alapján várható emisszió megadására.

Ennek a létrehozásához több lépésben közelítettem meg a problémát.

1. Első lépésként létrehoztam egy egyszerűsített szintetikus adatkészletet. Ezt felhasználva különböző modell típusokat teszteltem és kiválasztottam a legjobban működő konfigurációt.
2. Ezután növeltem a felhasznált plazmaprofilok komplexitását fluktuációk hozzáadásával, valamint teszteltem a modelleket különböző BES nyalábok esetén.
3. A BES rendszerek mintavételezési frekvenciája egy viszonylag tág spektrumon mozoghat, a viszonylag lassú 100kHz-es mintavételezéstől az 1MHz-es gyors mérésekig. Harmadik lépésként megvizsgáltam a modell teljesítményét egy 2D plazmaturbulencia modell, a HESEL[4] által generált realisztikus, 100kHz-el, illetve 1MHz-el mintavételezett plazmaprofilokra.
4. Miután megbizonyosodtam a háló pontosságáról a 1-3. pontokban említett szintetikus adatkészleteken, megvizsgáltam a teljesítményt egy W7-X-től származó, kísérleti adatok visszaállításából kapott adatkészleten is.

5. Végző lépésként azt vizsgáltam, hogy a kapott hálók mennyire alkalmazhatók valós BES méréseket kiértékelő programcsomagokban.

5a) Először a háló kimenetek folytonosságát vizsgáltam, ami a BES kiértékelő programcsomagok esetén legalább olyan fontos mint a pontosság.

5b) Az 1-4. pontokban a fluktuációk vizsgálatához ideális magas numerikus felbontással dolgoztam. A BES rendszerek saját felbontása általában ennél egy nagyságrenddel kisebb. Ebben a pontban azt vizsgáltam, hogy az általam kifejlesztett hálók hogyan teljesítenek a felbontás csökkentésével.

5. Eredmények

5.1. Adathalmazok

A munkám során több adathalmazzal is dolgoztam melyek között voltak általam létrehozott szintetikus adathalmazok, a HESEL plazmamodellező kód által létrehozott szintetikus adathalmazok, valamint egy kísérleti adatokon alapuló visszaállított adathalmaz a W7-X[32] berendezéstől.

Az összes szintetikus adathalmaz esetén tiszta, kvázi-semleges plazmát feltételeztem, ami azt jelenti, hogy a plazma csak elektronokból és egy ion populációból áll, amelyek ebben az esetben hidrogén ionok. A kvázisemlegesség miatt e két komponens térbeli sűrűségeloszlása megegyezik. Így a plazma sűrűségprofiljai egy függvénnyel megadhatóak.

A BES nyalábfejlődés sokkal kisebb mértékben függ a hőmérsékleti profiltól mint a sűrűségtől, ezért az adathalmazok létrehozásánál a hőmérsékletet nem tekintettem független változónak. A megfelelő plazmasűrűségprofilok átskálázott változatát használtam a fizikailag releváns hőmérséklet bemeneti értékeként.

A gépi tanuláson alapuló nyalábemissziós megoldót több BES-hez használt nyalábra is vizsgáltam. Elsősorban lítium és nátrium nyalábokkal dolgoztam, amelyek a BES-mérésekhez széles körben használt diagnosztikai nyalábok. Lítiumot használnak, illetve használtak például az EAST[33], az ASDEX-Upgrade [34] és a JET [35] tokamakokban. A W7-X sztellarátorban [32] pedig nátriumot. Ezek mellett vizsgáltam a módszer alkalmazhatóságát deutériumból álló fűtőnyalábok esetén is.

5.1.1. Tangens hiperbolikus adathalmazok generálása

Első lépésként létrehoztam három szintetikus adathalmazt a három általam vizsgált nyalábtípushoz, amik a lítium, nátrium és deutérium nyalábok voltak.

A nyalábparamétereket a következők voltak:

- **Lítium nyaláb:** A nyaláb teljes kinetikus energiáját 50 keV-ra állítottam be, ami elterjedt [36]. (Az energia és a hőmérséklet esetében az eV a fúziós plazmafizikában általánosan használt mértékegység $1\text{ eV} = 1.16 \cdot 10^4\text{ K}$).
- **Nátrium nyaláb** A nyaláb teljes kinetikus energiáját 40 keV-ra állítottam be, ami a jelenleg W7-X-en működő BES rendszer által használt energia tartományba esik. A választással az volt a cél, hogy a szintetikus nátrium adatkészlet és a kísérleti W7-X adatkészlet könnyen összevethető legyen.
- **Deutérium nyaláb** Ebben az esetben a nyalábenergia 80 keV-re volt állítva, ami a KSTAR[37] tokamakon működő deutérium nyaláb használati tartományába esik [38].

A nyalábparaméterek kivételével az adathalmazok létrehozása mind a három esetben ugyan azon a módon történt. Mind a hőmérséklet, mind a sűrűség profilt 400 pontban adtam meg, amelyeket egyenlően mintavételeztem a nyaláb egy 0.4 m-es szakasza mentén, ami egy elég tipikus nyalábhossz. Azért választottam ezt a felbontást (10 pont/centiméter), hogy jó szemcsézettséget biztosítsak felesleges túlmintavételezés nélkül (a diagnosztika térbeli felbontása általában egy centiméteres tartományban van).

A plazma sűrűségprofiljának szimulációjához tangenshiperbolikus függvényt használtam, amely jó közelítés a stacionárius sűrűségprofilokhoz [39]. Egy példa profil látható az 6/a ábrán, itt az alacsonyabb értékek a határréteg plazmát képezik, ahol a mágneses mezővonalak nyíltak, míg a magasabb értékek a zárt fluxusfelületű, zárt plazma régiót fedik le. A tanh függvény inflexiós pontja az utolsó zárt fluxusfelület, a zárt és nem zárt plazma közötti határvonal helyzete. A profil előállításához a (9). egyenletet használtam.

$$\rho(x) = \rho_{min} + \frac{\rho_{max} - \rho_{min}}{2} \cdot [1 + \tanh(g \cdot (x - x_0))] \quad (9)$$

Ahol a következő paraméterek változtak:

- ρ_{min} a határréteg plazma sűrűség (a $\rho(x)$ függvény alsó határértéke).
- ρ_{max} a szélplazma sűrűség (a $\rho(x)$ felső határértéke).
- g a sűrűségprofil meredeksége az inflexiós pontban.
- x_0 az utolsó zárt fluxusfelület helyzete

Ezeket a paramétereket rögzített lépésekben változtattuk az 2. táblázatban megadott tartományokban. Ez 6 lehetséges paramétert értéket eredményezett a határréteg plazma sűrűségére és 12-t az szélplazma sűrűségére. A profil meredekségének 10 lehetséges értéke volt, míg az utolsó zárt fluxusfelület helyzetének 11. A felsorolt paraméterbeállítások minden lehetséges kombinációját kipróbáltam, és ennek eredményeképpen 7920 egyedi plazmasűrűség-profilt kaptam.

Paraméter	Minimum érték	Maximum érték	Lépés méret
Határréteg plazma sűrűség ρ_{min}	10^{17} m^{-3}	10^{18} m^{-3}	$2 \cdot 10^{17} \text{ m}^{-3}$
szélplazma sűrűség ρ_{max}	10^{18} m^{-3}	10^{20} m^{-3}	$2 \cdot 10^{17} \text{ m}^{-3}$
A profil meredeksége g	15 m^{-4}	60 m^{-4}	5 m^{-4}
Az utolsó zárt fluxusfelület helyzete x_0	0.05 m	0.15 m	0.01 m

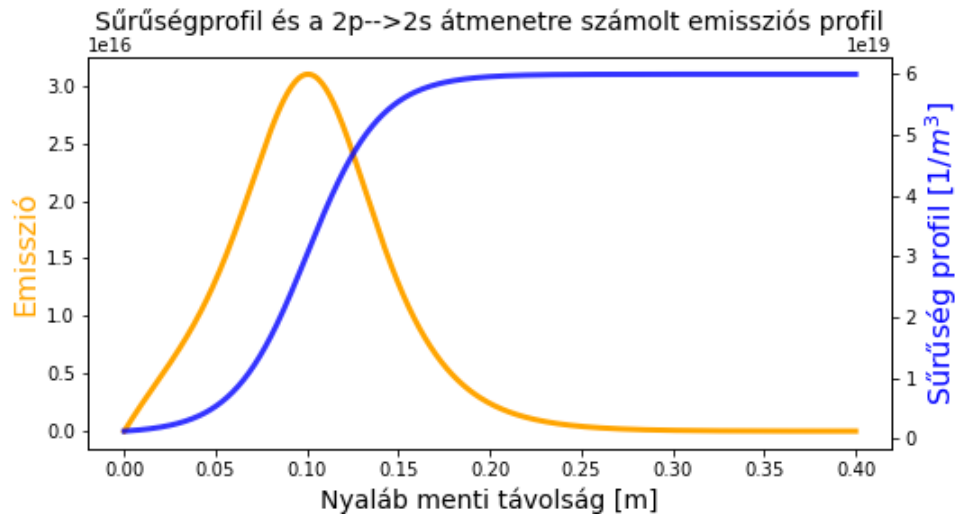
1. táblázat. Az 1. adathalmaz létrehozásához használt paramétertartomány. Ezeket a paraméterértékeket úgy választottam ki, hogy reális és fizikailag releváns eredményeket kapjak.

Mint korábban említettem, az egyenlet által generált profilokat vettem, és a függvény alsó és felső határértékeinek újraskálázásával létrehoztam belőlük a megfelelő hőmérsékleti profilokat. Ezeket az értékeket minden esetben 1 eV és 3 keV értékre állítottam be, ami reális plazma hőmérsékletnek felel meg.

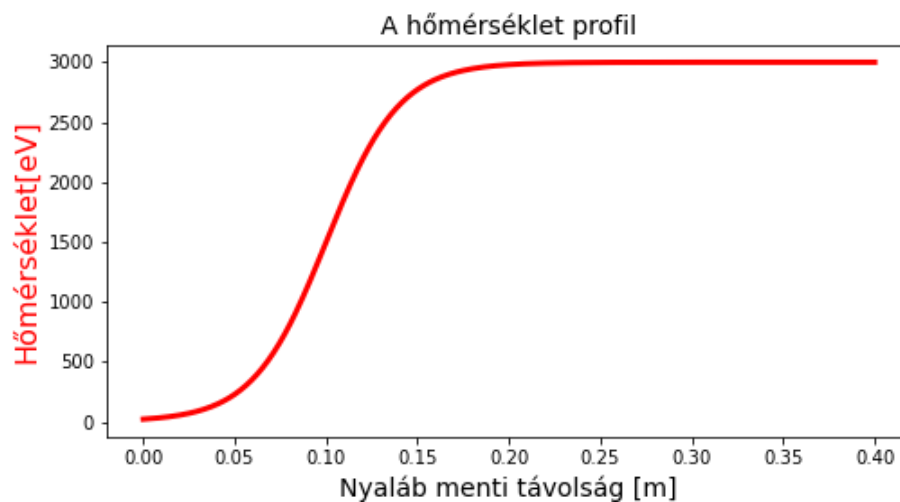
Ezután a kapott hőmérséklet- és plazmaprofilokat (kék és piros profilok a 6. ábrán) használva a RENATE-OD segítségével kiszámítottam a megfelelő emissziós profilokat. Mint a 2.1. fejezetben említettem Lítium nyaláb esetén a $2p \rightarrow 2s$ átmenet által generált emissziót használtam (sárga profil a 6 ábrán), nátrium esetén a $3p \rightarrow 3s$ -t, deutérium esetén pedig a hidrogénnél is használt Balmer alfa vonalat ($3n \rightarrow 2n$).

A 6 ábrán látható egy példa a generált plazma- és hőmérsékleti profilra a számított emissziós profillal. Mivel a hőmérsékleti profilt ebben a munkában nem kezeltem független bemenetként, a későbbi ábrákból kihagytam, hogy átláthatóbbá tegyem azokat.

Mint már említettem, a bemeneti paraméterek minden lehetséges kombinációjának felhasználásával 7920 profilt hoztam létre az egyes nyalábokhoz.



(a) Egy példa a sűrűségprofilok közül (kék) és a hozzá számított emissziós profil (sárga) a $2p \rightarrow 2s$ átmenethez, lítium nyaláb esetén.



(b) A fenti sűrűséghez tartozó hőmérséklet profil

6. ábra. Egy példa profil az 1. (statikus) adathalmazból lítium nyaláb esetén.

5.1.2. Fluktuációs tangens hiperbolikus adathalmazok generálása

Két fluktuációkkal ellátott tangens hiperbolikus profil alapú adathalmazt generáltam a 5.1.1. fejezetben leírtakat felhasználva, a statikus profilokra fluktuációkat helyezve. A fluktuációs adathalmazokat a lítium és a nátrium nyalábhoz hoztam létre. A profilok alapja (a tanh függvény) és a hozzá használt paraméterek teljes mértékben azonosak az előző alfejezetben leírtakkal. Ezekre az alapokra minden esetben pontosan egy Hann ablak alakú fluktuációt helyeztem. A fluktuációkat három paraméterrel lehet jellemezni, x_f a fluktuációcsúcs pozíciója a nyaláb mentén (méterben), w_f a fluktuáció szélessége (méterben), amp_f pedig a fluktuáció amplitúdója (ez az alap profil sűrűségéhez hozzáadódik az x_f pontban és így kapjuk a végső sűrűség értéket).

A fluktuációk értékét (F_{fluc}) a 10. egyenlet adja meg az x_f körüli w_f tartományon, azon kívül pedig 0 lesz.

$$F_{fluc}(x) = \text{amp}_f \cdot \left(0.5 - 0.5 \cdot \cos \left(2 \cdot \pi \left(\frac{x - \left(x_f - \frac{w_f}{2} \right)}{w_f} \right) \right) \right) \quad (10)$$

Paraméter	Minimum érték	Maximum érték	Lépés méret
Fluktuációcsúcs pozíció x_f	0.01 m	0.2 m	0.01 m
Fluktuáció szélesség w_f	0.005 m	0.001 m	0.02 m
Fluktuáció amplitúdó amp_f	10^{18} m^{-3}	10^{19} m^{-3}	$1 \cdot 10^{18} \text{ m}^{-3}$

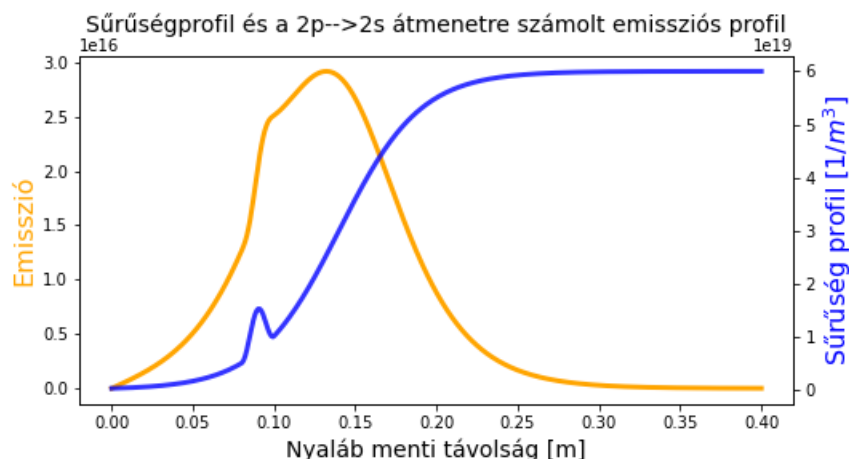
2. táblázat. A 2. adathalmaz létrehozásához használt fluktuációs paraméterek. Ezeket a paraméterértékeket úgy választottuk ki, hogy reális és fizikailag releváns eredményeket kapjunk.

A fluktuációs adathalmazok generálásánál már sokkal több kombináció lenne lehetséges az új paraméterek miatt, mindkét nyaláb esetén 60000 profilt generáltam, amelyekhez a paramétereket egy pseudo random szám generátor segítségével választottam ki.

Ezek az adathalmazok sok extrém esetet is tartalmaznak, ahol többek közt például "Edge Localized Mode"-nak [42] nevezett plazmakilökődéssel járó események közben látható plazmaprofilok, vagy extrém nagy fluktuációk is vannak. Ezen extrém esetek a BES mérések szempontjából kevésbé relevánsak, ám remekül használhatóak

a neurális hálók viselkedésének vizsgálatára.

Egy (nem az extrém esetek közé tartozó) példa a generált fluktuációs sűrűségprofilok közül a 7. ábrán látható a hozzá tartozó hőmérséklettel, valamint az ezekből számított emisszióval.

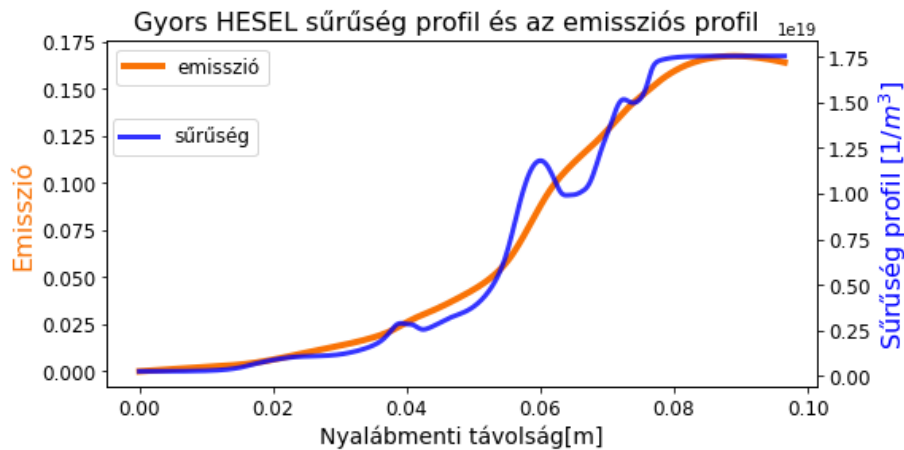


7. ábra. Egy példa a fluktuációs sűrűségprofilok közül (kék) és a hozzá számított emissziós profil (sárga) a $2p \rightarrow 2s$ átmenethez, lítium nyaláb esetén.

5.1.3. HESEL adathalmazok

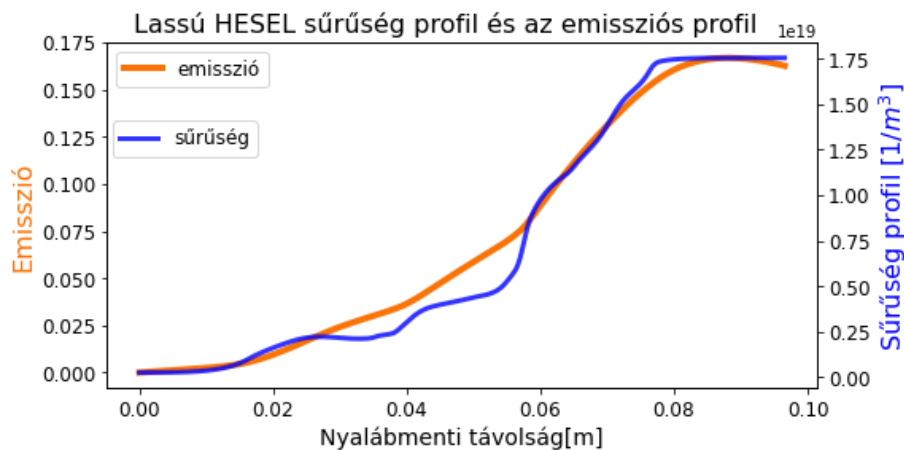
Mint említettem a munkám során dolgoztam a HESEL kód által létrehozott komplex fluktuációs profilokkal is, pontosabban az alább leírt 2 adatkészlettel. Mindkét esetben lítium nyalábot használt a HESEL kód a modellezéshez. A HESEL adathalmazokat (a tangens hiperbolikus adatkészletekkel ellentétben) nem én hoztam létre, csupán feldolgoztam és felhasználtam őket.

Gyors HESEL adathalmaz A gyors HESEL adathalmaz 32014 profilból állt 384 pontos numerikus felbontással, amik 10cm hosszan helyezkedtek el a nyaláb mentén. A profilok 1 MHz-en mintavételezett határréteg plazma sűrűség profilok voltak a hozzájuk tartozó emissziós profillal. Ezen az időfelbontáson meglehetősen sok turbulencia látható a profilokon, mint azt a 8. ábra is szemlélteti.



8. ábra. Egy példa az 1MHz-el mintavételezett gyors HESEL sűrűségprofilok közül (kék) és a hozzá tartozó emissziós profil (sárga) a $2p \rightarrow 2s$ átmenethez (Lítium BES rendszer).

Lassú HESEL adathalmaz A lassú HESEL adathalmaz szintén 32014 profilból állt 384 pontos numerikus felbontással, amik 10cm hosszan helyezkedtek el a nyaláb mentén. Ebben az esetben az 1MHz-en mintavételezett sűrűségprofilok ki voltak integrálva 0,5ms-os intervallumokra. Így a lassú HESEL adatkészlet simább volt, inkább lassabb BES mérésekre emlékeztető profilokkal. Ezt a 9. ábra is szemlélteti.



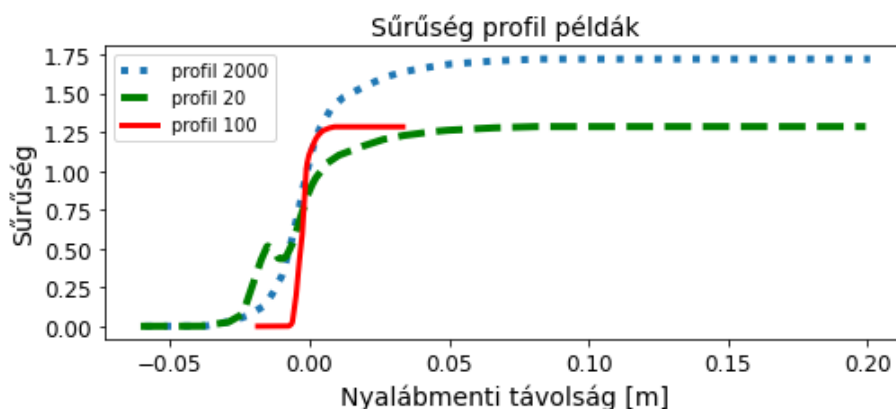
9. ábra. Egy példa a 0.5ms-ra integrált lassú HESEL sűrűségprofilok közül (kék) és a hozzá tartozó emissziós profil (sárga) a $2p \rightarrow 2s$ átmenethez (Lítium BES rendszer).

5.1.4. W7-X adathalmaz

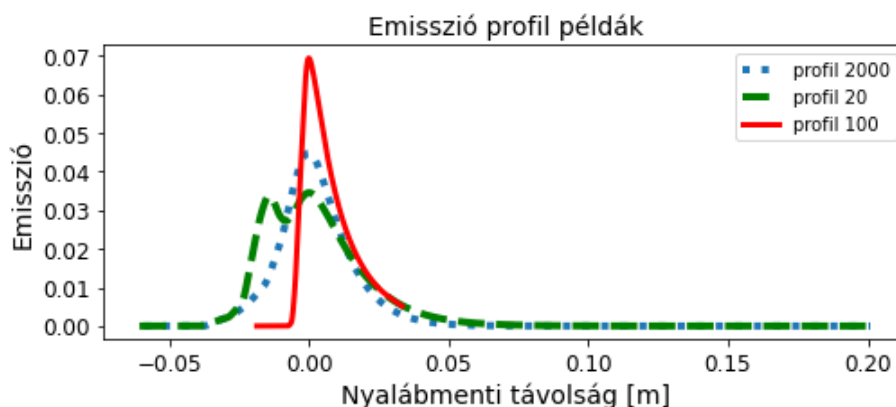
A W7-X berendezéstől kapott kísérleti méréseken alapuló rekonstruált adatkészlet amivel dolgoztam 16046 profilból állt. A BES rendszer ebben az esetben nátrium nyalábot használt.

Ennek az adatkészletnek a sajátossága, hogy bár minden profil 1000 pontból állt, ezek elhelyezkedése a nyaláb mentén és a pontok közötti távolság profilonként eltérő volt a sűrűség profil tulajdonságaitól függően. A nyalábmenti távolság 0 pontjának az utolsó zárt fluxusfelület volt kijelölve. Ezek jól láthatóak a 10.-11. ábrákon.

Mind a sűrűség, mind az emissziós profilok normálva voltak ebben az adatkészletben. Ez, mint később látjuk majd, az újratanított modellnek nem jelent különösebb problémát, így a normált adatot nem alakítottam vissza.



10. ábra. Három sűrűség profil a W7-X adatszettből. Az ábrán jól látható, hogy a profilokhoz tartozó térbeli mintavételezés eltér, viszont az utolsó zárt fluxusfelület minden esetben a 0 pontba esik.



11. ábra. A 10. ábrán látható sűrűség profilokhoz tartozó három emisszió profil a W7-X adatszettből.

5.2. Adattisztítás

5.2.1. Szintetikus adathalmazok

A 5.1.1.-5.1.3. fejezetekben bemutatott szintetikusan létrehozott adathalmazok esetében nem volt szükség adattisztításra, és az adatelőkészítés is meglehetősen egyszerű volt.

A megfelelő sűrűség és emissziós profilokat a python sklearn[40] könyvtár "train_test_split" függvényét használva 60%-20%-20%-os felosztással tanító, validációs és teszt halmazokra osztottam. Ez a függvény pszeudo-véletlen felosztást biztosít, ahol a random seed rögzítésével reprodukálható eredményeket kaptam.

5.2.2. W7-X adathalmaz

A W7-X adathalmaz esetében a változó felbontás miatt több módszer is felmerülhet a profilok standardizálására. Én a profilok extra- illetve interpolálása mellett döntöttem. Ezt az alábbi módon elvégeztem.

Minden sűrűség és emissziós profilt az adatbázisban előforduló leghosszabb térbeli szakaszra extrapoláltam -0.06m -tól 0.223m -ig. Ebben a lépésben egy interpolálást is elvégeztem, mivel az összes profil numerikus felbontását 400 pontra csökkentettem (mivel ennyit használtunk a tangens hiperbolikus szerű adatkészletek esetén is, és ez a felbontás kellően magas is hogy leírja a profilokon látható fluktuációkat).

A következő inter és extrapolációkat alkalmaztam:

- **Sűrűség profilok** esetén lineáris interpolációt alkalmaztam és konstans extrapolációt. Vagyis a profil első, illetve utolsó értékét vettem a kiterjesztett tartományon, ahol szükséges volt. Ez az extrapolációs módszer illik legjobban a sűrűség profilokhoz, amik az eredetileg megadott tartomány szélén bekonvergálnak egy-egy értékhez (hiszen eredetileg a tartomány az alapján volt kiválasztva, hogy mekkorák a gradiensek a nyaláb mentén).
- **Emisszió profilok** esetén lineáris interpolációt alkalmaztam és lineáris extrapolációt. Tehát a profil első/utolsó két pontjára illesztett egyenest használtam a kiterjesztett tartományon, ahol szükséges volt. Ez az extrapolációs módszer jobban illik az emisszió profilokhoz, mivel a sűrűséggel ellentétben az emisszióknak mindenképpen le kell csengenie a nyalábgyengülés miatt.

Az emisszió profiloknál alkalmazott lineáris extrapoláció több esetben negatív emissziós értékekhez is vezetett az extrapolált tartományon. Ezeket 0-ra állítottam, hiszen a negatív emisszió nem bír fizikai értelemmel, csupán az extrapolálás mellékterméke.

Végző lépéseként az adatokat 60%-20%-20%-os felosztással tanító, validációs és teszt halmazokra osztottam.

5.3. Teljesítmény metrikák

A modellek végző teljesítményét elsősorban a tesztkészlet predikcióihoz kapott relatív kumulatív hibák átlagának kiszámításával értékeltem. Ezt a hibát a (11) egyenletben leírtak szerint számoltam ki.

$$Error = \frac{1}{N} \cdot \sum_{i=0}^N \left(\frac{\sum_{j=0}^M |y_{i,j} - p_{i,j}|}{\sum_{j=0}^M y_{i,j}} \right)_i \quad (11)$$

Ahol N a profilok száma a teszt adatszettben (ez minden esetben a teljes adatszett 20%-a volt), M az egyes profilok pontjainak száma (ez a numerikus felbontás, a legtöbb esetben 400 volt), $y_{i,j}$ a $i^{i,j}$ profil j^{th} pontja és $p_{i,j}$ a $i^{i,j}$ profilra vonatkozó előrejelzés j^{th} pontja.

5.4. Modellek tesztelése statikus profilokra

Első lépésként a lítium nyalábhoz létrehozott tangens hiperbolikus adatkészletet használtam a megfelelő modell architektúra kiválasztásához.

Erre az adatkészletre több perceptron és ELM hálót illesztettem és optimalizáltam,

valamint egy lineáris regressziót is illesztettem összehasonlítási alapként.

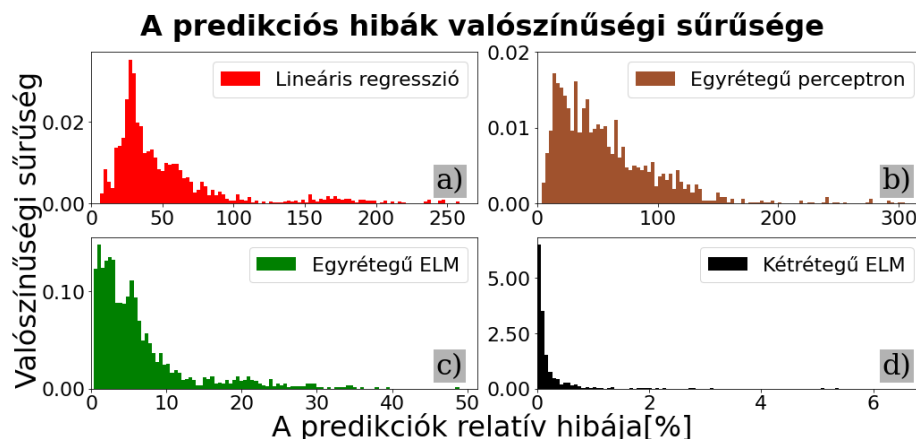
A 3. táblázatban a legjobban teljesítő egy- illetve kétrétegű perceptron és ELM háló, illetve a lineáris regresszió eredményei láthatóak. A 12. ábra pedig az ezekhez a modellekhez tartozó hiba histogrammokat tartalmazza (a kétrétegű perceptron kivételével, ami a legrosszabbul teljesítő modell volt).

Modell	Átlagos Abszolút Hiba $\left[\frac{\text{foton}}{\text{s}\cdot\text{m}}\right]$	Átlagos Relatív Kumulatív Hiba	Predikciós idő 1000 profilra	Teljesítmény RENATE-hoz képest
RENATE-OD (kiindulási alap)	-	-	1448.14 s	-
Lineáris regresszió	$3.79 \cdot 10^{15}$	52.46%	0.02 s	72407-szer gyorsabb
Egyrétegű Perceptron	$4.18 \cdot 10^{15}$	56.12%	0.19 s	7622-ször gyorsabb
Kétrétegű Perceptron	$5.69 \cdot 10^{15}$	77.44%	0.17 s	8518-szor gyorsabb
Egyrétegű ELM	$4.52 \cdot 10^{14}$	6.43%	0.10 s	14481-szer gyorsabb
Kétrétegű ELM	$1.61 \cdot 10^{13}$	0.27%	0.17 s	8518-szor gyorsabb

3. táblázat. A kifejlesztett modellek pontosságának és teljesítményének összehasonlítása a RENATE-számítással. A kétrétegű ELM-háló alkalmazásával a számítás 8518-szor gyorsabban elvégezhető, miközben a pontosság csupán 0,27%-al romlik.

A lineáris regresszióval kapcsolatban kezdettől fogva alacsony elvárásaim voltak, mivel várható volt, hogy a bemenet és a kimenet közötti kapcsolat nem írható le pusztán lineáris összefüggésekkel. Ez a sejtés az 52.46%-os hibával be is bizonyosodott.

Ami meglepő eredmény volt, hogy a perceptron hálózatok a lineáris regressziónál rosszabbul teljesítettek, míg az ELM hálózatok jól működtek. A perceptron hálózatok elméletileg képesek pontosan ugyanazokat a súlyokat megtanulni a képzés során amelyeket az ELM hálózatok alkalmaznak, mivel a két modell szerkezetileg hasonló. Ennek ellenére a probléma a perceptronok számára túl nehéznek bizonyult.



12. ábra. A relatív előrejelzési hibák valószínűségi sűrűsége a tesztkészleten a különböző modellekkel. A legrosszabbul teljesítő modell (a kétrétegű perceptron) nem látható ezen az ábrán.

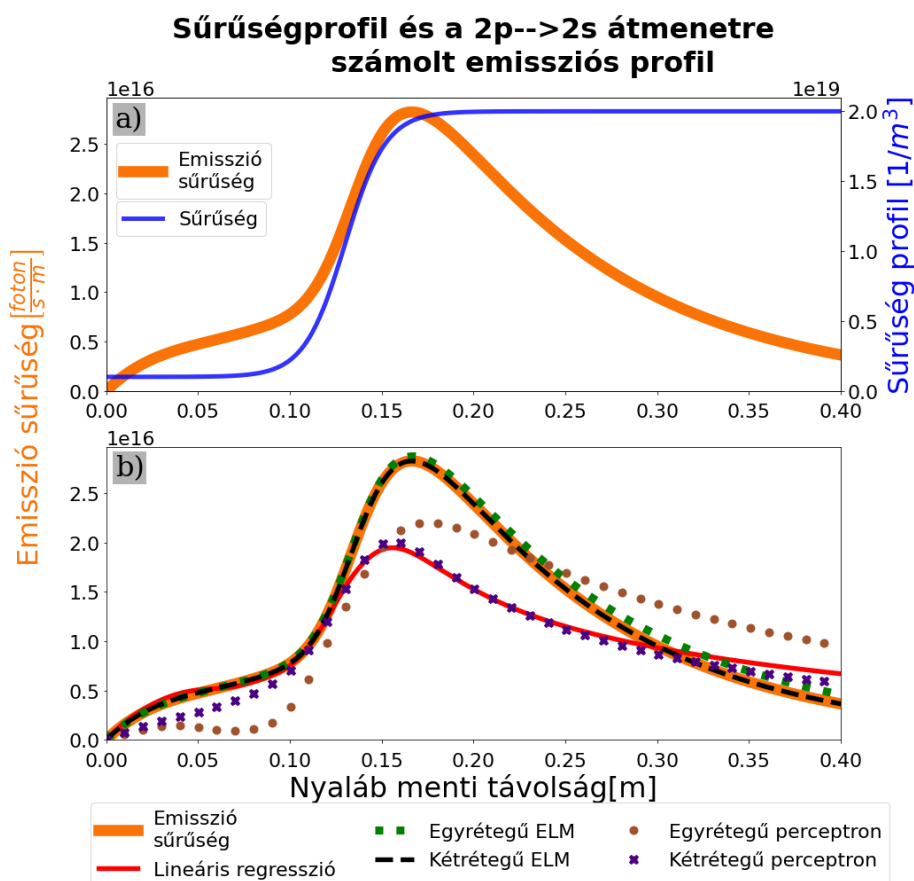
Valószínűleg túl sok helyi minimum van a többdimenziós veszteségfelületen ahol a perceptronhálózatok a képzés során veszteségminimalizálást végeznek, ezért ezek a modellek megrekedtek ezekben a helyi minimumokban, és nem tudták elérni a helyes paraméterbeállításokat. Az ELM-hálózatok robusztussága azonban hasznosnak bizonyult, ezek a modellek nem ütköztek ugyanabba a problémába, mint a perceptronok, mivel ezek nem próbáltak minden paramétert backpropagation-el optimalizálni. Az ELM-hálózatok sikerét az is lehetővé tette, hogy a rögzített súly és bias paramétereket sikerült egy, a problémához illeszkedő eloszlásból mintavételezni a batch intrinsic plasticity segítségével. Így történhetett, hogy a kétrétegű ELM háló képes volt az emissziós profilok prediktálására, csupán 0.27%-os hibával.

A 12. ábra azt mutatja, hogy ezek a megállapítások szinte minden tesztprofil esetében helyt állnak. A lineáris regresszió és az egyrétegű perceptron modellek általában nagy hibákkal rendelkeznek, bár az e modellek által adott hibák közötti szórások is nagyok. Ezzel szemben az ELM-hálózatok általában jól teljesítenek, legfőképpen a kétrétegű. Ennek a modellnek nemcsak az átlagos hibája nagyon alacsony, hanem a hibák szórása is.

A modellek 1000 profilra vonatkozó predikciós idejét szintén a 3. táblázat tartalmazza. Itt láthatjuk, hogy bár a lineáris regresszió által adott előrejelzési idő a legalacsonyabb (amely messze a legegyszerűbb modell) minden modell jelentősen gyorsabb mint a RENATE-OD, amelynek 1000 profilra vonatkozó emisszió számítás

1448.14 másodpercet vesz igénybe.

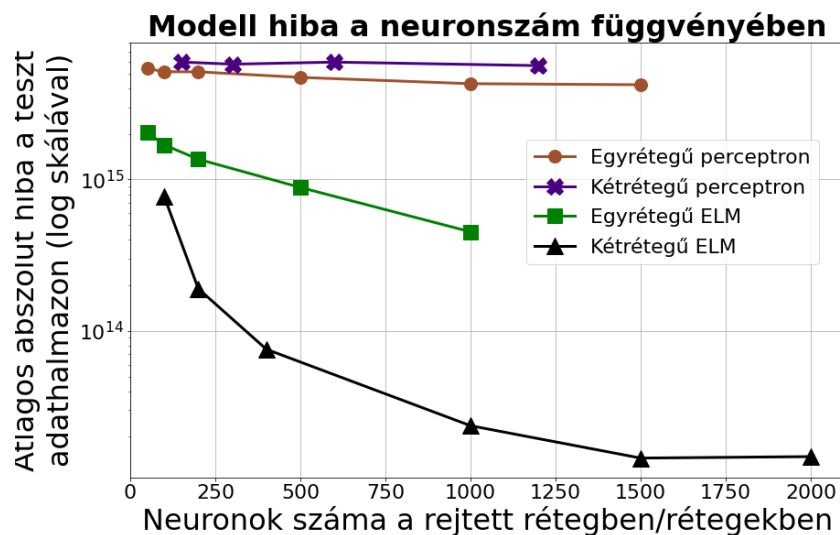
A 13/a ábrán egy példa sűrűség profil és a hozzá tartozó emisszió látható, a 13/b ábra ugyanezt az emisszió profilt és a különböző modellek által adott predikciókat tartalmazza. Az ELM hálózatok pontossága és a többi modell pontatlansága itt is jól látható.



13. ábra. Egy példa profil a lítium nyalábos tangens hiperbolikus tesztkészletből és a különböző modellek által megjósolt profilok. Egyértelmű, hogy az ELM-hálózatok jól teljesítenek ebben az esetben, míg a többi modell nem tudta helyesen megjósolni az emisszió profilt.

A 14. ábrán a neurális hálózatok előrejelzési hibája a rejtett rétegek neuron-számának függvényeként van ábrázolva, ami a modell komplexitásának mértékét írja le. Ez az ábra világosan mutatja, hogy míg a perceptron hálózatok nem igazán

profitálnak a komplexitás növeléséből, addig az ELM hálózatok teljesítménye jelentősen nő a neuronok számának növelésével. Az optimális teljesítményt a kétrétegű ELM 1500 neuron körül éri el, nincs értelme tovább növelni a neuronok számát, nem növeli a teljesítményt.



14. ábra. A modell teljesítménye a rejtett réteg neuronjainak számától függően. Ez az ábra azt mutatja, hogy a modell teljesítménye hogyan javul a kipróbált neuronhálózatok összetettségének növelésével.

Ezeket az eredményeket vizsgálva egyértelmű, hogy a kétrétegű ELM hálózat nagymértékben felülmúlja a többi kipróbált modellt: átlagosan csak 0,27% relatív hibát okoz, miközben a teljesítménye még mindig 8518-szor felülmúlja a RENATE-OD által végzett számításét. Ebből az eredményből kiindulva a munkám következő fejezeteiben kétrétegű ELM modelleket fogok alkalmazni.

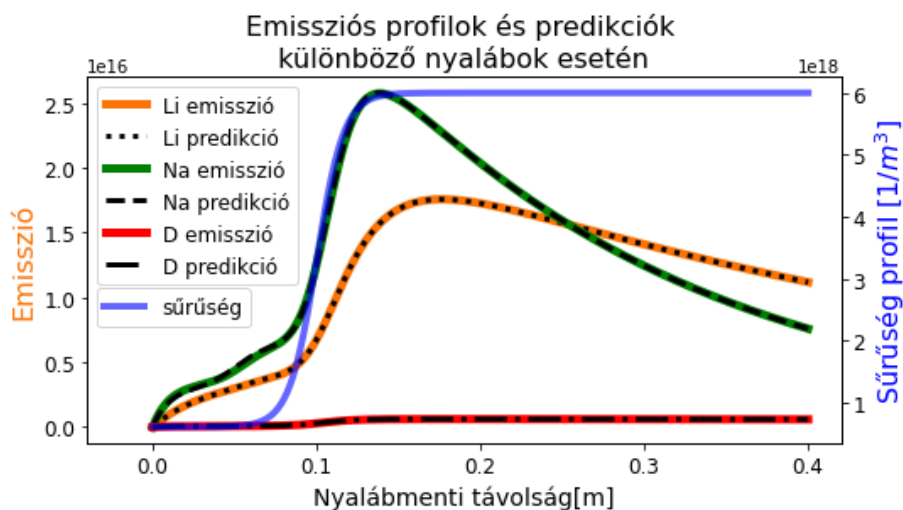
5.5. ELM-ek teljesítménye a különböző adatszetekre

A munkám következő részében a 5.4. fejezet eredményeiből kiindulva végig kétrétegű ELM hálókat alkalmaztam, amiket tanítással és hiperparaméter optimalizálással illesztettem a különböző adatkészletekre.

5.5.1. Modell eredmények a tangens hiperbolikus adathalmazokon

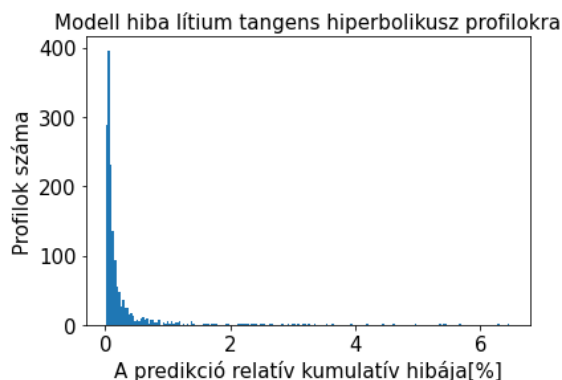
A 5.4. fejezetben már bemutattam, hogy lítium nyaláb BES és tangens hiperbolikus sűrűségprofilok esetén az ELM hálók rendkívül gyorsan és pontosan képesek a sűrűség profilok alapján várható emisszió meghatározására. Ezen eredményekből kiindulva a modellek alkalmazását kiterjesztettem a deutérium és nátrium nyalábokra is.

A három különböző nyalábtípusal létrehozott adatkészletre három különböző ELM modellt optimalizáltam és tanítottam. Az 15. ábrán egy példa sűrűség profil látható a hozzá tartozó emissziós profilokkal a különböző nyalábok esetén, valamint a három ELM modell által prediktált emissziós profilok.

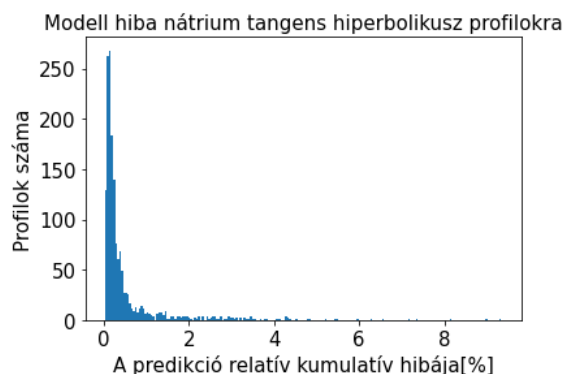


15. ábra. Egy példa sűrűség profil és a hozzá tartozó emissziós profilok a három különböző nyaláb esetén.

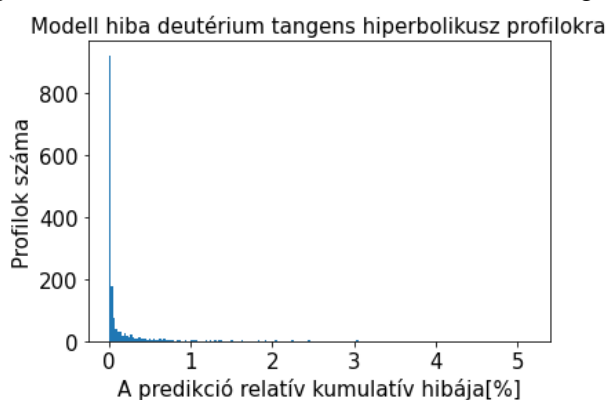
A kiragadott példán jól látszik, hogy a modellek által adott predikciók tökéletesen illeszkednek a RENATE-OD által számolt emissziós profilokra. Hogy meg tudjuk vizsgálni az ELM-ek teljesítményét az egész adathalmazon a 16. ábrán ábrázoltam a modellekhez tartozó hiba hisztogramokat.



(a) Modell hiba elszolás a lítium nyalábos adatkészleten. Az átlagos hiba 0.2631%.



(b) Modell hiba elszolás a nátrium nyalábos adatkészleten. Az átlagos hiba 0.4821%.



(c) Modell hiba elszolás a deutérium nyalábos adatkészleten. Az átlagos hiba 0.1740%.

16. ábra. Modellhiba a tangens hiperbolikus alakú sűrűségprofilokon a három különböző anyagú nyaláb esetén.

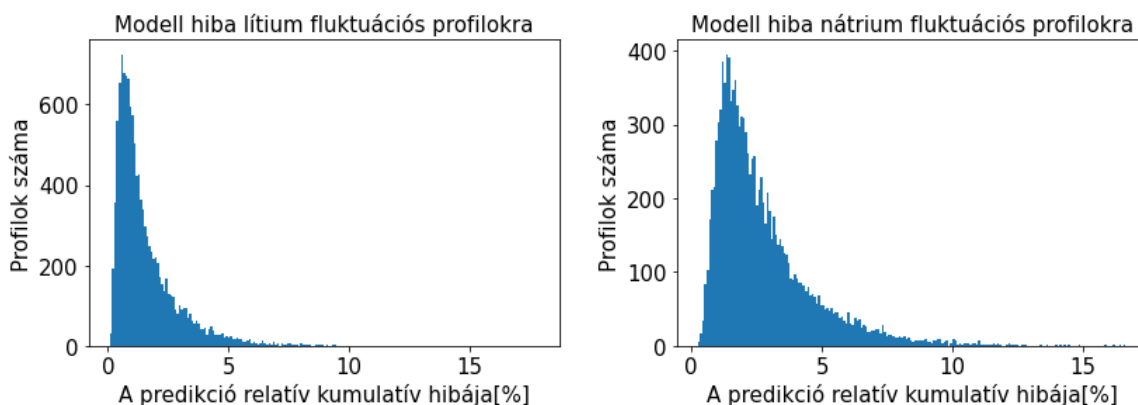
Az ábrák alapján jól látszik, hogy mind a nátrium mind a deutérium esetében a modell pontossága összemérhető a korábban már vizsgált lítium nyalábos esetével. Fontos emellett azt is megjegyezni, hogy a modell hiba a legrosszabb esetekben is csupán 6,8, illetve 5 százalék körüli a különböző nyalábok esetén, bár a profilok túlnyomó többségénél mind a három esetben jóval kisebb hibát vét a modell. Összehasonlításképpen a rátaegyenletek megoldására használt numerikus ütközési sugárzási modellek eredményeit nagyságrendileg 10%-os bizonytalanság terheli [41].

Ezek alapján kijelenthető, hogy az ELM hálók jól általánosíthatóak különböző anyagú BES nyalábokra, mint a diagnosztikai nyalábként elterjedt lítium és nátrium

nyalábok, vagy a deutérium fűtőnyalábok. A modellek hibája pedig kellően alacsony ahhoz, hogy a valóságban is felhasználhatóak legyenek.

5.5.2. Modell eredmények a fluktuációs tangens hiperbolikus adathalmazokon

A következő lépésben a két fluktuációkkal ellátott tangens hiperbolikus adatkészletre illesztettem és optimalizáltam ELM hálókat. A kettő különböző nyalábtípushoz tartozó modellek hiba hisztogramjai a 17. ábrán láthatók.



(a) Modell hiba elszólás a lítium nyalábos fluktuációs tangens hiperbolikus adatkészleten. Az átlagos hiba 1.5600%.

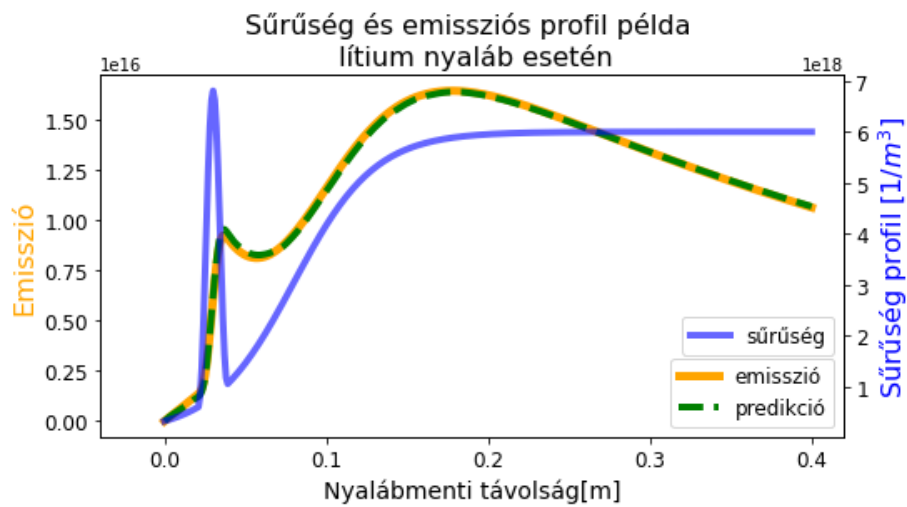
(b) Modell hiba elszólás a nátrium nyalábos fluktuációs tangens hiperbolikus adatkészleten. Az átlagos hiba 2.7430%.

17. ábra. Modell hiba a fluktuációs tangens hiperbolikus adatkészleteken.

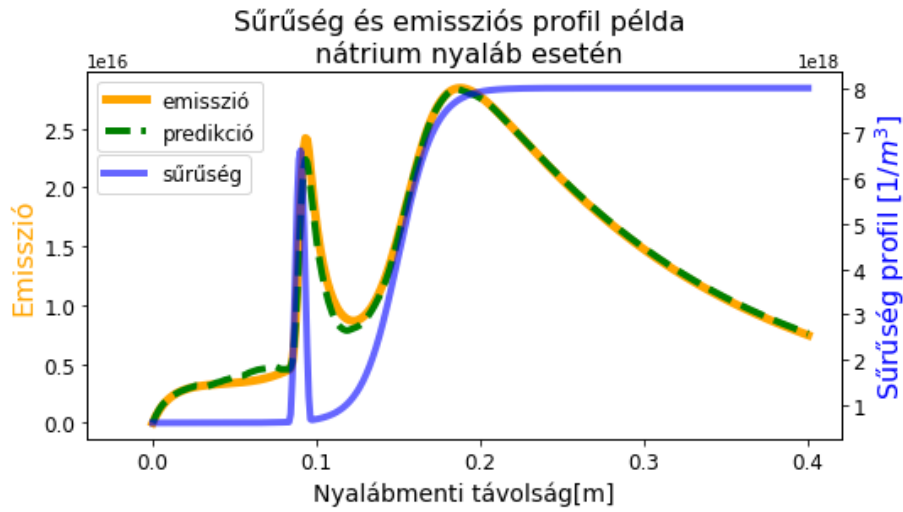
A hisztogramokról jól látáható, hogy bár a fluktuációs adathalmazokon a modellek hibája szignifikánsan növekedett a 16. ábrán látható eredményekhez képest, a hibák túlnyomó többsége továbbra is 10% alatti. A tendencia itt is hasonló mint amit a 5.5.1. fejezetben láthattunk, nátrium nyaláb esetén a modell hiba nagyjából kétszeres a lítiumhoz képest, ez az arány további modell optimalizálással talán javítható.

Mind a lítium, mind a nátrium esetében a magasabb hibával rendelkező esetek (nagyjából a 3% illetve 5% százalék feletti) az 5.1.2. fejezetben említett, BES esetén kevésbé releváns extrém esetek. Ezeket megvizsgálva elmondható, hogy bár a modell teljesítmény romlik ilyen esetekben, a kapott predikciók itt sem állnak túl távol a várt eredményektől. A neurális hálók ilyen extrém esetekben sem adnak vissza teljesen értelmetlen választ.

Egy-egy kiragadott példa a 18. és a 19. ábrán látható. Mindkét ábra nagy fluktuációt tartalmaz, de ezeknél extrémebb eseteket is vizsgáltam. A 18. ábrán a predikció gyakorlatilag elválaszthatatlan az eredeti emissziós profiltól. A 19. ábrán a modell hiba már szemmel látható. A határréteg plazmában lévő nagyméretű, összetartott plazmától elvált fluktuáció környezetében a modell predikciója kismértékben eltér a várt emissziós profiltól.



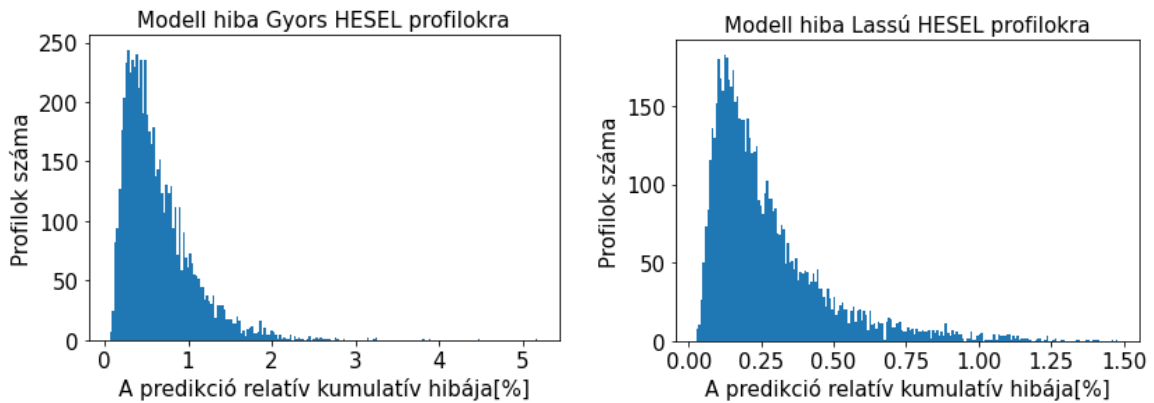
18. ábra. Egy példa a fluktuációs lítium adatkészletből.



19. ábra. Egy példa a fluktuációs nátrium adatkészletből.

5.5.3. Modell eredmények a HESEL adathalmazokon

Létrehoztam két ELM modellt az 5.1.3. fejezetben bemutatott HESEL adatkészletekre is, amik a fluktuációs tangenshiperbolikus adatkészleteknél komplexebb plazma-profilokat tartalmaztak. A hiba hisztogrammok a 20. ábrán láthatóak.



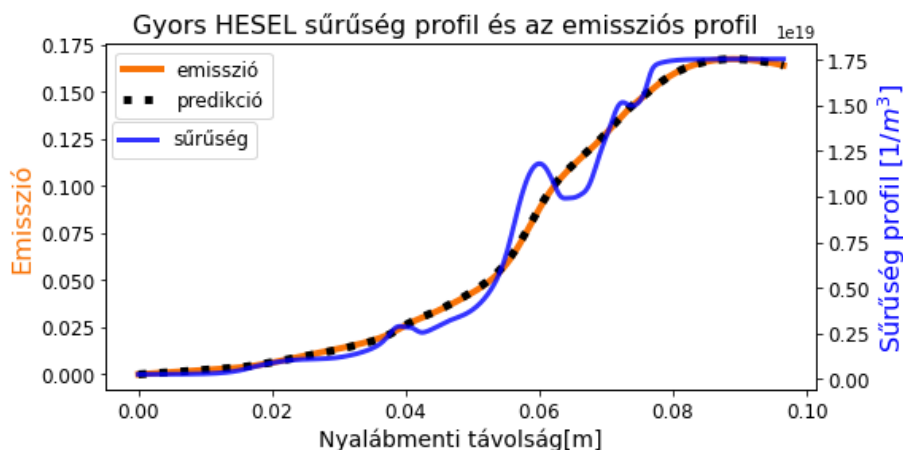
(a) Modell hiba az 1MHz-es mintavételezéssel szimulált gyors HESEL adatkészletre. Az átlagos hiba 0.6352%.

(b) Modell hiba az 1MHz-en mintavételezett és 0.5ms-ra kiintegrált lassú HESEL adatkészletre. Az átlagos hiba 0.2726%.

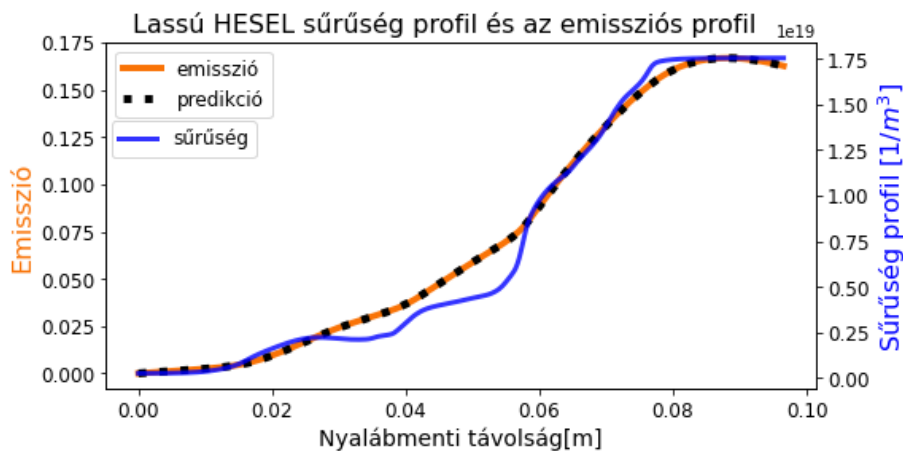
20. ábra. Modell hiba a HESEL adatkészleteken.

A hisztogramok alapján jól látható, hogy bár mindkét HESEL adatkészlet komplex plazmaprofilokat tartalmazott, a modell teljesítménye meghaladta azt amit a

fluktuációs tangens hiperbolikus adatoknál tapasztaltunk. Ez alapvetően a fluktuációs tangens hiperbolikus adatokban lévő számos extrém esetre vezethető vissza.



21. ábra. Egy példa az 1MHz-es mintavételezéssel modellezett gyors HESEL adatkészletből.

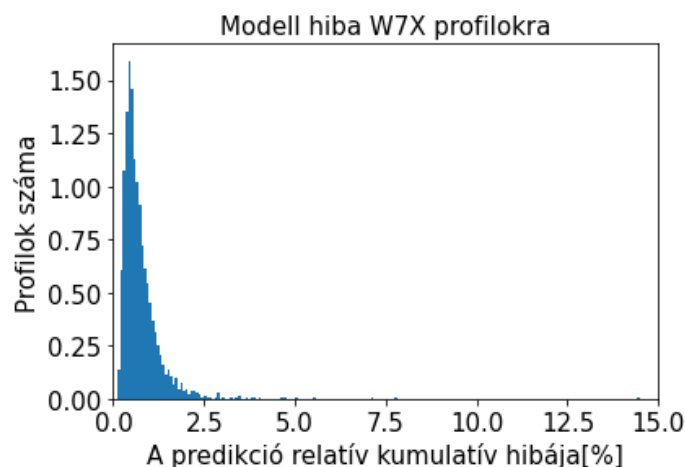


22. ábra. Egy példa az 1MHz-es mintavételezéssel modellezett és 0.5ms-ra kiintegrált lassú HESEL adatkészletből.

A 21-22. ábrákon látható egy-egy példa a gyors és a lassú HESEL adatkészletből. Jól látszik, hogy a 21. ábra sok fluktuációt tartalmaz, amik közül az egyik kifejezetten nagy. Ennek ellenére mint azt a hisztogramok is mutatják, a modell képes gyakorlatilag tökéletesen prediktálni a várható emissziót.

5.5.4. Modell eredmények a W7-X adathalmazon

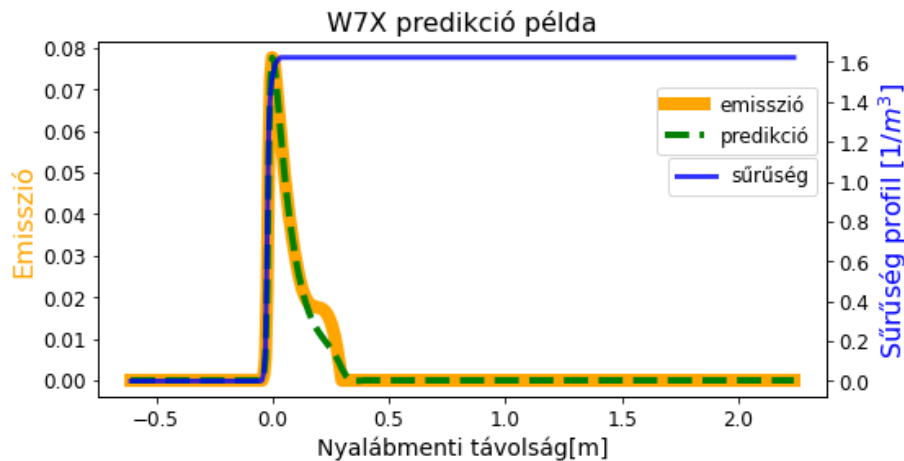
A kísérleti profilokból visszaállított W7-X adatkészletet először feldolgoztam a 5.2.2. fejezben leírtak szerint, majd erre is tanítottam és optimalizáltam egy ELM modellt. A modellhez tartozó hiba hisztogram a 23. ábrán látható.



23. ábra. Modell hiba a kísérleti adatokból visszaállított W7-X adatkészleten. Az átlagos modell hiba 0.7372%.

A modell jól teljesített ezen az adathalmazon is, a 0.7372%-os nagyjából a gyors HESEL adatkészleten elért eredményekkel egyezik meg.

Az profilok között volt 5 darab kiugró eset, ahol a modell hiba 5% fölé esett. Ezeket megvizsgáltam, és mind az öt profil esetén sok extrapolációt kellett alkalmazni az adatok feldolgozása során. A modell által vétett hiba alapvetően abból ered, hogy az extrapolálás csupán egy durva közelítést tudott nyújtani, az így kapott profilok nem fizikailag pontosak. Az extrapolálás más (5%-os hiba alatti) profilok esetén is negatívan hat a modell által adott predikcióra, ám ezekben az esetekben nem túlságosan nagy az eltérés. Mindazonáltal meg kell jegyezni, hogy az extrapoláció amit alkalmaztam nem tökéletes, ami az eredményeken is tapasztalható.



24. ábra. Egy példa a W7-X sűrűség profilok közül, ahol a sárgával jelölt emisszió feltételezhetően hibás.

A profilok közül egy érdekes példa látható a 24. ábrán. Itt egy egészen extrém gradienssel rendelkező tangens hiperbolikus alakú profilt láthatunk, amihez egy érdekes emisszió profil társul. A sárgával jelölt emisszió profil első csúcsa után látható egy sokkal kisebb csúcs is az emisszió lecsengésének végén. Ez a második csúcs a kékkel jelölt sűrűség profil alapján nem indokolt, valószínűleg a kísérleti adatok feldolgozása és a visszaállítás során került oda, hibásan.

Az, hogy az ELM háló predikciója ezt a második csúcsot nem tartalmazza kifejezetten jó eredmény. Ez azt mutatja, hogy a létrehozott háló helyesen tanulta meg a háttérben lévő fizikai kapcsolatokat.

5.6. Modell kimenetek "folytonossága"

Ahhoz, hogy az itt vizsgált ELM hálók kompromisszumok nélkül alkalmazhatóak legyenek a manapság használatos numerikus emisszió számító kódok helyett, nem csak a pontosságra van szükség. Sok esetben a modell által adott megoldások "folytonossága" ugyan olyan fontos mint az eredmény pontossága. Természetesen egy véges numerikus felbontással rendelkező profil esetén a matematikai definíció szerinti folytonosság nem értelmezhető. Itt ez a kifejezés alatt inkább azt kell érteni, hogy a profil kellően sima, a predikció (akárcsak az eredeti emisszió profilok) viszonylag lassan, a numerikus felbontásnál nagyobb skálán változik.

Ennek a tulajdonságnak a számszerűsítésére a profil göröngyösségének nevezett metrikát találtam ki. Ez a metrika a következő módszerrel számolható:

1. Első lépésként vesszük az összes emisszió predikciót amit a modell adott a teszt adatkészletre.

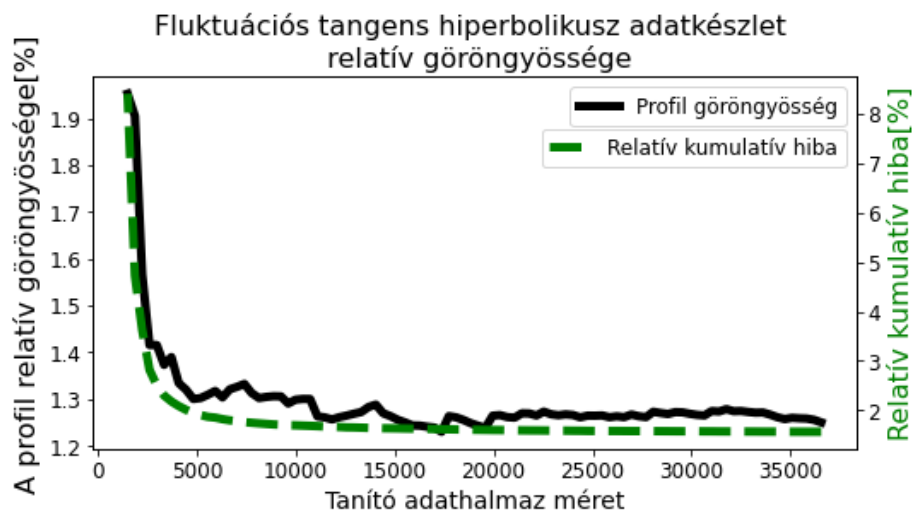
2. Minden predikció esetén egymás után eltávolítunk 1-1 pontot a profilból és a maradék pontokra egy görbét illesztünk. (ezt elvégezzük majd a profil minden pontjára kivéve a kezdő és a végpontot)
3. Kiszámoljuk a relatív különbséget az eltávolított ponthoz tartozó eredeti emisszió érték, és a predikcióra illesztett görbe között.
4. Minden egyes prediktált emisszió profil esetén vesszük a maximális relatív különbséget. Ez lesz a profilhoz tartozó relatív göröngyösség érték.
5. A teszt adatszett göröngyösség értéke pedig az adatszettben lévő maximális profil göröngyösség érték lesz.

Ezzel a módszerrel kapunk egy könnyen értelmezhető értéket arra, hogy mennyire vannak ugrások a modell által adott predikciókban. Mivel végső soron az egész adatkészletet a legrosszabb ponthoz tartozó értékkel jellemezzük, ez az érték egy erős felülbecslése lesz az átlagos modell teljesítménynek.

A göröngyösség vizsgálatát a fluktuációs lítium és a két HESEL adatkészletre végeztem el. A vizsgálat során nem csupán kiszámítottam mennyi a göröngyösség, hanem azt is vizsgáltam hogy hogyan változik ez az érték és a háló pontossága a tanító adathalmaz méretének függvényeként.

5.6.1. Fluktuációs tangens hiperbolikus adatkészlet göröngyössége

Először a fluktuációkkal ellátott tangens hiperbolikus-szerű lítium nyalábos adatszett esetén vizsgáltam a modell által adott predikciók göröngyösségét a 5.6. fejezetben leírtak szerint. Az eredmények a 25. ábrán láthatók.



25. ábra. A prediktált profilok göröngyössége és a modell hiba a tanító adathalmaz méretének függvényében a fluktuációs lítium adatszetre.

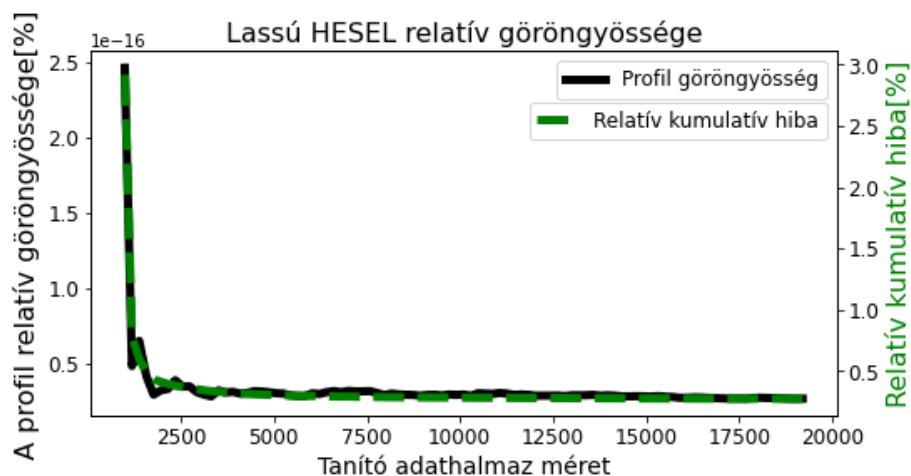
Az ábrán láthatjuk, hogy a feketével jelölt göröngyösség nagyjából öt-tíz ezer profilból álló tanító adathalmaz felett bekonvergál a minimális értékéhez ami 1.2-1.4%. Ugyen ez elmondható a modell pontosságára (az ábrán zöld), ami 1.5-2.0%-hoz konvergál.

Bár az 1.2-1.4%-os göröngyösség a következő fejezetekben bemutatott HESEL adatokra kapott eredményekhez képest soknak tűnhet, ez csupán az extrém esetekből ered. Hiszen a göröngyösség megadásánál az egész teszt adatkészlet esetén tapasztalt legnagyobb értéket vesszük, ami ebben az esetben egy extrém nagy amplitúdójú és oldal irányban kis kiterjedésű fluktuáció csúcsában lesz. Ezek mellett az 1% közeli érték önmagában is kellően kicsi, hogy a modellre azt mondhassuk "folytonos" kimeneteket ad.

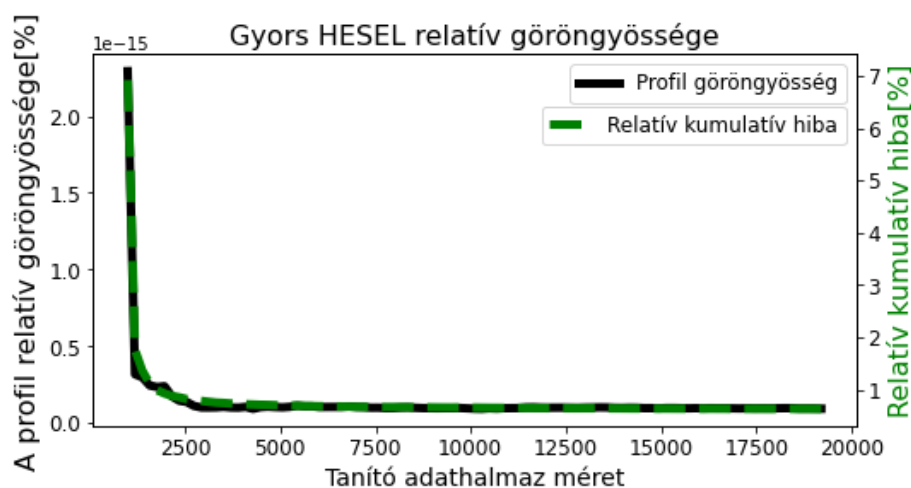
Egy másik nagyon fontos dolog amit láthatunk az ábráról, hogy mind a pontosság, mind a göröngyösség öt-tíz ezer profilból álló tanító adathalmaz után konvergál. Ez azt jelenti, hogy nincs szükség a nagyméretű 35000 profilos tanító adathalmazra, közel ugyan ezek az eredmények 10000 profillal is elérhetőek.

5.6.2. HESEL adatkészletek göröngyössége

A HESEL adatszettekre is elvégeztem a göröngyösség vizsgálatát, az eredmények a 26-27. ábrákon láthatók.



26. ábra. A prediktált profilok göröngyössége és a modell hiba a tanító adathalmaz méretének függvényében a lassú HESEL adatszetre.



27. ábra. A prediktált profilok göröngyössége és a modell hiba a tanító adathalmaz méretének függvényében a gyors HESEL adatszetre.

Az ábrákon azt láthatjuk, hogy a HESEL adatok esetén (amik nem rendelkeznek extrém méretű fluktuációkkal mint a fluktuációs lítium adatszett) még alacsonyabb lesz a profilok göröngyössége. Nagyjából 2500 profilos tanító adatszett felett mind a két model göröngyössége konvergál, a lassú HESEL nagyjából $0.5 \cdot 10^{-16}\%$ -hoz, míg a gyors HESEL $0.2 \cdot 10^{-15}\%$ -hoz, amik teljesen elhanyagolható

értékek. Ez alapján elmondható, hogy az összes modell predikció kellően "folytonos", valamint az adat numerikus felbontása elég magas, hogy a plazma profilok lassan változónak látszanak rajta.

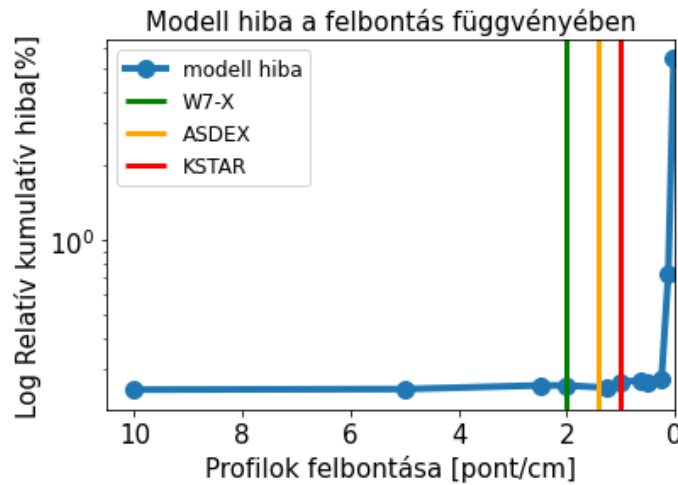
Nagyjából 2500 profilos tanító adatszett után a modell hibák is konvergálnak, így az is elmondható hogy ekkora tanító adatkészlet elég a kívánt eredmények eléréséhez.

5.7. Felbontásra való érzékenység

A korábbi fejezetekben a szintetikus diagnosztikákra jellemző viszonylag magas (10 pont/cm) numerikus felbontással dolgoztam, hogy kellő részletességgel tudjam vizsgálni a plazma profilokat és a rajtuk elhelyezkedő fluktuációkat. Ám ilyen felbontás általában csak a profilok rekonstruálása után áll rendelkezésre, a BES diagnosztikák jóval alacsonyabb felbontással dolgoznak. A W7-X berendezésen lévő BES rendszer például 40 csatornán figyeli meg az emissziót, ami 2 pont/cm-es felbontásnak felel meg[32].

Emiatt azt is megvizsgáltam, hogy hogyan viselkednek az ELM modellek, ha csökkenem a sűrűség és az emisszió profilok felbontását. Erre a célra a tangens hiperbolikus alakú lítium adatszettet használtam, amiből interpoláció nélkül alakítottam ki a kisebb adatszeteket. Ezt (a kívánt mérettől függően) a 400 pontból (10 pont/cm) minden x -edik pont kiválasztásával értem el. Így elkerültem a szükségtelen interpolációból hozzáadott hibát.

Az eredmények a 28. ábrán láthatók. Az ábrán látható, hogy a profilok felbontása 1 pont/cm (az ábrán is látható tipikus mérési felbontás) alá csökkenthető anélkül, hogy növelnénk a modellhibát. Ez alapján kijelenthető, hogy az ELM modellek képesek akár a fúziós berendezésekbe épített diagnosztikák saját felbontásával is dolgozni, nincs szükség a profilok előzetes feldolgozására.



28. ábra. Modell hiba a sűrűség és emisszió profilok numerikus felbontásának függvényében. Az ábrán a W7-X, ASDEX, illetve KSTAR fúziós berendezéseken lévő BES diagnosztikák mérési felbontását is jelöltem.

5.8. Modell gyorsaság

A 5.4-5.7 fejezetekben bemutattam, hogy milyen pontossággal rendelkeznek az ELM modellek amiket létrehoztam, milyen "folytonossággal" tudnak prediktálni, valamint mennyire érzékenyek a felbontásra. Ezek mind nagyon fontos paraméterek, ám a gyorsaságról nem minden alfejezetben esett külön szó.

Ebben a dolgozatban az összes bemutatott kétrétegű ELM modell rendkívül hasonló architektúrával rendelkezik, így a gyorsaságuk megegyező lesz (a különböző felbontással dolgozó modellek közt lesznek különbségek).

Mint a 3. táblázatban is említettem, ezek a modellek (400 pontos felbontás esetén) az általam használt hardver és szoftver környezetben 0.17 másodperc alatt tudnak 1000 emisszió predikciót előállítani. Ugyan ebben a szofver és hardver környezetben az adatgeneráláshoz használt RENATE-OD kód 1448.14 másodperc alatt képes 1000 emisszió profil kiszámítására.

Ez azt jelenti, hogy az ELM hálók egy 8518-szor gyorsabb megoldást jelentenek, miközben az esetek túlnyomó többségében elhanyagolható hibával rendelkeznek. Mindezek mellett az ELM hálók sajátos architektúrája miatt (mivel a tanításuk során nincs szükség backpropagation-re[24]) a tanításuk is csupán másodperceket vesz igénybe, nem igényel sok számítási kapacitást.

6. Konklúzió

A nyalábemissziós spektroszkópia (BES) a plazmasűrűség mérésére használt aktív plazmadiagnosztikai módszer, ezen rendszerek modellezésére számításigényes szintetikus diagnosztikákat használunk, mint például a RENATE-OD.

A RENATE-OD egy ütközéses sugárzási modellt (CRM) használ egy semleges nyaláb mentén történő emisszió kiszámításához, adott plazmasűrűség- és hőmérsékletprofil mellett. Ez a számítás egy differenciálegyenlet-rendszer megoldását foglalja magában, ami erősen korlátozza a diagnosztika alkalmazhatóságát a 3D-nyaláb modellezésében, ahol a számítást alkalmazástól függően 10^6 – 10^{10} -szor is szükséges lehet elvégezni.

A gépi tanulási megközelítések viszonylag újak a fúzió területén, céljuk általában a különböző plazmafizikai és diagnosztikai modellek számítási idejének javítása, mint például a RENATE-OD. Munkám fő célja az volt, hogy értékeljem a gépi tanulási módszerek alkalmazhatóságát a BES forward modellezési folyamatára.

Munkám során lineáris regressziót, illetve perceptron és Extreme Learning Machine (ELM) hálókat vizsgáltam és megmutattam, hogy egyszerűsített sűrűség profilokkal lítium BES esetén a 0.2631%-os bizonytalansággal rendelkező ELM hálók 8500-szorosan felgyorsíthatják a BES forward modellezést.

Ezek után megmutattam, hogy ez kiterjeszthető nátrium és deutérium nyalábokat használó BES rendszerekre is 0.4821%-os és 0.1740%-os bizonytalansággal, illetve fluktuációkkal ellátott szintetikus plazmaprofilokra is.

Vizsgáltam továbbá az ELM modellek alkalmazhatóságát egy 2D plazmaturbulencia modell (HESEL[4]) által létrehozott fluktuáló plazmaprofilokra 100kHz-es illetve 1MHz-es mintavételezés esetén, ahol a hálók szintén képesek voltak az emisszió prediktálására csupán 0.2726%-os és 0.6352%-os bizonytalansággal.

Ezek után megmutattam, hogy a szintetikus adatokon kívül a modellek használhatóak W7-X kísérleti adatokból visszaállított sűrűség profilok esetén is, csupán 0.7372%-os bizonytalansággal.

Azt is demonstráltam, hogy az ELM modellek által adott eredmények folytonossága a numerikus megoldók megoldásainak folytonosságához mérhető. Emiatt a BES mérésértékeléshez használt kódokba való integrációjuk nem okoz folytonossági problémákat. Emellett azt is felismertem, hogy az általam vizsgált esetekben a modellek tanítására kis méretű, esettől függően 2500-10000 profilból álló adatszetek is elégségesek.

Végül megmutattam, hogy az ELM modellek nem érzékenyek a sűrűség és emisszió profilok felbontására, képesek akár a BES mérésekben tipikus 1 pont/cm-es felbontási szinten is megbízhatóan működni.

7. Köszönetnyilvánítás

Szeretnék köszönetet mondani Asztalos Örsnek és Pokol Gergőnek minden segítségért és támogatásért, illetve hogy bármilyen fúziós plazmafizikát érintő kérdésemmel fordulhattam hozzájuk. Szeretnék köszönetet mondani Azarakhsh Jalalvandnak, hogy segített nekem a gépi tanulással és neurális hálózatokkal kapcsolatos kérdéseimben. Szeretném megköszönni Csabai Itvának, Oz Kilimnek és a Gépi Tanulási és Adat-Intenzív Tudományok Kutatócsoportnak, hogy megtanították nekem a gépi tanulás alapjait és mindig segítettek az ezzel kapcsolatos kérdéseimben.

A TDK dolgozatomban bemutatott munkát a HUN-REN Energiatudományi Kutatóközpont és az EUROfusion támogatta.

8. Bibliográfia

Hivatkozások

- [1] D. M. Thomas *et al.*, Fusion Sci. Technol. **53.**, 487-527, (2008).
- [2] D. Guszejnov, G. I. Pokol, I. Pusztai, et al., “Three-dimensional modeling of beam emission spectroscopy measurements in fusion plasmas,” Review of Scientific Instruments, vol. 83, no. 11, p. 113 501, 2012. DOI: 10.1063/1.4764564.
- [3] Peter Steiner and Azarakhsh Jalalvand and Simon Stone and Peter Birkholz, “PyRCN: A Toolbox for Exploration and Application of Reservoir Computing Networks”, Engineering Applications of Artificial Intelligence 113 (2022) 104964
- [4] A.H. Nielsen, O. Asztalos *et al.*, Nuclear Fusion **59.**, 0029-5515, (2019).
- [5] Barbarino, M. A brief history of nuclear fusion. Nat. Phys. 16, 890–893 (2020). <https://doi.org/10.1038/s41567-020-0940-7>
- [6] International Atomic Energy Agency website, Matteo Barbarino, accessed 28 May 2023, <https://www.iaea.org/newscenter/news/what-is-nuclear-fusion>

- [7] G. Federici, L. Boccaccini, F. Cismondi, M. Gasparotto, Y. Poitevin, I. Ricipito, "An overview of the EU breeding blanket design strategy as an integral part of the DEMO design effort", *Fusion Engineering and Design*, Volume 141, 2019, Pages 30-42, <https://doi.org/10.1016/j.fusengdes.2019.01.141>.
- [8] J. D. Lawson, "Some criteria for a power producing thermonuclear reactor", *Proceedings of the Physical Society. Section B*, vol. 70, pp. 6-10, jan 1957.
- [9] Ongena, J., Koch, R., Wolf, R. et al. Magnetic-confinement fusion. *Nature Phys* 12, 398–410 (2016). DOI: 10.1038/nphys3745
- [10] John Wesson, D. J. Campbell, J. W. Connor, R. D. Gill, N. C. Hawkes *Tokamaks*(4th Edition) Published 2011 by Oxford University Press ISBN: 978-0-19-959223-4
- [11] Beidler, Craig D., et al. "Stellarator fusion reactors—An overview." *Proceedings of Toki Conference ITC12*. 2001.
- [12] R. Wolf, "Fusionsforschung: Auf dem Weg zu einer neuen Primärenergiequelle." Presentation.
- [13] K. McCormick, S. Fiedler, G. Kocsis, J. Schweinzer, and S. Zoletnik, *Fusion Eng. Des.* 34–35, 125 (1997.), DOI: 10.1016/s0920-3796(96)00685-0
- [14] Asztalos Örs, "Modell-aided design and interpretation of beam emission spectroscopy measurements on fusion devices." (2022).
- [15] M. Lampert, G. Anda, A. Czopf, G. Erdei, D. Guszejnov, Á. Kovácsik, G. I. Pokol, D. Réfy, Y. U. Nam, and S. Zoletnik, "Combined hydrogen and lithium beam emission spectroscopy observation system for Korea Superconducting Tokamak Advanced Research", *Review of Scientific Instruments*, vol. 86, no. 7, p. 073501, 2015.
- [16] D. M. Thomas, "Development of lithium beam emission spectroscopy as an edge fluctuation diagnostic for DIII-D (invited)", *Review of Scientific Instruments*, vol. 66, no. 1, pp. 806-811, 1995.
- [17] S. Zoletnik, G. Anda, M. Aradi, O. Asztalos, S. Bató, A. Bencze, M. Berta, G. Demeter, D. Dunai, P. Hacek, S. Hegedűs, G. H. Hu, T. Križsan óczy, M. Lampert, D. Nagy, J. Németh, M. Otte, G. Petravich,

- G. I. Pokol, D. Réfy, B. Tál, and M. Vécsei, "Advanced neutral alkali beam diagnostics for applications in fusion research (invited)", *Review of Scientific Instruments*, vol. 89, no. 10, p. 10D107, 2018.
- [18] J. Schweinzer, E. Wolfrum, F. Aumayr, M. Pockl, H. Winter, R. P. Schorn, E. Hintz, and A. Unterreiter, "Reconstruction of plasma edge density profiles from Li I (2s-2p) emission profiles", *Plasma Physics and Controlled Fusion*, vol. 34, pp. 1173-1183, jul 1992.
 - [19] R. Hecht-Nielsen, "Neurocomputing: picking the human brain," in *IEEE Spectrum*, vol. 25, no. 3, pp. 36-41, March 1988, DOI: 10.1109/6.4520.
 - [20] Sumit Kumar Dam, Choong Seon Hong, Yu Qiao, Chaoning Zhang, "A Complete Survey on LLM-based AI Chatbots", *arXiv preprint arXiv:2406.16937*, 2024.
 - [21] Su, X., Yan, X. and Tsai, C.-L. (2012), Linear regression. *WIREs Comp Stat*, 4: 275-294., DOI: 10.1002/wics.1198
 - [22] Towards AI website, the websites editorial team, accessed 28 May 2023, <https://towardsai.net/p/l/perceptron-a-basic-neural-network-model-for-deep-learning>.
 - [23] da Silva, I.N., Hernane Spatti, D., Andrade Flauzino, R., Liboni, L.H.B., dos Reis Alves, S.F. "Multilayer Perceptron Networks" in *Artificial Neural Networks*, pp. 55-115, 2017, ISSB: 978-3-319-43162-8. DOI: 10.1007/978-3-319-43162-8_5.
 - [24] Chauvin, Y., & Rumelhart, D. E. (Eds.). (1995). *Backpropagation: Theory, architectures, and applications*. Lawrence Erlbaum Associates, Inc.
 - [25] Zang, Shaofei & Wang, Xuesong & Yan, Yongyi. (2021). Transfer Extreme Learning Machine with Output Weight Alignment. *Computational Intelligence and Neuroscience*. 2021. 1-14. DOI: 10.1155/2021/6627765.
 - [26] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489-501, 2006, *Neural Networks*, ISSN: 0925-2312. DOI: 10.1016/j.neucom. 2005.12.126

- [27] J. Tang, C. Deng, and G. Huang, "Extreme Learning Machine for Multilayer Perceptron," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 4, pp. 809–821, Apr. 2016, ISSN: 2162-2388. DOI: 10.1109/TNNLS.2015.2424995.
- [28] A. Jalalvand, O. Asztalos, M. Karacsonyi and G. I. Pokol, "Application of neural networks in beam emission spectroscopy modelling," 2023 International Joint Conference on Neural Networks (IJCNN), Gold Coast, Australia, 2023, pp. 1-8, DOI: 10.1109/IJCNN54540.2023.10192040.
- [29] B. Schrauwen, M. Wardermann, D. Verstraeten, J. J. Steil, and D. Strobandt, "Improving reservoirs using intrinsic plasticity," *Neurocomputing*, vol. 71, no. 7, pp. 1159–1171, 2008, ISSN: 0925-2312. DOI: 10.1016/j.neucom.2007.12.020.
- [30] Suchuan Dong, Zongwei Li, "A modified batch intrinsic plasticity method for pre-training the random coefficients of extreme learning machines", *Journal of Computational Physics*, vol. 445, 110585, 2021, ISSN 0021-9991, DOI: 10.1016/j.jcp.2021.110585.
- [31] Klaus Neumann, Jochen J. Steil, "Optimizing extreme learning machines via ridge regression and batch intrinsic plasticity", *Neurocomputing*, vol 102, 2013, pp. 23-30, ISSN 0925-2312, DOI: 10.1016/j.neucom.2012.01.041.
- [32] G. Anda, D. Dunai, T. Krizsanóczy, D. Nagy, M. Otte, S. Hegedűs, M. Vécsei, S. Zoletnik, G. Gárdonyi, "Measurement of edge plasma parameters at W7-X using alkali beam emission spectroscopy", *Fusion Engineering and Design*, Volume 146, Part B, 2019, Pages 1814-1819, ISSN 0920-3796, <https://doi.org/10.1016/j.fusengdes.2019.03.042>.
- [33] H. J. Wang, Y. Yu, R. Chen, Y. F. Wu, B. D. Yuan, S. B. Gong, Q. J. Yu, B. Lyu, Y. J. Shi, M. Y. Ye, and B. N. Wan, "Development of beam emission spectroscopy diagnostic on EAST", *Review of Scientific Instruments*, vol. 88, no. 8, p. 083505, 2017.
- [34] M. Willensdorfer, G. Birkenmeier, R. Fischer, F. M. Laggner, E. Wolf- rum, G. Veres, F. Aumayr, D. Carralero, L. Guimaraes, B. Kurzan, and ASDEX Upgrade Team, "Characterization of the Li-BES at ASDEX Upgrade", *Plasma Physics and Controlled Fusion*, vol. 56, p. 025008, jan 2014.

- [35] M. Brix, D. Dodt, A. Korotkov, P. Morgan, D. Dunai, R. Fischer, A. Meigs, I. S. Nedzelskiy, J. Schweinzer, J. Vince, and S. Zoletnik, "Upgrade of the lithium beam diagnostic at JET", *Review of Scientific Instruments*, vol. 81, no. 10, p. 10D733, 2010.
- [36] D. I. Réfy, M. Brix, R. Gomes, et al., "Sub-millisecond electron density profile measurement at the JET tokamak with the fast lithium beam emission spectroscopy system," *Review of Scientific Instruments*, vol. 89, no. 6, p. 069 902, 2018. DOI: 10.1063/1.5043551
- [37] H.K. Park and M.J. Choi and S.H. Hong and Y. In and Y.M. Jeon and J.S. Ko and W.H. Ko and J.G. Kwak and J.M. Kwon and J. Lee and J.H. Lee and W. Lee and Y.B. Nam and Y.K. Oh and B.H. Park and J.K. Park and Y.S. Park and S.J. Wang and M. Yoo and S.W. Yoon and J.G. Bak and C.S. Chang and W.H. Choe and Y. Chu and J. Chung and N. Eidietis and H.S. Han and S.H. Hahn and H.G. Jhang and J.W. Juhn and J.H. Kim and K. Kim and A. Loarte and H.H. Lee and K.C. Lee and D. Mueller and Y.S. Na and Y.U. Nam and G.Y. Park and K.R. Park and R.A. Pitts and S.A. Sabbagh and G.S. Yun and the KSTAR Team, "Overview of KSTAR research progress and future plans toward ITER and K-DEMO", *Nuclear Fusion*, Volume 59., Number 11., 2019., DOI: 10.1088/1741-4326/ab20e2
- [38] S. H. Jeong; D. H. Chang; T. S. Kim; S. R. In; K. W. Lee; J. T. Jin; D. S. Chang; B. H. Oh; Y. S. Bae; J. S. Kim; H. T. Park; K. Watanabe; T. Inoue; M. Kashiwagi; M. Dairaku; H. Tobar; M. Hanada, "First neutral beam injection experiments on KSTAR tokamak", *Review of Scientific Instruments*, Volume 83., Issue 2., 2012., <https://doi.org/10.1063/1.3660254>
- [39] K. Hallatschek, A. Zeiler; Nonlocal simulation of the transition from ballooning to ion temperature gradient mode turbulence in the tokamak edge. *Physics of Plasmas* 1 June 2000; 7 (6): 2554–2564. <https://doi.org/10.1063/1.874096>
- [40] Scikit-learn: Machine Learning in Python, Pedregosa et al., *JMLR* 12, pp. 2825-2830, 2011.
- [41] E Delabie and M Brix and C Giroud and R J E Jaspers and O Marchuk and M G O'Mullane and Yu Ralchenko and E Surrey and M G von Hellermann and K D Zastrow and JET-EFDA Contributors, "Consistency of atomic data for the interpretation of beam emission

spectra”, Plasma Physics and Controlled Fusion, Volume 52., 2010.,
DOI: 10.1088/0741-3335/52/12/125008

- [42] Leonard, A. W. (2014). Edge-localized-modes in tokamaks. Physics of Plasmas, 21(9), 090501. doi:10.1063/1.4894742

A. Saját eredmények

A munkám során a saját eredményeim adathalmazok létrehozásából, gépi tanulási modellek létrehozásából és tanításából, valamint az ezek által elért eredmények kiértékeléséből álltak.

Listaszerűen összefoglalva az alábbi saját eredményeket értem el:

1. Tangens hiperbolikus alakú, szintetikus nyalábemissziós spektroszkópia adathalmazok létrehozása lítium, nátrium és deutérium nyalábok esetén.
2. Tangens hiperbolikus alakú, fluktuációkkal ellátott szintetikus nyalábemissziós spektroszkópia adathalmazok létrehozása lítium és nátrium nyalábok esetén.
3. A kutatótársaktól kapott nyers W7-X adathalmaz feldolgozása és ebből neurális hálók tanítására alkalmas adathalmaz létrehozása.
4. Teljesítmény metrika definiálása a modell pontosságának megadásához.
5. Különböző gépi tanulás modellek (lineáris regresszió, perceptron hálók, ELM hálók) fejlesztése, tanítása, és kiértékelése a tangens hiperbolikus alakú lítium nyalábos adatkészletre. Annak bemutatása, hogy az ELM hálók különösen jól alkalmazhatóak erre a feladatra.
6. Az ELM típusú hálók tanítása és kiértékelése más (nátrium, deutérium) nyalábot használó adatkészletek esetén, valamint fluktuációs, turbulens és kísérleti adatokra.
7. Az ELM hálók által adott kimenetek "folytonosságának" vizsgálata, ehhez saját módszer kifejlesztése. Annak a bemutatása, hogy ezen eredmények minden vizsgált esetben kellően "folytonosak".
8. Az ELM hálók numerikus felbontásra tanúsított érzékenységének vizsgálata, annak a bemutatása, hogy a hálók a mérési felbontási szinten is jól működnek.
9. A nyalábemissziós spektroszkópia számításokban ELM modellekkel elérhető számításgyorsítás számszerűsítése.