

УНИВЕРЗИТЕТ У БЕОГРАДУ
МАТЕМАТИЧКИ ФАКУЛТЕТ



Давид Живковић

МОДЕЛИРАЊЕ ТЕРМОДИНАМИЧКЕ
СТРУКТУРЕ ФОТОСФЕРЕ СУНЦА
МЕТОДАМА МАШИНСКОГ УЧЕЊА

мастер рад

Београд, 2026.

Ментор:

др Иван Милић, доцент
Универзитет у Београду, Математички факултет

Чланови комисије:

др Анђелка КОВАЧЕВИЋ, ванредни професор
Универзитет у Београду, Математички факултет

др Ђорђе САВИЋ, виши научни сарадник
Астрономска опсерваторија у Београду

Датум одбране:

Захвалница

Пре свега, желим да се захвалим свом ментору, др Ивану Милићу, на сирпиљењу, подршци и неуморном менторству, као и на ведрини и позитивној атмосфери коју је стварао током нашеј заједничкој рада. Неизмерно сам захваљан што сам имао прилику да учим и радим са изузетним научником и човеком као што је он.

Захваљујем се Теодору Костићу (АОБ), пријатељу и колеги, на свим савештама, дискусијама до касно у ноћ и пажљивим коментарима који су допринели овом раду. Захвалност дућујем и Ђорђу Мијовићу на помоћи и корисним разговорима. Хвала мом пријатељу Константину Костићу на подршци, као и свим мојим пријатељима који су били уз мене током израде овог рада.

Посебну захвалност дућујем својој породици, која ми је омогућила да се посветим студијама и током свих ових година пружала ми подршку и веровала у мене, чак и у тренуцима када ми није ишло онако како сам желео.

На крају, захваљујем се својој девојци Катрини на неиресханој подршци, разумевању и томе што је увек била уз мене.

Наслов мастер рада: Моделирање термодинамичке структуре фотосфере Сунца методама машинског учења

Резиме: Посматрањима Сунчеве атмосфере не добијамо њену пуну тродимензионалну структуру. Мерени интензитет представља решење једначине преноса зрачења дуж правца визуре, па су доприноси различитих слојева испреплетани, а информација о великом делу домена остаје недоступна. Циљ ове тезе јесте да се испита у којој мери физички информисане неуронске мреже могу да реконструишу пун скуп физичких параметара Сунчеве фотосфере из непотпуног скупа података. Као тест проблем коришћена је MURAM симулација Сунчеве фотосфере димензија $300 \times 300 \times 48$ тачака, у којој су познате вредности температуре, притиска, поља брзина и магнетног поља у свакој тачки домена. Мрежа је постављена као непрекидна репрезентација домена, пресликавање просторних координата у осам физичких величина, при чему улазне координате пролазе кроз енкодер Фуријеових карактеристика. Реконструкција је испитана тако што су слојеви константне висине намерно уклањани из скупа за тренирање, а недостајући део домена реконструисан уз помоћ додатних чланова у функцији губитка изведених из магнетохидродинамичких једначина. Показано је да увођење услова $\nabla \cdot \vec{B} = 0$ смањује грешку реконструкције у слојевима без података у просеку за 11%, при чему је добитак највећи управо за компоненте магнетног поља, до 35% за B_z . Стационарна једначина континуитета доноси просечно побољшање од 5%, а највише погодује пољу брзина, до 11% за v_z . У оба случаја побољшање у слојевима без података праћено је извесним погоршањем у слојевима са подацима, што одговара очекиваном компромису између слагања са мерењима и са физичким моделом. Свака једначина при томе највише помаже управо оним величинама које експлицитно спреже. Резултати показују да физичка ограничења могу делимично надоместити недостајуће податке, што представља корак ка реконструкцији слојева Сунчеве атмосфере недоступних директним посматрањима.

Кључне речи: Сунце, фотосфера, магнетохидродинамика, машинско учење

Садржај

1	Увод	1
1.1	Сунце	5
1.1.1	Сунчева атмосфера	5
1.2	Симулације Сунчеве атмосфере	8
2	Машинско учење	10
2.1	Надгледано учење	10
2.2	Неуралне мреже	11
2.2.1	Вишеслојни перцептрон	12
2.3	Енкодер Фуријеових карактеристика	14
2.4	Физички-информисане неуралне мреже	16
3	Подаци и метод реконструкције	18
3.1	Подаци	18
3.2	Постављање неуралне мреже	21
3.2.1	Архитектура	21
3.2.2	Избор параметара	23
3.3	Увођење физике	24
3.3.1	Избор једначина за реконструкцију	25
4	Резултати	30
4.1	Учење над целокупним подацима	31
4.2	Уклањање слојева	41
4.2.1	Хомогено уклањање слојева	42
4.2.2	Нехомогено уклањање слојева	44
4.3	Додавање физичких једначина	47
4.3.1	Поређење синтетисаних спектра	53

САДРЖАЈ

5	Закључак и даљи рад	55
	Библиографија	58

Глава 1

Увод

Астрофизика је посматрачка наука, што значи да можемо користити само информације које прикупимо, не можемо да интерагујемо са системима које изучавамо. Основни начин прикупљања тих информација јесу посматрања, при чему готово сва (преко 99 %)¹ посматрања долазе у облику електромагнетног зрачења које системи које изучавамо емитују или на неки начин модификују. Имајући то у виду, посебно је важан развој метода којима из прикупљених посматрања процењујемо физичке параметре система који изучавамо.

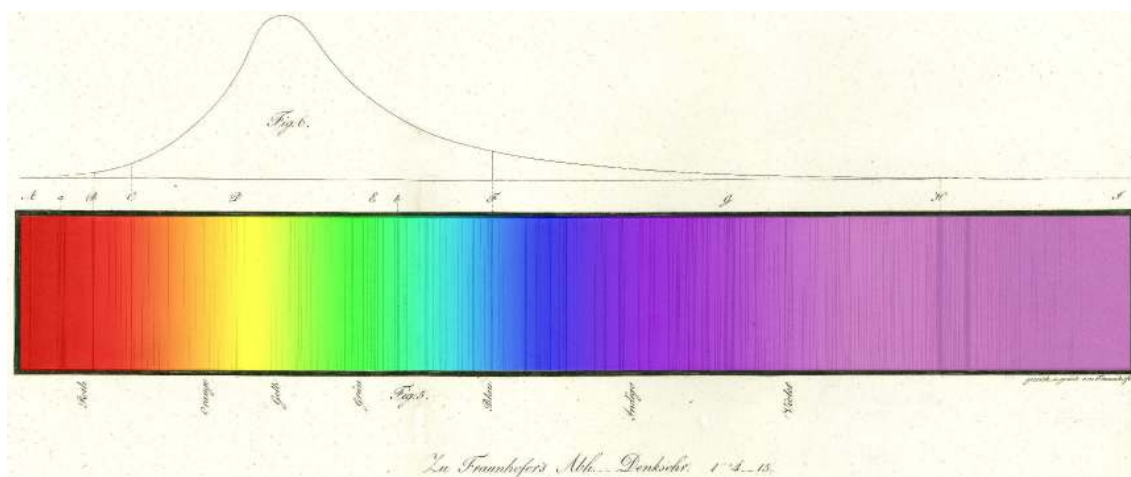
Једина звезда чију ”површину“ човек може посматрати јесте Сунце. Земља је трећа планета од Сунца и налази се на просечној удаљености од $1,5 \times 10^8$ km од Сунца. Нама најближа звезда после Сунца је Проксима Кентаури и налази се на око 4×10^{13} km. На тој удаљености са тренутном технологијом не можемо да разлучимо површину звезде, већ је видимо као материјалну тачку (Stix, 2002). Та, за астрономске појмове мала, удаљеност чини Сунце најпосебнијом звездом за нас и омогућава нам да детаљно изучавамо физичке процесе који се на њој одвијају. Ипак, и код Сунца непосредно посматрамо само његову атмосферу, будући да је унутрашњост непрозрачна за зрачење.

Научно проучавање Сунца почиње око 1610. године, са првим телескопским посматрањима Сунчевих пега (Ј. Фабрицијус, Галилео Галилеј, К. Шajnер и други). Голим оком тамне мрље на Сунчевом диску биле су забележене у Кини вековима раније, али је тек телескоп омогућио њихову систематску

¹Изузетак чине ретки други преносиоци информација: гравитациони таласи, неутрини, космички зраци и физички узорци попут метеорита.

анализу. Пратећи дневно кретање пега преко диска, Галилеј је показао да су пеге тамне творевине на самој Сунчевој површини, а не тела која прелазе преко диска (Foukal, 2004).

Прекретницу у соларним посматрањима донели су спектроскоп и фотографија у 19. веку. Фраунхофер је 1814. године конструисао први спектроскоп погодан за квантитативну анализу и измерио положаје преко 500 тамних линија у Сунчевом спектру (Слика 1.1). Исте тамне линије касније су уочене и у спектрима других звезда, што је поставило темељ спектралној класификацији звезда и показало да се методи развијени за проучавање Сунца могу применити на звезде уопште (Foukal, 2004).



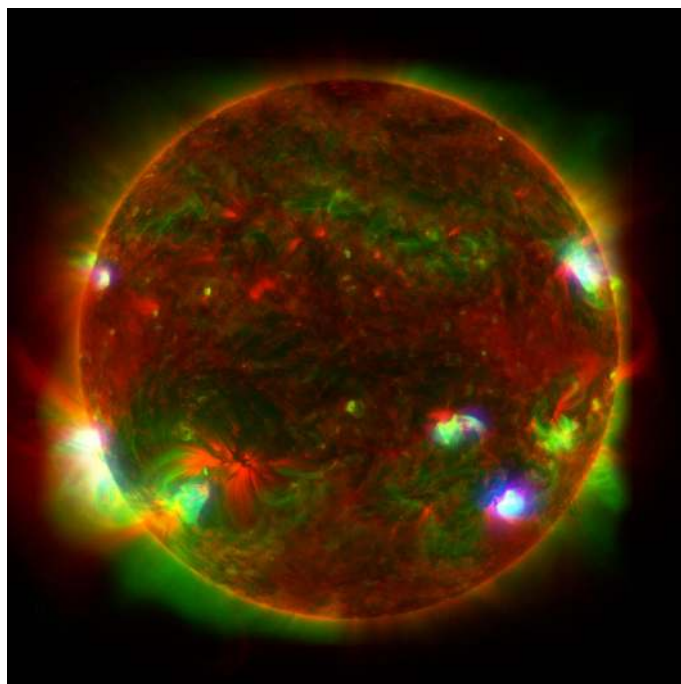
Слика 1.1: Фраунхоферова мапа Сунчевог спектра из 1814. године. Приказана верзија са додатом бојом преузета је из Pasachoff & Suer (2010).

Флукс Сунчевог зрачења је довољно велики да омогућава студије које за друге звезде нису изводљиве, посебно оне које истовремено захтевају врло високу спектралну, просторну и временску резолуцију (Stix, 2002). Ипак, и соларна посматрања имају ограничења, пре свега због Земљине атмосфере (сеинг², екстинкција), али и због ограничења самих инструмената.

Чак и код Сунца број расположивих фотона постаје проблем када се посматра у тамној пеги или у центру јаке спектралне линије. Ова ограничења

²Сеинг (*seeing*) је замућење и подрхтавање слике услед турбуленције у Земљиној атмосфери, које ограничава просторну резолуцију телескопа независно од његовог отвора.

превазилазе се већим апертурама и ефикаснијим детекторима, док се атмосферска ограничења заобилазе посматрањима из космоса. Од раних мисија попут SOHO-а и TRACE-а, преко Hinode-а и опсерваторије SDO, до данашњег Solar Orbiter-а, свемирске мисије омогућиле су непрекидна посматрања и приступ деловима спектра који до Земљине површине уопште не допиру (Stix, 2002). Пример таквих посматрања приказан је на слици 1.2.



Слика 1.2: Композитна слика Сунца, сачињена уз помоћ NUSTar, Hinode и SDO телескопа, на основу два канала у X домену и једног у ултраљубичастом домену³.

Ипак и поред напретка у посматрачким техникама (инфрацрвена спектроскопија, спескле-полариметрија, адаптивна оптика), посматрања имају суштинска ограничења. Оно што заправо меримо, интензитет зрачења, представља решење једначине преноса зрачења дуж правца визуре кроз атмосферу она даје нелинеарну везу између физичких параметара атмосфере и измереног интензитета. Услед те нелинеарности, мерени сигнал нам не даје једнозначну интерпретацију, а сви слојеви који нису доступни директним посматрањима остају тешки за реконструисати. Један од начина за ову реконструкцију је допуњавање измерених величина (температуре, магнетног поља, или поља брзина) једначинама магнетохидродинамике. На овом принципу

раде, нпр. технике за екстраполацију магнетног поља (Wiegelmann et al., 2014; Jarolim et al., 2023).

Поред тога, мале просторне скале остају испод резолуције тренутно постојећих инструмената. Оно што бисмо желели јесте приступ целој тродимензионалној структури атмосфере. Управо зато користимо нумеричке симулације које дају увид у комплетну структуру. Симулације тако помажу у изградњи конзистентних модела физичких процеса који стоје иза посматраних појава (Vögler et al., 2005).

Симулације су решења система парцијалних диференцијалних једначина које описују посматрани физички систем. За разлику од аналитичких решења, која постоје само за упрошћене случајеве, ова решења се добијају нумерички (Bodenheimer et al., 2007). Континуалне једначине се дискретизују на мрежи тачака и решавају нумеричким методама. Резултат је потпун тродимензионални модел у коме су вредности свих физичких величина познате у свакој тачки домена и у сваком тренутку. Тако успевамо да добијемо оне информације које посматрањима нису директно доступне. Овакве симулације су рачунарски веома захтевне, а поред тога њихова припрема и извођење траже много времена. Коначно, тешко је симулирати специфичан, посматрани догађај или феномен, пре свега због нелинеарности у једначинама магнетохидродинамике.

Машинско учење се овде надовезује тако што омогућава да се веза између симулираних атмосфера и спектроскопских или спектрополариметријских посматрања “научи” и затим примени на реална посматрања, чиме се рачунарски захтеван поступак замењује знатно бржим (Asensio Ramos et al., 2023). Ово ћемо детаљније објаснити у даљем тексту.

У овом раду испитујемо један такав приступ кроз проверу концепта (*proof-of-concept*) на симулацијама. Полазимо од потпуног тродимензионалног модела атмосфере, затим намерно уклањамо поједине слојеве и тренирамо неуралну мрежу да их реконструише из преосталих података. Циљ ове провере није директна примена на посматрања, већ провера да ли мрежа може да научи просторну структуру атмосфере на начин који би у будућности могао да се развије ка реконструкцији слојева недоступних реалним посматрањима.

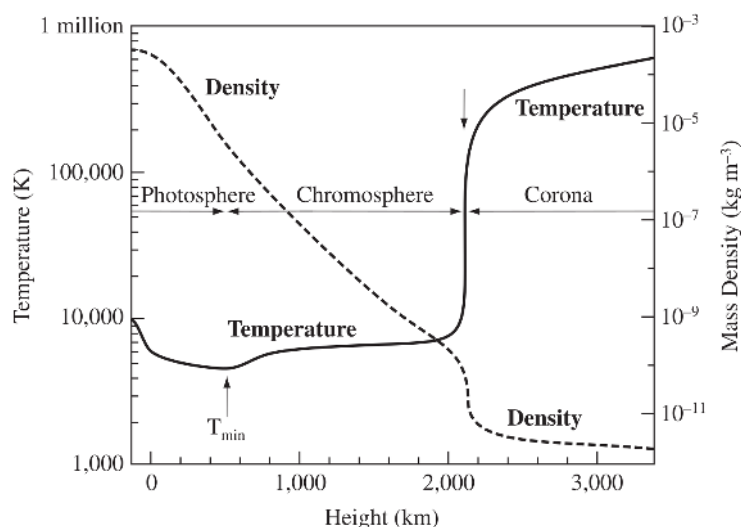
1.1 Сунце

Сунце је стабилна звезда спектралне класе G2V, масе $M_{\odot} = 2 \times 10^{30} \text{ kg}$ ⁴, и тиме спада у звезде главног низа. Своју енергију производи у центру термонуклеарним реакцијама (фузија водоника у хелијум). Унутрашњост Сунца можемо у основи да поделимо на језгро, радијативну зону и конвективну зону. Језгро је компактно и садржи око половине Сунчеве масе, карактеришу га високе температуре и притисци. Енергија се у језгру ствара у процесима термонуклеарних реакција и преноси се ка споља радијативно, при чему температура опада са повећањем радијуса до око $0,7 R$, након тога се пренос енергије наставља конвективно. Повећана непрозрачност и температурни градијент стварају конвекцију (Stix, 2002).

1.1.1 Сунчева атмосфера

Основни извор информација за физику Сунца јесте Сунчев спектар, који даје увид у физичке услове у Сунчевој атмосфери. Звездана атмосфера је прелазни регион из унутрашњости звезде у међузвездани простор, и дели се на фотосферу, хромосферу и корону (Слика 1.3). Кроз фотосферу температура опада са висином, до температурног минимума од око 4400 K на њеној горњој граници. Изнад минимума температура поново расте, најпре постепено кроз хромосферу, а затим нагло у прелазној области, где са неколико хиљада келвина скаче на милион келвина у корони (Stix, 2002). Механизам одговоран за то загревање и данас представља једно од отворених питања физике Сунца.

⁴Маса Сунца, $M_{\odot}$, представља стандардну јединицу у астрофизици, преко које се често изражавају масе других звезда и објеката.



Слика 1.3: Приказ зависности температуре и густирне од висине у атмосфери Сунца (Priest, 2014).

Фотосфера

Највећи део видљивог зрачења ствара се управо у фотосфери. Фотосфера дели непрозрачну Сунчеву унутрашњост од прозрчане атмосфере и тако дефинише Сунчеву „површину“. Простире се око 300 km. Фотосферу карактеришу појаве које називамо грануле. Грануле представљају конвективне структуре у којима се врела плазма из дубљих слојева Сунца уздиже ка површини брзином од око 1 km/s. Ови млазеви имају време трајања од 5 до 10 минута, затим се хладе и падају назад у унутрашњост. Грануле се приближно униформно простиру по Сунчевој површини (типична величина гранула је око 1000 km). Поред гранула, у фотосфери се јављају и друге структуре и појаве повезане са конвекцијом и магнетним пољем, као што су супергранулација, концентрације магнетног поља, факуле и сунчеве пеге (Stix, 2002).

Кроз фотосферу Сунце израчи већину своје емитоване енергије у облику непрекидног спектра (континуума). На континууму бележимо нагли мањак интензитета на уским распонима таласних дужина, то су апсорпционе (Фраунхоферове) линије, које носе информацију о физичким условима у слојевима у којима се формирају.

Хромосфера

Изнад фотосфере налази се слој под именом хромосфера, који се простире око 3000 km изнад фотосфере и знатно је ређи од ње. Температура прво опада до температурног минимума (око 4400 K), а затим са порастом висине почиње да расте. Механизам овог загревања је још увек једно од отворених питања физике Сунца. Близу горње границе хромосфере постоји нагли скок температуре са 10^4 на 10^6 K, тај танки слој називамо прелазна област (*transition region*) (Stix, 2002).

Корона

Изнад прелазне области налази се корона, најтоплији и најређи слој атмосфере Сунца. Много је блеђа од фотосфере, али је најпространији слој Сунчеве атмосфере. Нема оштру спољашњу границу, простире се на више Сунчевих радијуса и континуално прелази у Сунчев ветар. Димензије и облик короне, као и процеси који се у њој одвијају, у великој мери зависе од магнетне активности на Сунцу. Због мале густине, на плазму у корони битно утичу и слаба магнетна поља, па она одређују динамику и структуру короне. У време минимума активности корона је на половима сабијена, док је дуж екватора издужена, у периоду максимума активности корона се шири скоро симетрично у сваком правцу. Температура короне достиже максимум од око 2×10^6 K на висини од приближно 0,1 R изнад фотосфере, након чега постепено опада. Ако се Сунчев ветар третира као продужетак короне, онда у близини Земље она достиже температуре око 10^5 K (Stix, 2002).

Значај фотосфере

Магнетно поље представља везу између Сунчеве унутрашњости и виших слојева у атмосфери, оваква веза се простире чак и до спољних граница Сунчевог система. Оно настаје у унутрашњости, са највећом тачношћу може се измерити у фотосфери, а контролише већину физичких процеса у Сунчевој атмосфери (Wiegelmann et al., 2014). Управо у фотосфери и највишим слојевима конвективне зоне интеракција конвективних токова и магнетног поља доводи до низа процеса значајних како за физику Сунца, тако и у ширем астрофизичком контексту (Vögler et al., 2005).

Информација о структури и динамици фотосферских поља представља кључан улазни податак за разумевање магнетне повезаности целокупне Сунчеве атмосфере. Док се фотосферско магнетно поље рутински мери, његово простирање у више слојеве по правилу није доступно директним посматрањима, па се тродимензионална расподела поља у корони не мери, већ процењује методама екстраполације и симулације, у којима фотосферски магнетограм служи као гранични услов на доњој граници рачунске запремине (Jarolim et al., 2023). Ова зависност је утолико значајнија што је познавање те расподеле неопходно за разумевање настанка и покретања Сунчевих ерупција, као и за предвиђање високоенергетских догађаја (Jarolim et al., 2023).

Реконструкција пуног скупа физичких услова у фотосфери, а не само магнетног поља, стога није значајна само сама по себи, већ и као основа за конзистентније моделовање слојева изнад ње.

Мотивација за ову тезу је управо ограничена могућност реконструкције параметара у фотосфери, и покушај да се овај проблем реши методама машинског учења. У тези разматрамо тест проблем симулиране сунчеве фотосфере, где ћемо претпоставити да познајемо неке мерене вредности а недостајуће покушавамо да реконструишемо физички-информисаним неуралним мрежама.

1.2 Симулације Сунчеве атмосфере

Интеракција конвективних токова и магнетног поља у Сунчевој фотосфери и највишим слојевима конвективне зоне ствара низ процеса значајних како у соларном тако и у ширем астрофизичком контексту (Schüssler & Knölker, 2001). Како посматрања ове процесе региструју само посредно, у облику сигнала интегрисаног дуж линије видика у коме су доприноси различитих слојева испреплетани, нумеричке симулације магнетохидродинамичких (МХД) процеса представљају начин да се добије целокупна тродимензионална структура магнетних конфигурација. Да би се симулирали реални соларни услови симулације морају да обухвате и сложену физику као што је пренос зрачења, који је главни покретач конвекције и битно утиче на температурну структуру и сјај, интензитет области појачаног магнетног поља, и делимичну јонизацију која јако утиче на ефикасност преноса енергије кон-

векцијом (Vögler et al., 2005).

Један од кодова развијен управо за ову намену јесте тродимензионални МХД код MURAM⁵ за симулацију Сунчеве атмосфере и конвективне зоне, који укључује нелокални и несиви пренос зрачења (Vögler et al., 2005; Rempel, 2014). MURAM решава систем МХД једначина (1.1)–(1.4) нумерички у три просторне димензије за инерцијални референтни систем са константним гравитационим убрзањем у декартовим координатама. Ротација Сунца је занемарена како Кориолисова сила нема превелики утицај на скалама дужине специфичне за грануле (око 1000 km).

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \vec{v}), \quad (1.1)$$

$$\begin{aligned} \frac{\partial \rho \vec{v}}{\partial t} &= -\nabla \cdot (\rho \vec{v} \vec{v}) + \frac{f_{v_A}}{4\pi} \nabla \cdot \left(\vec{B} \vec{B} - \frac{1}{2} \vec{I} B^2 \right) \\ &= -\nabla P + \rho \vec{g}, \end{aligned} \quad (1.2)$$

$$\begin{aligned} \frac{\partial E_{HD}}{\partial t} &= -\nabla \cdot [\vec{v} (E_{HD} + P)] + \rho \vec{v} \cdot \vec{g} + \frac{\eta}{4\pi} \left(\nabla \times \vec{B} \right)^2 \\ &\quad + \vec{v} \cdot \frac{f_{v_A}}{4\pi} \nabla \cdot \left(\vec{B} \vec{B} - \frac{1}{2} \vec{I} B^2 \right) + Q_{rad}, \end{aligned} \quad (1.3)$$

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times \left(\vec{v} \times \vec{B} - \eta \nabla \times \vec{B} \right). \quad (1.4)$$

где су ρ густина, $\vec{v}$ вектор брзине, P притисак гаса, $\vec{B}$ вектор магнетне индукције, а $\vec{g} = -2,74 \times 10^4 \text{ cm s}^{-2}$ гравитационо убрзање. Величина E_{HD} означава хидродинамичку густину енергије (збир унутрашње и кинетичке енергије по јединици запремине), Q_{rad} је радијативни извор/попори енергије (дивергенција радијативног флукса), η је магнетна дифузивност, а $\vec{I}$ јединични тензор. Фактор f_{v_A} представља корекциони фактор за ограничавање Алфвенове брзине (Rempel, 2014), који у областима јаког поља и мале густине спречава да временски корак постане непрактично мали. Једначине (1.1)–(1.4) су редом једначина континуитета (одржање масе), једначина одржања количине кретања, једначина одржања енергије и индукциона једначина која описује еволуцију магнетног поља.

⁵Назив је акроним од *MPS/University of Chicago Radiative Magnetohydrodynamics*.

Глава 2

Машинско учење

Са развјетком посматрачких инструмената из године у годину се повећава количина података прикупљених у астрофизици. Док је пре само неколико деценија једна ноћ посматрања на великом телескопу давала свега неколико мегабајта података, данас су уобичајени гигабајти и терабајти, а надолазећи инструменти попут опсерваторије Vera C. Rubin (LSST) и радио-интерферометра SKA најављују петабајтне количине података (Norris, 2022). Прегледи неба (survey-и) доносе превише информација да би оне могле бити обрађене од стране човека. У последњих 10–20 година показала се велика корист рачунарских техника из области као што је машинско учење за обраду података у астрофизици. Машинско учење се данас успешно користи у многим астрофизичким областима за различите проблеме, од класификације извора и укрштања каталога на различитим таласним дужинама, до аутоматске обраде података у реалном времену (Norris, 2022).

У овом раду бавимо се надгледаним машинским учењем, које представља један од најчешће коришћених приступа у соларној физици (Asensio Ramos et al., 2023). Оно се ослања на статистичко учење и заснива се на прилагођавању модела познатим улазно-излазним подацима, тако да модел што боље апроксимира њихову међусобну везу (Hastie et al., 2009).

2.1 Надгледано учење

Надгледано учење (*supervised learning*) има за задатак да научи пресликавање између улаза (чији се скуп често означава са X) и излаза (означених са

Y), за које су доступни примери улазно-излазних парова. Из перспективе вероватноће, циљ надгледаног учења јесте моделовање условне расподеле $p(y|x)$ (Asensio Ramos et al., 2023).

Задатак учења се оквирно може формулисати овако: за дату вредност улазног вектора X , направити добру предикцију излаза Y , коју означавамо са $\hat{Y}$. Ако Y узима вредности у $\mathbb{R}$, исто важи и за $\hat{Y}$. За конструисање правила предикције потребни су подаци, и то често у великој количини. Претпостављамо стога да је доступан скуп мерења (x_i, y_i) , $i = 1, \dots, N$, познат као скуп података за тренирање (*training data*), на основу кога конструишемо предикцију (Hastie et al., 2009).

Циљ је, дакле, пронаћи функцију f која пресликава улазе у излазе, $f : X \rightarrow Y$. За разлику од приступа који сваки улазно-излазни пар обрађује независно, надгледано учење користи податке X и Y у целини да би прилагодило модел који апроксимира f . Ово се обично поставља као проблем оптимизације облика

$$f^* = \arg \min_f \mathcal{L}(f(X), Y), \quad (2.1)$$

односно, тражи се функција f таква да функција губитка $\mathcal{L}$, која мери неслagaње између предвиђених вредности $f(X)$ и стварних вредности Y , буде минимизована (Asensio Ramos et al., 2023). Функција f се изражава преко скупа параметара (нпр. тежина и одступања (*bias*) у неуралној мрежи), а прилагођавање се своди на подешавање тих параметара тако да се $\mathcal{L}$ минимизује, овај корак се назива **тренирање** модела.

2.2 Неуралне мреже

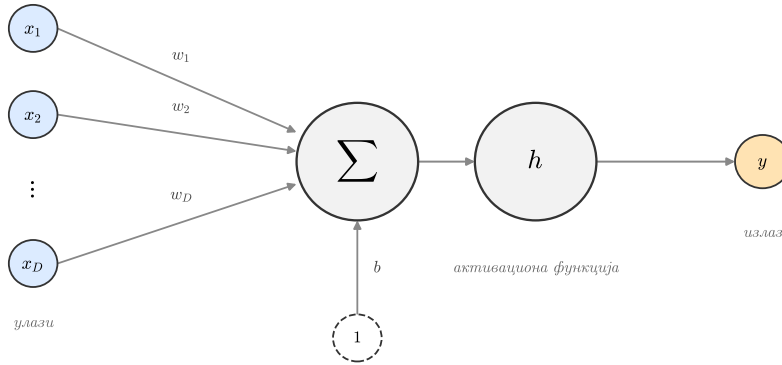
Остаје питање како функцију f конкретно параметризовати. Најједноставнији избор био би линеаран модел, у коме је излаз линеарна комбинација улаза. Како је веза између физичких величина у Сунчевој атмосфери нелинеарна такав модел неће бити одговарајућ. Потребан нам је довољно флексибилан облик функције, који може да опише сложену нелинеарну зависност, а да при томе остане описан коначним бројем параметара које можемо подешавати. Неуралне мреже управо то омогућавају (Bishop, 2006).

Неурон

Основни градивни елемент неуралне мреже јесте неурон (Слика 2.1). Неурон формира линеарну комбинацију улаза са одговарајућим тежинама и одступањима (*bias*), а затим је пропушта кроз активациону функцију h . Математички,

$$y = h\left(\sum_j w_j x_j + b\right), \quad (2.2)$$

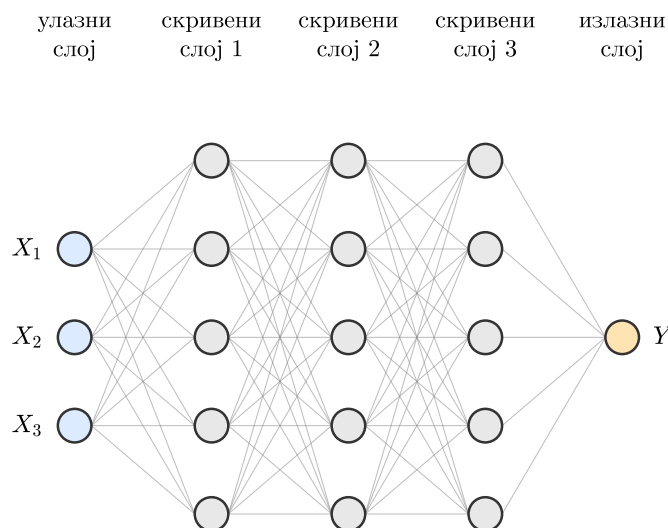
где су x_j улазне вредности, w_j тежине, а b одступање (Asensio Ramos et al., 2023). Излаз једног неурона постаје улаз наредног, који врши сличну операцију, па се неурална мрежа може посматрати као композиција великог броја оваквих једноставних нелинеарних функција.



Слика 2.1: Пример једног неурона.

2.2.1 Вишеслојни перцептрон

Неурони се групишу у слојеве, а број слојева одређује дубину мреже. Најједноставнија коришћена архитектура јесте потпуно повезана (*fully connected*) мрежа са пропагацијом унапред (*feed-forward*), у којој је сваки излаз претходног слоја повезан са сваким неуроном наредног слоја, а информација тече само у једном смеру, од улаза ка излазу. Такву мрежу називамо вишеслојним перцептроном (*multilayer perceptron*, MLP) (Слика 2.2) (Bishop, 2006).



Слика 2.2: Пример једног вишеслојног перцептрона који слика три улаза у један излаз, на основу три тзв. скривена слоја.

Мрежа се састоји од улазног слоја, једног или више скривених слојева и излазног слоја. Слојеве између улаза и излаза називамо скривеним јер њихове вредности нису непосредно посматране, већ представљају међурезултате које мрежа сама формира током тренирања. Свака јединица једног слоја рачуна своју вредност по изразу (2.2), узимајући као улаз излазе свих јединица претходног слоја.

Активационе функције

Активациона функција излазног слоја бира се према природи проблема. За регресионе проблеме, какав је наш, користи се идентитет, тако да излаз мреже може узети било коју реалну вредност, код класификационих проблема користе се сигмоидне (Bishop, 2006). Историјски су најчешће коришћене сигмоидне функције, попут логистичке сигмоиде

$$\sigma(a) = \frac{1}{1 + e^{-a}}, \quad (2.3)$$

и хиперболичког тангенса

$$\tanh(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}}. \quad (2.4)$$

Обе функције су глатке и монотоне, али се разликују по опсегу, логистичка сигмоида пресликава улаз у интервал $(0, 1)$, док $\tanh$ даје вредности у интервалу $(-1, 1)$ и центрирана је око нуле (LeCun et al., 1998).

Тренирање мреже

Оптималне вредности тежина и одступања нису познате унапред, оне се пре тренирања иницијализују случајним вредностима (Glorot & Bengio, 2010). Тренирање се затим своди на следеће: за парове улаза и излаза $(\mathbf{x}_i, \mathbf{y}_i)$ израчуна се функција губитка $\mathcal{L}$ из израза (2.1), која мери колико се предикције мреже разликују од стварних вредности, а параметри се потом ажурирају тако да се та вредност смањи. За то се користе методе засноване на градијентном спусту, параметри се померају у смеру супротном од градијента функције губитка.

Градијенти функције губитка по свим параметрима мреже рачунају се алгоритмом пропагације уназад (*backpropagation*) (Rumelhart et al., 1986), који систематски примењује правило извода сложене функције кроз слојеве мреже. Овај корак је могућ управо зато што су и активационе функције и сама мрежа диференцијабилне по параметрима. Иста особина, могућност ефикасног рачунања извода мреже (*automatic differentiation*) (Baydin et al., 2018), биће кључна и у наставку, када у функцију губитка будемо уводили физичка ограничења.

2.3 Енкодер Фуријеових карактеристика

Под енкодером подразумевамо фиксну трансформацију улаза која се примењује пре мреже. Уместо да сирове улазне вредности директно уђу у први слој, оне се најпре преводе у скуп трансформисаних величина, а мрежа затим ради над њима. Такво пресликавање нема подесивих параметара и не учи се током тренирања, већ представља начин да се проблем унапред постави у облику који је мрежи пријемчивији.

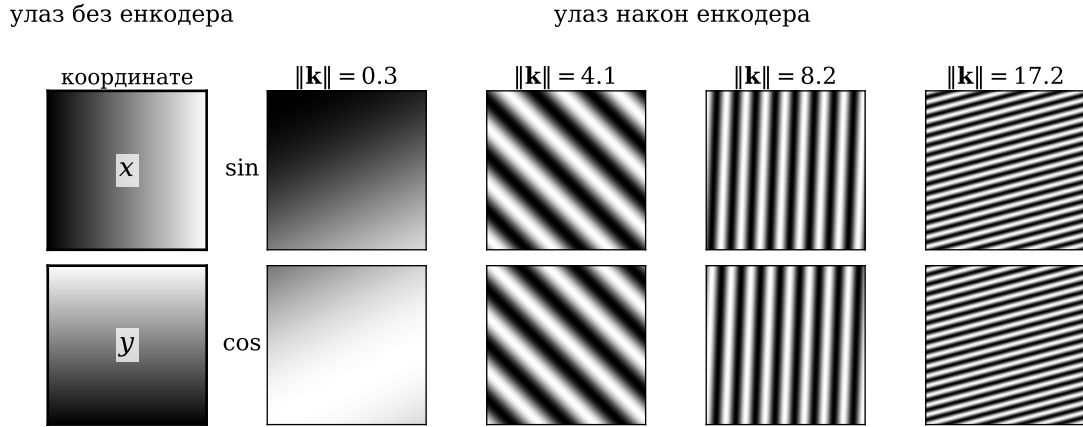
Ако се крећемо по домену, суседне тачке имају блиске координате, па мрежа природно тежи да им додели и блиске вредности. Такав опис добро репродукује глобалне структуре, али лоше нагле прелазе. То није случајност, мреже током тренирања најпре и најбрже уче глатке компоненте, док

фине структуре уче знатно спорије. Ова појава назива се спектрална пристрасност (*spectral bias*, Rahaman et al., 2019). За наше податке то је озбиљан проблем. Сунчева фотосфера истовремено садржи структуре врло различитих скала: гранулација уводи карактеристичну скалу реда 1000 km, али су прелази између гранула и међугрануларних трака оштри, док је магнетно поље концентрисано у уским областима које заузимају мали део површине. Мрежа склона глатким решењима дала би размазано поље у коме управо ти прелази и концентрације нису репродуковани, а они су физички најбитнији део решења.

Решење је да мрежи унапред понудимо структуре које би сама тешко направила. Уместо саме координате, у мрежу шаљемо скуп синуса и косинуса различитих фреквенција израчунатих из те координате

$$\gamma(\mathbf{x}) = [\mathbf{x}, \sin(2\pi\mathbf{K}\mathbf{x}), \cos(2\pi\mathbf{K}\mathbf{x})], \quad (2.5)$$

где је $\mathbf{K}$ матрица фреквенција, узоркована из нормалне расподеле и фиксирана пре тренирања. Ово називамо енкодером Фуријеових карактеристика (Tancik et al., 2020). Пример оваквог енкодера приказан је на слици 2.3.



Слика 2.3: Илустрација принципа Фуријеових енкодера у просторном домену.

Идеја је једноставна: синуси високих фреквенција већ садрже брзе промене, па мрежа не мора да их ствара — довољно је да им одреди амплитуде и да их међусобно комбинује. Показано је да овакво пресликавање омогућава

мрежама са координатним улазом да науче високофреквентне функције, што без њега није изводљиво (Tancik et al., 2020).

Кључан избор при томе јесте ширина фреквенцијског појаса, тј. колико високе фреквенције енкодер садржи. Преузак појас враћа нас на полазни проблем и даје замагљену реконструкцију, док преширок појас производи осцилаторно, зашумљено решење, будући да енкодер тада садржи фреквенције које подацима не одговарају. Избор појаса за наше податке описан је у одељку 3.2.1.

2.4 Физички-информисане неуралне мреже

До сада описано надгледано учење се у потпуности сводило на предвиђање задатих података. Мрежа се тренира тако да се минимизује функцију губитка која мери разлику између њене предикције и стварних вредности у тачкама у којима те вредности познајемо. У пракси подаци често нису потпуни, могу бити ретки, оптерећени шумом, или доступни само у делу домена. Тако мрежа у областима у којима података нема, нема ништа што би је ограничавало, па њена предикција тамо може бити произвољна и физички бесмислена. Управо ту физика може да надомести оно што подаци не покривају. Ако за одређени скуп података или за одређени део домена знамо да у њему важи неки физички закон, који је описан диференцијалном једначином (2.6).

$$\frac{\partial u}{\partial \mathbf{x}} + \mathcal{N}[u] = 0, \quad (2.6)$$

где је $\mathcal{N}[\cdot]$ нелинеаран диференцијални оператор. Такав закон можемо искористити као додатно ограничење, свако одступање предикције од њега представља грешку, коју је лако и измерити. Довољно је из излаза мреже израчунати резидуал одговарајуће једначине,

$$f := \frac{\partial u}{\partial \mathbf{x}} + \mathcal{N}[u], \quad (2.7)$$

који је по конструкцији једнак нули свуда где је једначина задовољена. Резидуал израчунавамо у N_c колокационих тачака $\mathbf{x}_i$ унутар домена и од њега формирамо део функције губитка заснован на једначини,

$$\mathcal{L}_f = \frac{1}{N_c} \sum_{i=1}^{N_c} |f(\mathbf{x}_i)|^2, \quad (2.8)$$

који представља средњу квадратну вредност резидуала. Захтевањем да $\mathcal{L}_f$ буде што мањи, мрежу усмеравамо ка решењима која боље задовољавају задату диференцијалну једначину. Овај члан затим додајемо делу функције губитка који описује одступање од података,

$$\mathcal{L} = \mathcal{W}_d \mathcal{L}_d + \mathcal{W}_f \mathcal{L}_f, \quad (2.9)$$

где $\mathcal{L}_d$ мери одступање предикције од података, а $\mathcal{L}_f$ одступање резидуала f од нуле. За разлику од $\mathcal{L}_d$, $\mathcal{L}_f$ не захтева никакве податке, потребно је само познавати једначину. Мреже трениране на овај начин називамо **физички информисаним неуралним мрежама** (*physics-informed neural networks*, PINNs).

Тежине $\mathcal{W}_d$ и $\mathcal{W}_f$ одређују релативни утицај ова два члана. Њиховим избором подешавамо колико се решење ослања на податке, а колико на физику. У различитим поставкама тај однос није исти, па се тежине бирају према томе колико поуздано у датом случају важе једно и друго.

Сам појам физички информисаних неуралних мрежа уведен је у раду (Raissi et al., 2019), мада идеја да се неурална мрежа искористи за решавање диференцијалних једначина потиче још од раније (Lagaris et al., 1998). У многим физичким проблемима прикупљање података је скупо, што отежава примену метода машинског учења. С друге стране, за велики број физичких система располажемо великом количином претходног знања, као што су физички закони. Ако се то знање угради у сам алгоритам учења, оно делује као механизам регуларизације који сужава простор допустивих решења, чиме се ефективно појачава информациони садржај расположивих података (Raissi et al., 2019).

Кључна техничка компонента јесте аутоматска диференцијација (*automatic differentiation*) (Baydin et al., 2018). Исто као што се рачуна градијент по параметрима уз помоћ аутоматске диференцијације, тако можемо да израчунамо и резидуал диференцијалне једначине директно из излаза мреже.

Овај приступ демонстриран је на низу класичних проблема математичке физике, од Бургерсове и Шредингерове до Навије-Стоксових једначина (Raissi et al., 2019). Показало се да мрежа може да реконструише решење из свега неколико стотина тачака.

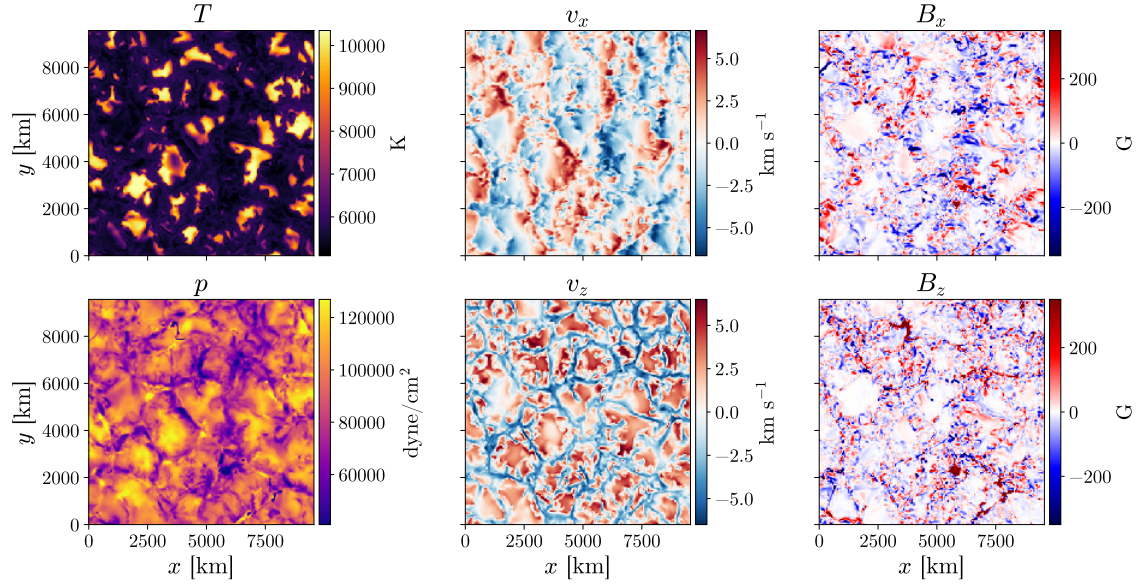
Глава 3

Подаци и метод реконструкције

У овој глави ћемо изложити коришћене податке за реконструкцију. У питању је једна симулација сунчеве фотосфере. За развитак метода често користимо симулације јер је онда познат тачан одговор па је лакше оценити тачност и перформансе метода.

3.1 Подаци

За тренирање мреже коришћени су подаци добијени MURAM симулацијом Сунчеве фотосфере, описаном у одељку (1.2). Расположиви скуп података представља тродимензионални домен димензија $300 \times 300 \times 48$ тачака (Слика 3.1), при чему су прве две координате хоризонталне, а трећа вертикална. Растојање између суседних тачака мреже износи $\Delta x = \Delta y = \Delta z = 32 \text{ km}$ у сва три правца, што одговара хоризонталном пресеку од приближно $9,6 \text{ Mm} \times 9,6 \text{ Mm}$ и вертикалном опсегу од -288 km до 1216 km . Фотосферски ниво налази се на деветом слоју висине, тако да домен обухвата и подфотосферске слојеве конвективне зоне и слојеве изнад фотосфере.



Слика 3.1: Приказ физичких параметара $(T, p, v_x, v_z, B_x, B_z)$ из симулације на висини $z = 0 \text{ km}$.

У свакој тачки домена познате су вредности температуре T , притиска P , густине ρ , све три компоненте вектора брзине (v_x, v_y, v_z) и све три компоненте вектора магнетне индукције (B_x, B_y, B_z) . Циљ је реконструисати сва ова поља осим густине. Мрежа стога представља пресликавање

$$f : (x, y, z) \mapsto (T, P, v_x, v_y, v_z, B_x, B_y, B_z), \quad (3.1)$$

дакле са три улаза и осам излаза.

Заменски модел за густину

Густину изостављамо из излаза главне мреже јер она није независна величина, већ је у симулацији одређена температуром и притиском преко једначине стања. Пошто ова једначина није дата у затвореном облику, засебно смо истренирали мању неуралну мрежу која служи као заменски модел (*surrogate model*) и пресликава пар (T, p) у густину. Архитектура и параметри тренирања овог модела приказани су у табели (3.1). Због великог распона вредности, мрежа је тренирана у логаритамском простору. Добијена вредност NRMSE_σ износи 1,538%, док је Пирсонов коефицијент корелације $r = 0,9999$.

Током тренирања главне мреже параметри заменског модела остају фиксирани, али градијенти кроз њега слободно теку назад у температуру и притисак.

Табела 3.1: Параметри заменског модела за једначину стања.

Параметар	Ознака	Вредност
Улазне величине	–	$\log T, \log p$
Излазна величина	–	$\log \rho$
Број скривених слојева	D	4
Број неурона у скривеном слоју	W	64
Активациона функција	–	$\tanh$
Оптимизатор	–	Adam
Почетна стопа учења	lr	2×10^{-3}
Распоред стопе учења	–	CosineAnnealingLR
Број епоха	–	3000
Величина серије података	–	200,000
Функција губитка	–	MSE у log простору

Нормализација

Просторне координате x , y и z , које представљају улазе мреже, пре тренирања линеарно нормализујемо на интервал $[0, 1]$ у сваком правцу. Физичке величине, које представљају излазе мреже, нормализујемо z -нормализацијом (3.2), чиме се њихови опсези, који се међусобно разликују за редове величине, свODE на вредности погодне за оптимизацију (LeCun et al., 1998).

$$Y_i^{\text{norm}} = \frac{Y_i - \mu_i(z)}{\sigma_i(z)}. \quad (3.2)$$

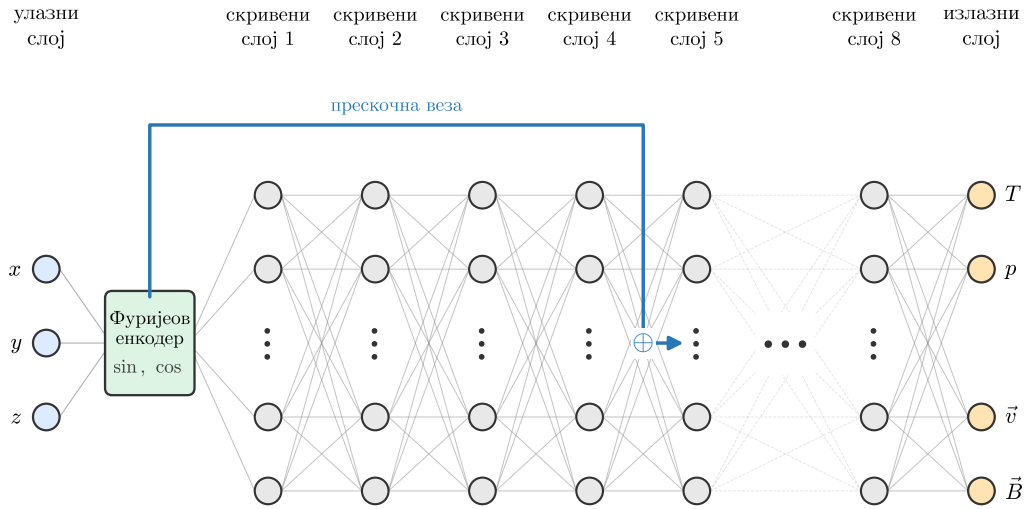
Овде Y_i означава i -ти од осам излаза мреже, а $\mu_i(z)$ и $\sigma_i(z)$ његову средњу вредност и стандардну девијацију. Нормализацију радимо засебно за сваки хоризонтални слој, будући да се средње вредности мењају за редове величине са висином.

Кључно је при томе да се профили $\mu_i(z)$ и $\sigma_i(z)$ рачунају искључиво из слојева доступних за тренирање. Како ћемо поједине z -слојеве уклањати из тренирања, за њих неће бити доступне ни вредности μ_i и σ_i . На уклоњеним слојевима профиле зато добијамо линеарном интерполацијом између доступних слојева, уз изузетак притиска, чији профил интерполирамо у логаритамској скали, будући да он опада експоненцијално са висином.

3.2 Постављање неуралне мреже

3.2.1 Архитектура

Мрежа коју користимо пресликава просторне координате у вредности физичких величина у тој тачки, $f : (x, y, z) \mapsto (T, p, v_x, v_y, v_z, B_x, B_y, B_z)$, дакле има три улаза и осам излаза (Слика 3.2).



Слика 3.2: Скица дате архитектуре коју смо користили.¹

Нормиране координате најпре пролазе кроз Фуријеов енкодер (Tancik et al., 2020), који их пресликава у скуп синусних и косинусних функција. За $N_f = 192$ добија се излазни вектор димензије $d = 3 + 2N_f = 387$. Матрица фреквенција генерише се насумично из нормалне расподеле и остаје фиксирана током тренирања.

Стандардне девијације расподела одређене су на основу спектра снаге симулираних података. Температури, притиску и брзинама одговара ниско-фреквентни опсег са $\sigma * \text{low} = (2,55, 2,55, 0,48)$, док магнетном пољу одговара

¹На слици смо $\vec{v}$ и $\vec{B}$ приказали као један излаз, ови вектори заправо сликају три излаза, што даје укупан број излаза једнак осам.

високофреквентни опсег са $\sigma * \text{high} = (9,23, 9,23, 0,48)$. Од укупног броја фреквенција, 15% узоркујемо из нискофреквентне, а 85% из високофреквентне расподеле. Високофреквентни опсег узоркујемо гушће због присуства ситнијих просторних структура у магнетном пољу.

Тако добијени вектор улази у вишеслојни перцептрон дубине $D = 8$ и ширине W . Као активациона функција свих скривених слојева коришћен је хиперболички тангенс, из разлога изложених у одељку (2.2.1), док излазни слој нема активациону функцију, са обзиром да је циљ регресија параметара. Тежине су иницијализоване Ксавијеровом униформном расподелом (Glorot & Bengio, 2010), а одступања постављена на нулу.

На средини мреже уведена је прескочна веза (*skip connection*) (Слика 3.2), излаз енкодера поново се придружује улазу четвртог скривеног слоја. Дубоке мреже теже да у узастопним слојевима постепено изгубе високофреквентну информацију садржану у улазу, поновним придруживањем енкодерског вектора та информација постаје директно доступна и другој половини мреже. Управо ту структуру користи и NeRF (Mildenhall et al., 2020), преузимајући је из Park et al. (2019). Прескочна веза се при томе оправдава побољшаним протоком информације кроз мрежу и бржом конвергенцијом током тренирања. Такође су исту архитектуру неуралне мреже користили за репрезентацију короналног магнетног поља (Jarolim et al., 2023).

Ефикасност мреже

Овако постављена мрежа представља континуалну апроксимацију дискретног симулационог домена, вредности физичких величина се чувају у параметрима мреже. Природна мера ефикасности такве репрезентације јесте однос броја вредности које чува оригинални сет података и броја параметара мреже

$$C = \frac{N_x N_y N_z N_{ch}}{N_{par}}, \quad (3.3)$$

који ћемо називати фактором компресије (Jarolim et al., 2023). Где су N_x, N_y, N_z број просторних тачака, N_{ch} број канала (*channels*), тј. број физичких параметара и N_{par} је број параметара мреже. Вредност $C > 1$ значи да репрезентација заузима мање места од самих података. Компресија при томе није без губитка, реконструкција увек носи одређену грешку, па се проблем своди на

тражење што већег C уз задржавање задовољавајуће тачности.

За архитектуру мреже описане горе, укупан број параметара износи

$$N_{par} = (D - 1)W^2 + 2dW + (D + N_{ch})W + N_{ch}. \quad (3.4)$$

3.2.2 Избор параметара

Тренирање једне физички информисане неуралне мреже није једноставан процес и зависи од већег броја параметара. На резултат истовремено утичу архитектура мреже, избор фреквенција у енкодеру, стопа учења, број корака и тежина физичког дела функције губитка. Због тога ови параметри нису бирани независно, већ је тражена комбинација која даје стабилну конвергенцију у времену доступном за тренирање.

За Фуријеов енкодер коришћено је $N_f = 192$ фреквенција као што је описано у одељку (2.3). Фреквенције су подељене на нискофреквентни и високофреквентни део, при чему је однос броја високих и ниских фреквенција постављен на 0,85. Карактеристичне фреквенције одређене су из просторних фреквенција присутних у симулираним подацима. Као заједничка вредност коришћена је медијана добијених карактеристичних фреквенција.

Дубина мреже постављена је на $D = 8$, док је за ширину одабрано $W = 384$ неурона у сваком скривеном слоју. Ова ширина је одабрана на основу испитивања различитих фактора компресије у одељку (4.1), при чему је дала најбољи однос тачности реконструкције, компресије и времена потребног за тренирање.

Функција губитка подељена је на део који мери одступање од података $\mathcal{L}_d$ и физички део $\mathcal{L}_f$, као у једначини (2.9). Физички део чини један или више претходно описаних резидуала, у зависности од тога које једначине користимо у датом тренингу. Његов утицај контролисан је тежином $\mathcal{W}_f$, која није била иста у свим тренинзима. Како физички члан не би имао превелики утицај на самом почетку оптимизације, његова тежина је постепено повећавана током првих 150 епоха, након чега достиже задату вредност.

За оптимизацију је коришћен оптимизатор Adam са почетном стопом учења $lr = 2 \times 10^{-3}$ или $lr = 4 \times 10^{-3}$, у зависности од конкретног тренинга. Стопа учења је током тренирања смањивана коришћењем *CosineAnnealingLR* распореда. За формирање серија података (*batch*) и иницијализацију Фу-

ријеовог енкодера коришћена су случајна семена² 1234 и 42, редом. Сваки тренинг трајао је 400 епоха, док се број оптимизационих корака разликовао у зависности од конкретне поставке тренинга. При поређењу различитих поставки нисмо захтевали исти број корака, већ смо водили рачуна да функција губитка достигне приближно исти степен конвергенције.

Табела 3.2: Параметри коришћени за тренирање неуралне мреже.

Параметар	Ознака	Вредност
Број Фуријеових фреквенција	N_f	192
Однос високих и ниских фреквенција	—	0,85
Дубина мреже	D	8
Ширина скривених слојева	W	384
Активациона функција	—	tanh
Оптимизатор	—	Adam
Почетна стопа учења	lr	зависи од тренинга
Распоред стопе учења	—	<i>CosineAnnealingLR</i>
Укупан број епоха	—	400
Постепено увођење физике	—	150 епоха
Тежина физичког члана	$\mathcal{W}_f$	зависи од тренинга
Број оптимизационих корака	—	зависи од тренинга
Случајна семена	—	1234 и 42

3.3 Увођење физике

До сада описана мрежа тренирана је на потпуном скупу података, у свакој тачки домена биле су познате вредности свих физичких величина. Таква поставка добро одговара задатку компресије, али не одговара ситуацији у којој бисмо метод стварно применили. Из посматрања никада не располажемо потпуном тродимензионалном структуром атмосфере. Међутим, разумевање и саме компресије је изузетно важно зато што нам омогућава да конфигуришемо комплексност мреже и да боље разумемо колико је, сама по себи, атмосфера комплексна, тј. колико комплексан модел можемо реконструисати на основу података задате тачности.

У најбољем случају добијамо свега неколико слојева, и то слојева константне оптичке дубине τ , а не константне геометријске висине z . Овде се

²Случајно семе (*seed*) одређује начин генерисања случајних вредности и омогућава понављање тренинга под истим почетним условима.

тим ограничењем бавимо у поједностављеном облику. Задржавамо мрежу константног z , будући да су нам подаци дати на њој, али намерно уклањамо део слојева и посматрамо да ли мрежа може да реконструише цео домен из онога што је преостало³. Наравно, мрежа није способна да то уради, барем не сама. Између задржаних слојева мрежа нема никакву информацију и ту реконструише са сличном тачношћу као и линеарна интерполација између слојева.

Управо ту улази физика, тј. додатна ограничења на основу МХД једначина. Величине које мрежа предвиђа нису међусобно независне, већ су повезане једначинама које важе у свакој тачки домена, без обзира на то да ли у тој тачки располажемо податком.

3.3.1 Избор једначина за реконструкцију

Систем МХД једначина (1.1)–(1.4) не можемо употребити у целини. Неке од њих садрже временске изводе, а неке важе само приближно, па их уводимо постепено.

Соленоидалност магнетног поља

Једначина (3.5), произлази из саме структуре Максвелових једначина и важи у свакој тачки и у сваком тренутку. Не садржи временски извод, захтева само прве просторне изводе и не уводи никакве додатне претпоставке, због чега га уводимо први. Исти услов као физичко ограничење у неуралној мрежи користе и Jarolim et al. (2023) при екстраполацији короналног магнетног поља.

$$\nabla \cdot \vec{B} = \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0. \quad (3.5)$$

Функцију губитка за ову једначину можемо записати у облику:

$$\mathcal{L}_{\text{div}B} = \frac{1}{N_c} \sum_{i=1}^{N_c} \left[\frac{1}{S_{\text{div}B}} \left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} \right) \right]_{\mathbf{x}_i}^2, \quad (3.6)$$

где је $S_{\text{div}B}$ карактеристична скала дивергенције магнетног поља, коју узимамо као:

$$S_{\text{div}B} = \frac{\langle |\vec{B}| \rangle}{\Delta x}. \quad (3.7)$$

³Разлика између слојева константног τ и константног z овде се не разматра; она уводи додатну сложеност коју остављамо за даљи рад.

Овај услов је примењен на одређен број колокационих тачака N_c . Колокационе тачке представљају тачке у домену на којима се израчунава резидуал једначине (3.5).

Стационарна једначина континуитета

Ако у једначини одржања масе (1.1) занемаримо временски извод, добићемо стационарну једначину континуитета,

$$\nabla \cdot (\rho \vec{v}) = 0. \quad (3.8)$$

Такође и:

$$\mathcal{L}_{\text{cont}} = \frac{1}{N_c} \sum_{i=1}^{N_c} \left[\frac{\nabla \cdot (\rho \vec{v})}{S_{\text{cont}}} \right]_{\mathbf{x}_i}^2. \quad (3.9)$$

Где карактеристичну скалу за ову једначину дефинишемо као:

$$S_{\text{cont}} = \left\langle \left| \frac{\partial(\rho v_x)}{\partial x} \right| + \left| \frac{\partial(\rho v_y)}{\partial y} \right| + \left| \frac{\partial(\rho v_z)}{\partial z} \right| \right\rangle. \quad (3.10)$$

Овако записана једначина је применљива на један тренутак, али важи само приближно, па из тог разлога приликом рачуна за укупну функцију губитка (2.9) морамо да поведемо рачуна коју тежину дајемо овом члану (3.9).

Стационарна једначина количине кретања

Полазимо од једначине одржања количине кретања (1.2), која је записана у конзервативном облику. За потребе физичког ограничења магнетни члан записујемо преко Лоренцове силе, користећи идентитет

$$\nabla \cdot \left(\vec{B} \vec{B} - \frac{1}{2} \vec{I} B^2 \right) = (\nabla \times \vec{B}) \times \vec{B}, \quad (3.11)$$

који важи за соленоидално магнетно поље, $\nabla \cdot \vec{B} = 0$. Уз $f_{v_A} = 1$ и занемаривање временског извода у (1.2), добијамо стационарни облик једначине кретања

$$\nabla \cdot (\rho \vec{v} \vec{v}) + \nabla p + \rho g_{\odot} \vec{e}_z - \frac{1}{4\pi} (\nabla \times \vec{B}) \times \vec{B} = 0, \quad (3.12)$$

где смо искористили $\vec{g} = -g_{\odot} \vec{e}_z$. Ова једначина описује стационарну равнотежу инерцијалног, притисног, гравитационог и магнетног члана. У даљем тексту називамо је стационарном MHD једначином кретања.

Ради једноставнијег записа, појединачне чланове означавамо као

$$A = \nabla \cdot (\rho \vec{v} \vec{v}), \quad G = \nabla p + \rho g_{\odot} \vec{e}_z, \quad L = \frac{1}{4\pi} (\nabla \times \vec{B}) \times \vec{B}. \quad (3.13)$$

Тада резидуал стационарне једначине количине кретања гласи

$$R_{\text{mom}} = A + G - L. \quad (3.14)$$

Као и код претходних једначина, резидуал се израчунава у колокационим тачкама, при чему се потребни просторни изводи добијају директно из излаза неуралне мреже. Пошто једначина има три просторне компоненте, одговарајућу функцију губитка можемо записати у облику

$$\mathcal{L}_{\text{mom}} = \frac{1}{3N_c} \sum_{i=1}^{N_c} \sum_{j \in \{x,y,z\}} \left[\frac{R_{\text{mom},j}(\mathbf{x}_i)}{S_{\text{mom},j}(z_i)} \right]^2, \quad (3.15)$$

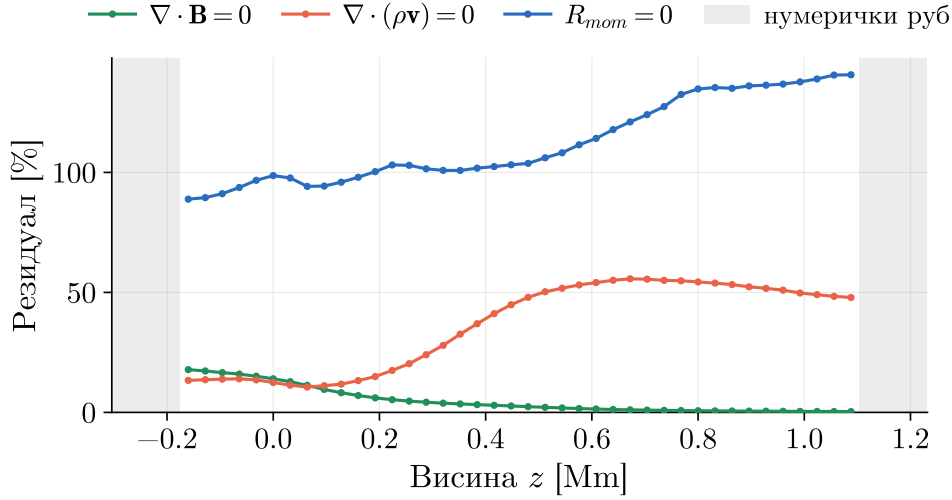
где $S_{\text{mom},j}(z_i)$ представља карактеристичну скалу силе за одговарајућу компоненту на висини z_i . Скала се рачуна засебно за сваки хоризонтални слој, чиме се спречава да слојеви са већим апсолутним вредностима физичких чланова доминирају функцијом губитка. Дефинисана је као

$$S_{\text{mom},j}(z) = \max \left(\langle |A_j| \rangle_{xy,z}, \langle |G_j| \rangle_{xy,z}, \langle |L_j| \rangle_{xy,z} \right), \quad j \in \{x, y, z\}, \quad (3.16)$$

где $\langle \cdot \rangle_{xy,z}$ означава усредњавање по хоризонталној равни на висини z . Оваква нормализација омогућава да различите компоненте резидуала и различити висински слојеви имају упоредив утицај на оптимизацију.

Задовољеност физичких једначина на симулираним подацима

Пре увођења физичких чланова у функцију губитка проверавамо у којој мери симулирани подаци задовољавају одабране једначине. Степен у коме су једначине задовољене на самим подацима одређује колико строго њихове резидуале можемо наметнути током тренирања и утиче на избор тежина одговарајућих физичких чланова.



Слика 3.3: Нормализовани резидуали једначина из (3.3.1) на симулираним подацима у зависности од висине.

Из приказаних резултата (Слика 3.3) видимо да је услов соленоидалности магнетног поља добро задовољен на свим висинама. Одступање је највеће у најнижим слојевима, где износи око 18%, а затим се са висином постепено смањује до вредности испод 1%. Због тога услов $\nabla \cdot \vec{B} = 0$ можемо применити као физичко ограничење на целом висинском домену са тежином $\mathcal{W}_{divB} = 1$.

Стационарна једначина континуитета најбоље је задовољена у области до приближно $z = 0,3$ Mm, где је одступање мање од 25%. Изнад ове висине одступање расте и достиже вредности веће од 50%. Због тога ову једначину можемо применити са пуном тежином једнаком 1 само у области испод 0,3 Mm. Друга могућност је да је применимо на целом висинском домену, али са смањеном тежином од 0,3, како њен резидуал у вишим слојевима не би имао превелики утицај на тренирање.

За разлику од претходне две једначине, једначина кретања, односно једначина одржања импулса, показује велико одступање на целом висинском домену. Њен резидуал у приказаним слојевима износи од приближно 89% до 141% и са висином углавном расте. Због тога ова једначина није погодна за увођење као јако физичко ограничење, јер би такво ограничење приморавало мрежу да задовољи једначину коју ни сами симулирани подаци не задовољавају довољно добро.

На основу задовољености једначина на симулираним подацима формирамо

четри поставке тренинга, приказаних у табели (3.3). У сваком оптимизационом кораку користимо исти скуп од $N_c = 8000$ насумично одабраних колокационих тачака из целог домена. Према томе, наведени број представља укупан број колокационих тачака по кораку, а не број тачака по слоју или појединачној једначини.

Табела 3.3: Поставке тренинга са физичким ограничењима.

Тренинг	Физички чланови	Област примене	N_c	Тежине
1	$\nabla \cdot \vec{B} = 0$	цео домен	8000	$\mathcal{W}_{\text{div}B} = 1$
2	$\nabla \cdot (\rho \vec{v}) = 0$	цео домен	8000	$\mathcal{W}_{\text{cont}} = 0,3$
3	$\nabla \cdot \vec{B} = 0,$ $\nabla \cdot (\rho \vec{v}) = 0$	цео домен	8000	$\mathcal{W}_{\text{div}B} = 0,7,$ $\mathcal{W}_{\text{cont}} = 0,3$
4	стационарна једначина момента	цео домен	8000	$\mathcal{W}_{\text{mom}} = 1$

У последњој поставци стационарну једначину кретања уводимо са тежином $\mathcal{W}_{\text{MHS}} = 1$ као посебан тест утицаја снажног физичког ограничења на реконструкцију, који је слабо задовољен на подацима.

Глава 4

Резултати

У овом поглављу представљамо резултате тренирања неуралних мрежа на задацима реконструкције атмосферских величина из симулираних података. Тачност реконструкције анализирамо помоћу релативне грешке NRMSE_σ ¹ дефинисане као:

$$\text{NRMSE}_\sigma = \frac{\sqrt{\frac{1}{N_{x,y}} \sum_{i=1}^{N_{x,y}} [\hat{Y}_i(z) - Y_i(z)]^2}}{\sigma_Y(z)} \times 100\%. \quad (4.1)$$

Овде су $Y_i(z)$ симулиране вредности посматране величине у слоју z , $\hat{Y}_i(z)$ одговарајуће предикције мреже, $N_{x,y}$ број тачака у слоју, а $\sigma_Y(z)$ стандардна девијација симулираних вредности у том слоју. Ову грешку приказујемо у зависности од висине, односно од положаја слоја дуж координате z . Поред тога, предикције мреже поредимо са симулираним подацима помоћу дводимензионалних мапа у xy и xz пресецима. Мапе компоненти v_y и B_y изостављамо из приказа, јер показују сличне карактеристике као одговарајуће компоненте v_x и B_x и не пружају додатне информације значајне за анализу.

Анализу започињемо тренирањем на целокупном скупу података у одељку (4.1), где испитујемо како сложеност мреже и фактор компресије утичу на квалитет реконструкције. Након тога постепено уклањамо слојеве доступне током тренирања и анализирамо способност мреже да реконструише недостајуће слојеве (4.2). На крају разматрамо утицај додавања физичких једначина у функцију губитка (3.3).

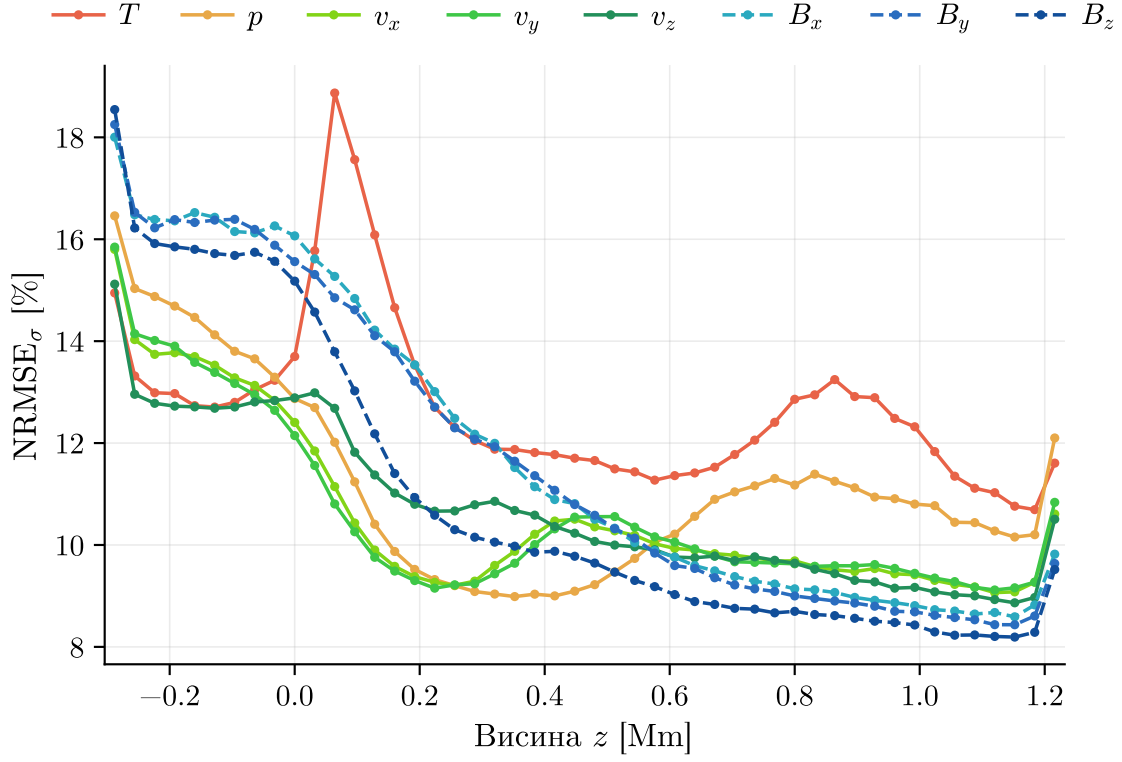
¹ NRMSE_σ (*normalized root mean square error*) означава корен средње квадратне грешке нормализован стандардном девијацијом симулираних вредности.

4.1 Учење над целокупним подацима

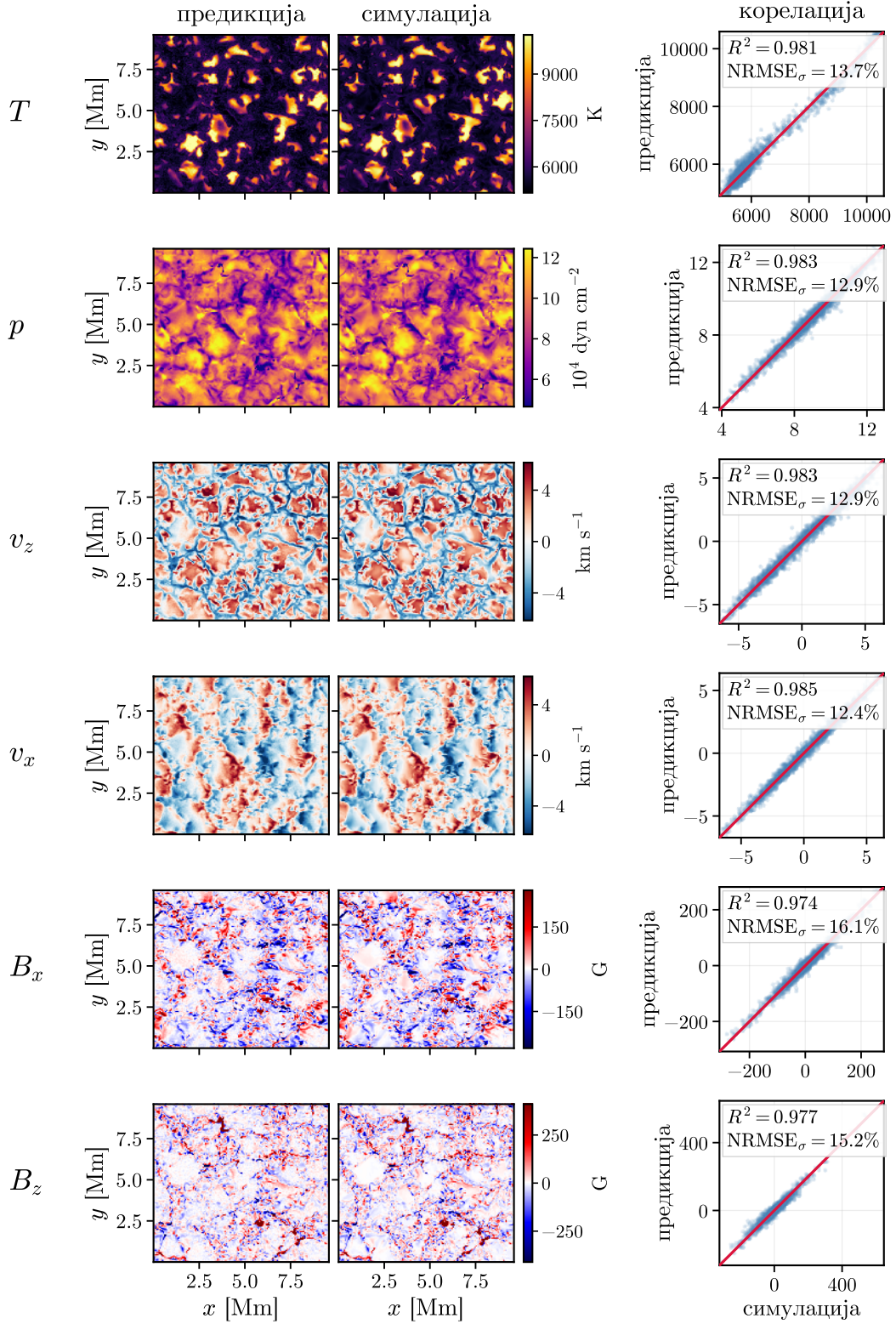
Овде представљамо три одвојена примера тренирања неуралне мреже на свим доступним подацима. То подразумева да се мреже тренирају на домену $300 \times 300 \times 48$ и имају на располагању вредности физичких параметара у свим тачкама домена. Мреже разматране у овом поглављу се искључиво разликују по томе што је број параметара мреже различит и самим тим компресија је другачија. Дакле, три наведена примера садрже само различит број неурона у скривеним слојевима.

Компресија $C \approx 15$

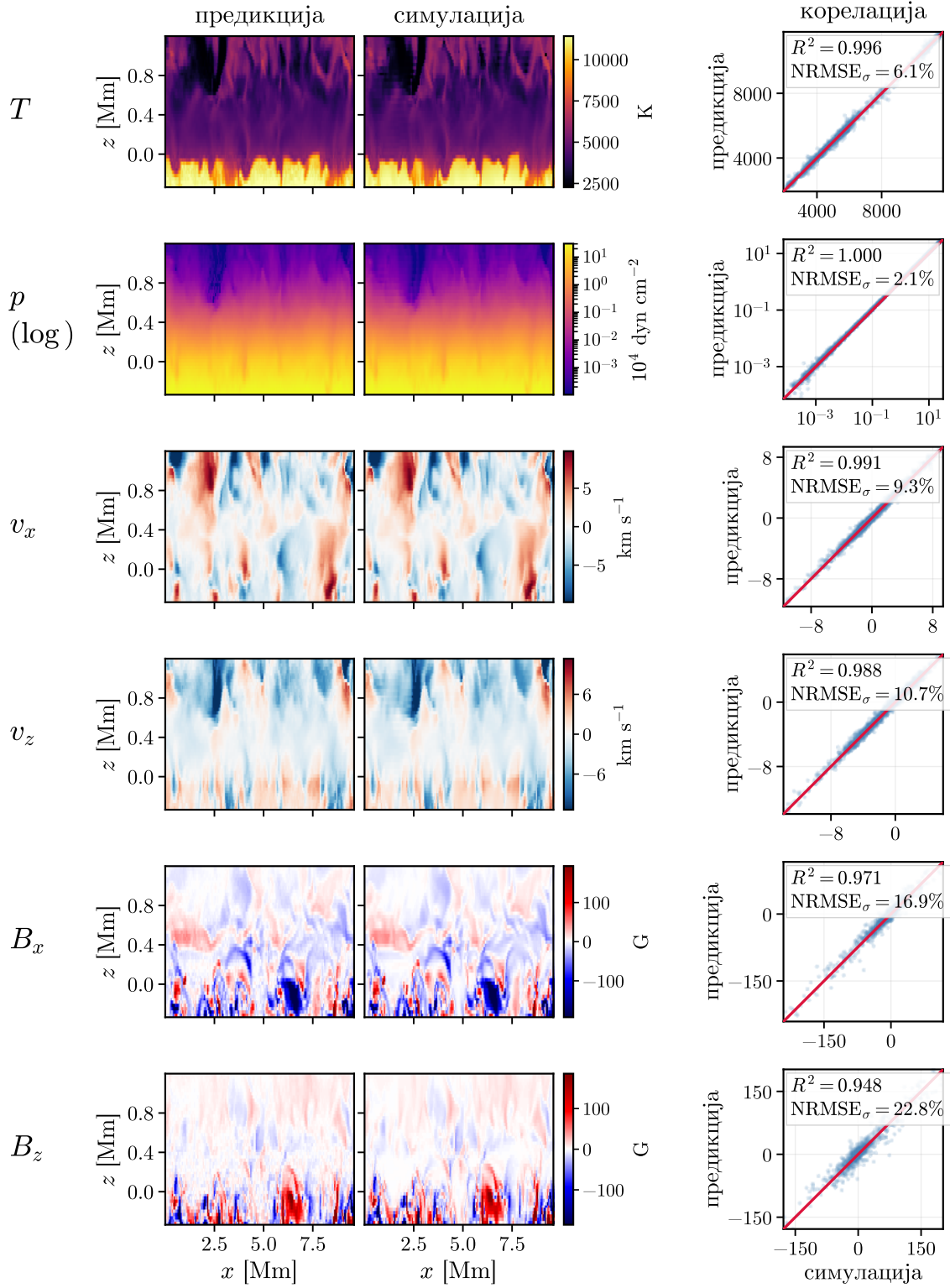
У овом примеру имамо 512 неурона у сваком скривеном слоју, који доводе до компресије од $C = 15.43$, користећи релацију (3.3). Ова компресија релативно добро репродукује све физичке величине кроз домен. Можемо да запазимо пораст грешке за температуру око десетог слоја (Слика 4.1). То је одступање које можемо запазити на сваком примеру у нашем раду. Овај образац се објашњава из структуре симулиране атмосфере. На тим слојевима долази до преласка из доње фотосфере у горњу и средња температура нагло пада са око 10,000 К на 5500 К. Мрежи је тешко да покупи ову наглу промену у структури података. Грешка не потиче само од великог температурног градијента, већ и од тога што су структуре у xu равни значајно другачије на овом прелазу. И даље, вертикална структура атмосфере коју мрежа успева да научи се добро слаже са оригиналом, што се може видети са слике (4.3).



Слика 4.1: Зависност грешке NRMSE_σ од висине за мрежу ширине $W = 512$ и фактор компресије $C = 15,43$.



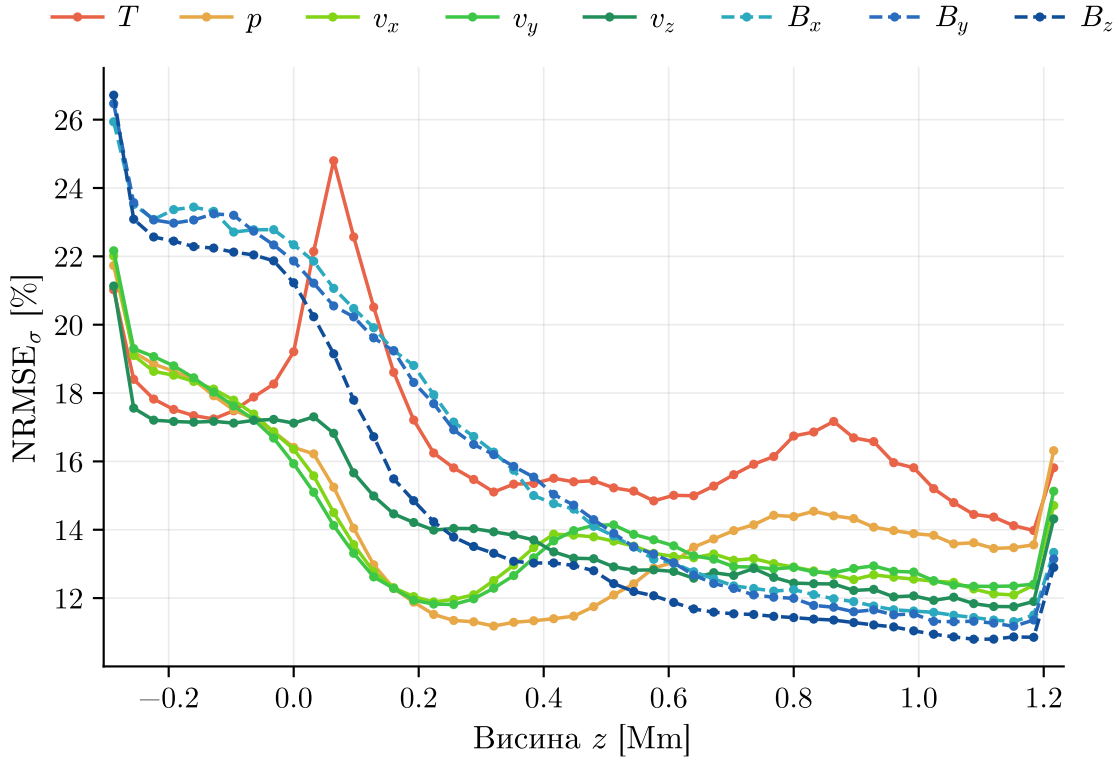
Слика 4.2: Поређење предикција мреже и симулираних података у xy пресеку на висини $z = 0$ km за мрежу ширине $W = 512$. У трећој колони приказани су дијаграми корелације, Пирсонов коефицијент корелације и вредност NRMSE $_{\sigma}$ на датом слоју.



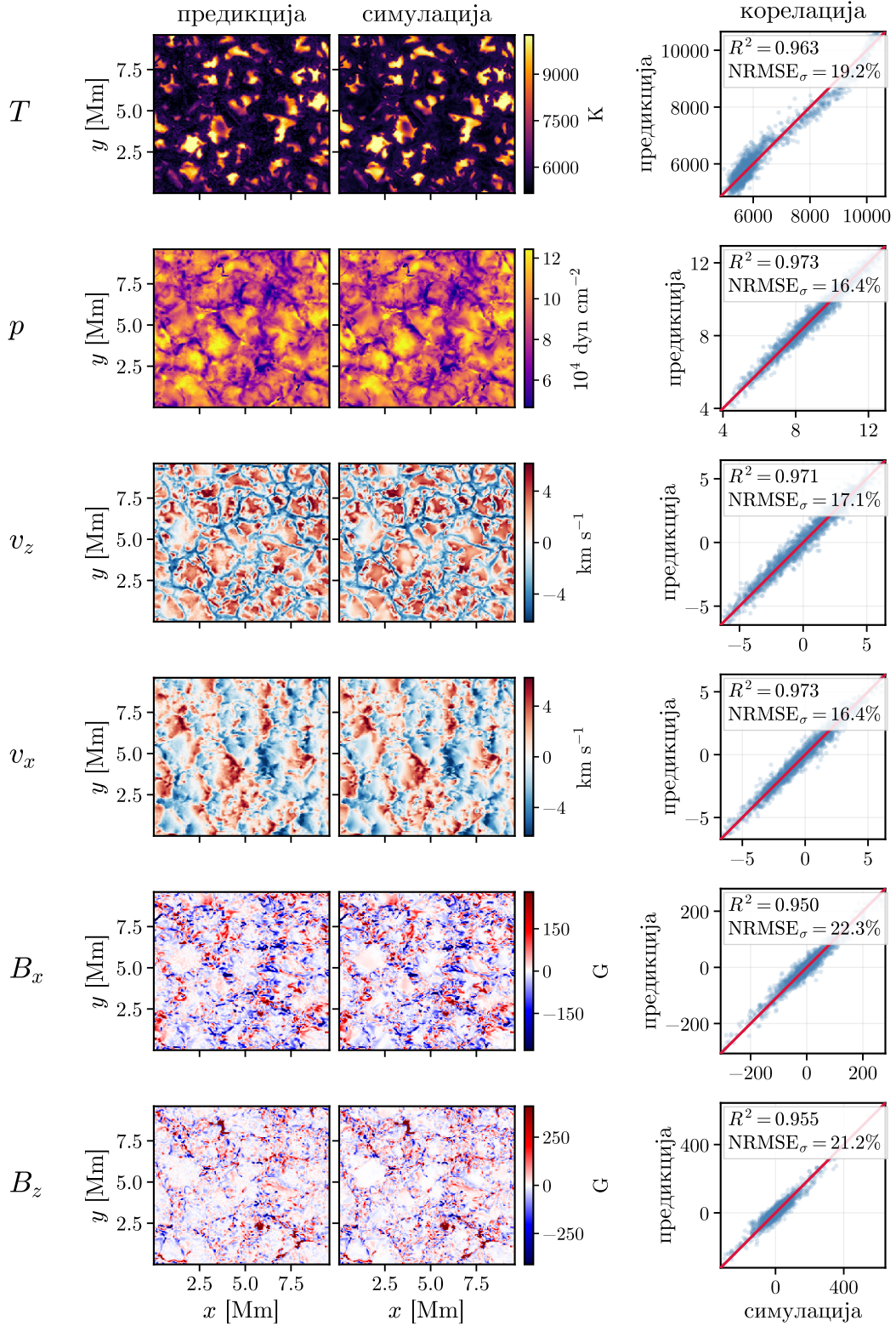
Слика 4.3: Поређење предикција мреже и симулираних података у xz пресеку на средишњем y слоју за мрежу ширине $W = 512$. У трећој колони приказани су дијаграми корелације.

Компресија $C \approx 25$

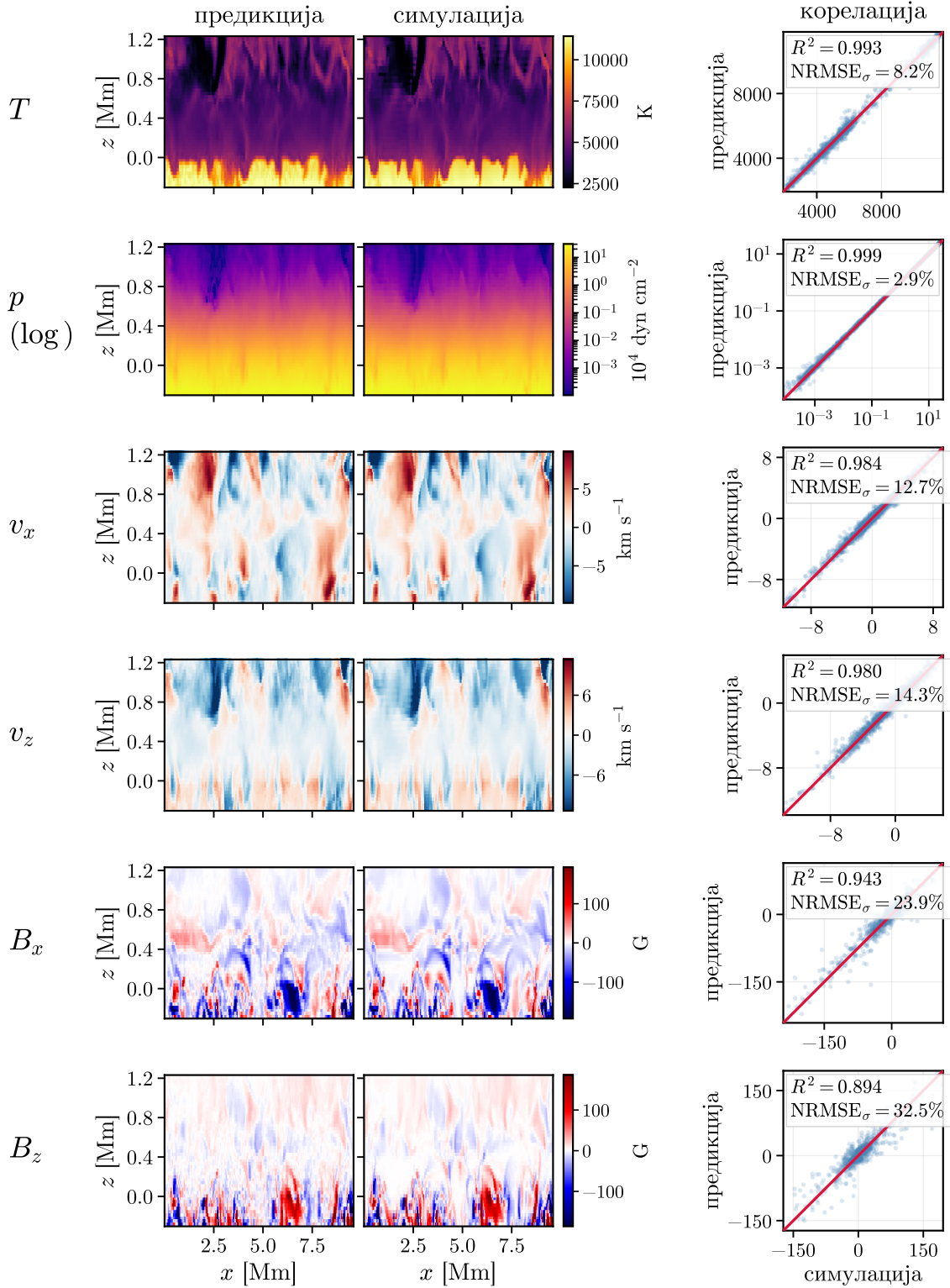
Број неурона у скривеним слојевима сада је $W = 384$, што даје фактор компресије $C = 25,88$. Приказујемо исте графике као у претходном примеру, како бисмо непосредно упоредили квалитет реконструкције при различитим факторима компресије.



Слика 4.4: Зависност грешке NRMSE_σ од висине за мрежу ширине $W = 384$ и фактор компресије $C = 25,88$.



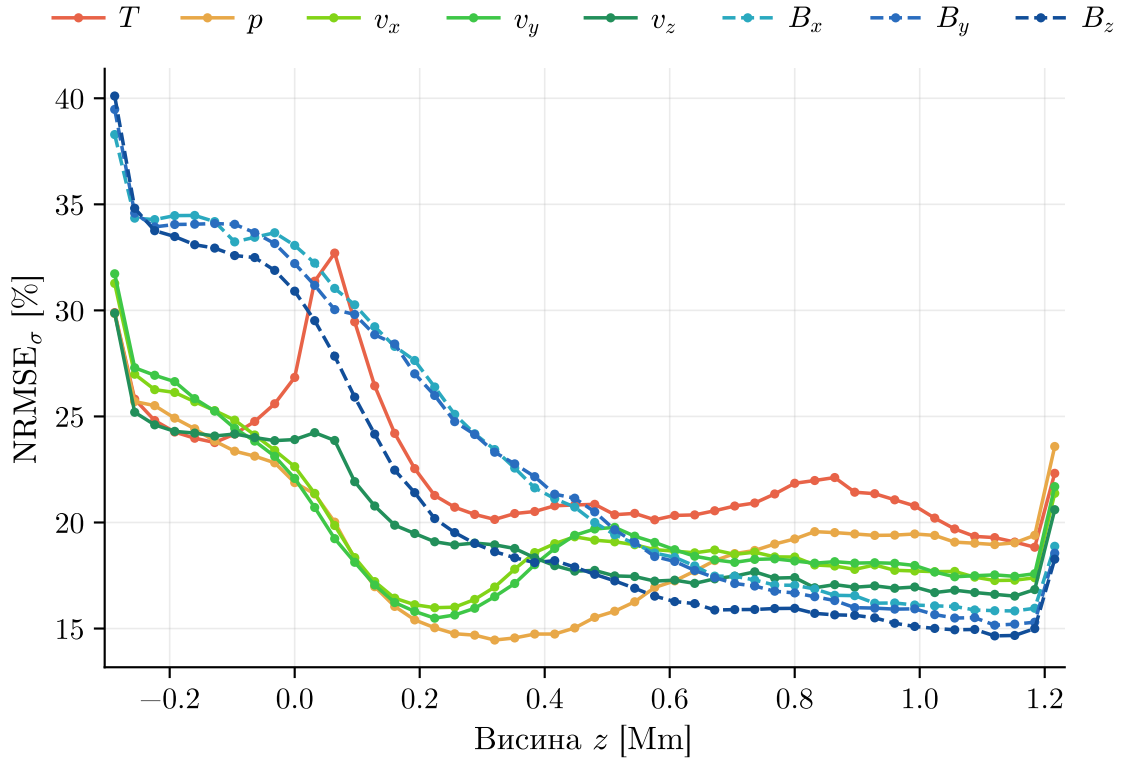
Слика 4.5: Поређење предикција мреже и симулираних података у xy пресеку на висини $z = 0$ km за мрежу ширине $W = 384$. У трећој колони приказани су дијаграми корелације, Пирсонов коефицијент корелације и вредност NRMSE $_{\sigma}$ на датом слоју.



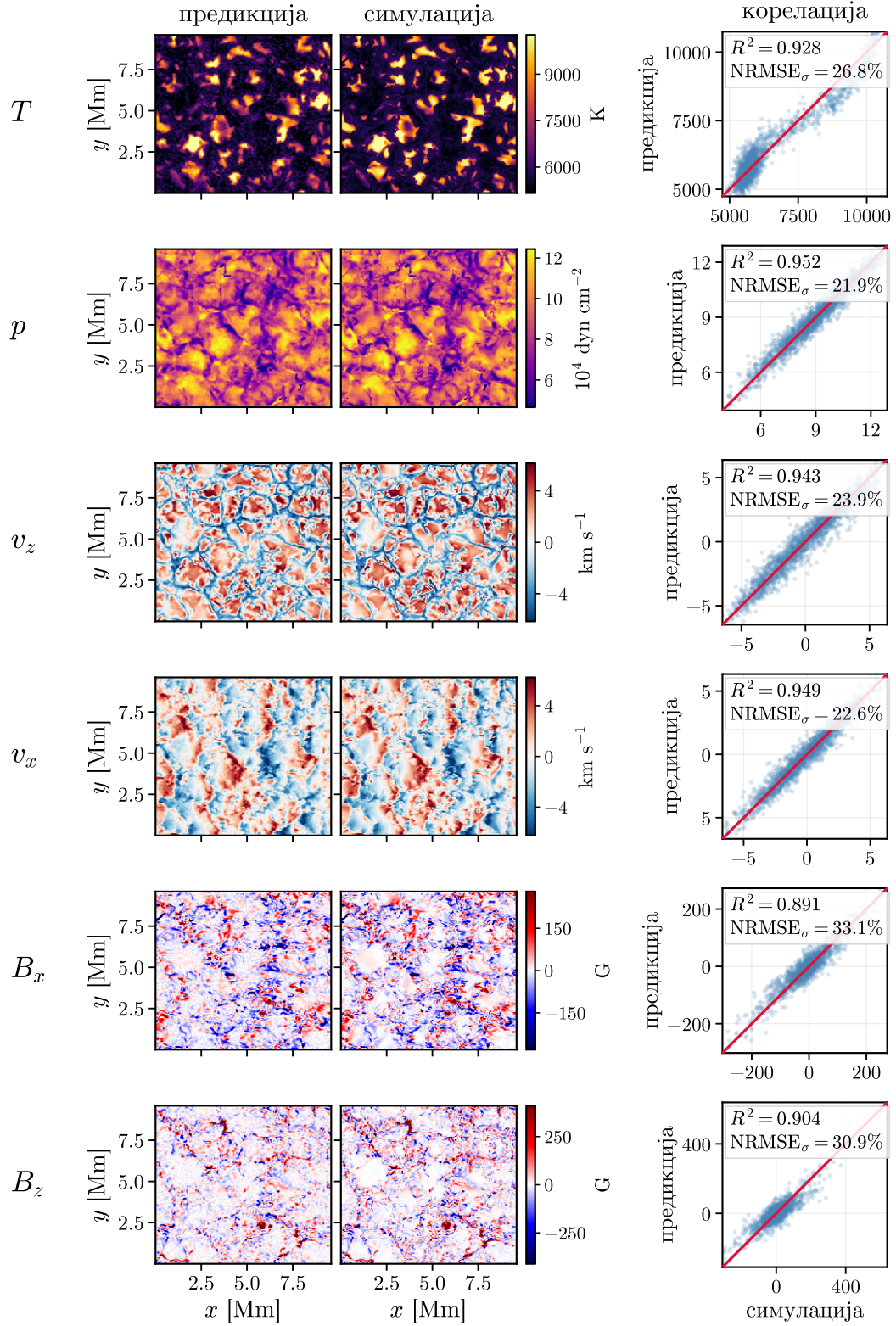
Слика 4.6: Поређење предикција мреже и симулираних података у xz пресеку на средишњем y слоју за мрежу ширине $W = 384$. У трећој колони приказани су дијаграми корелације.

Компресија $C \approx 52$

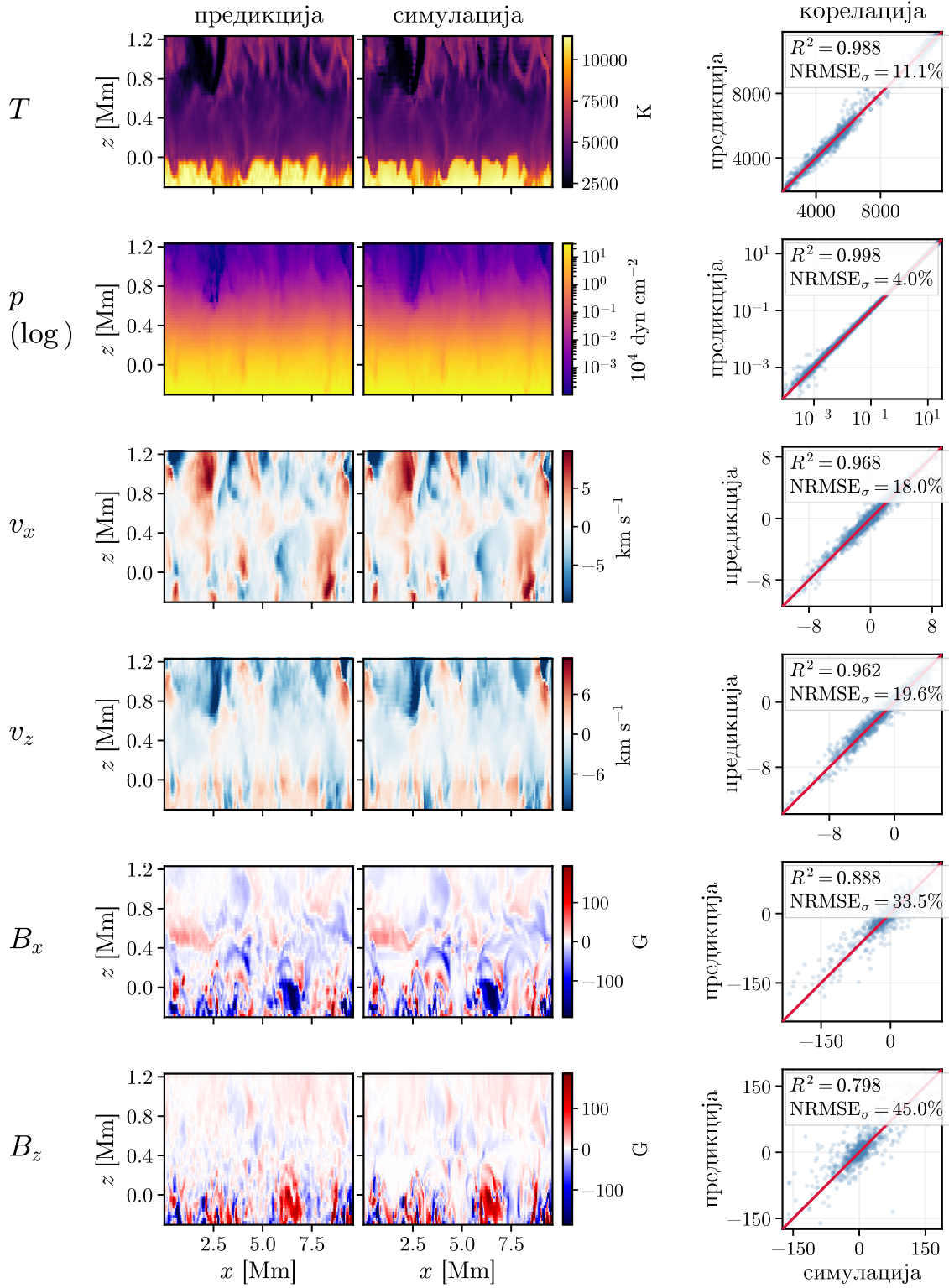
Преласком са $W = 384$ на $W = 256$ број неурона смањује се за исти износ као у претходном кораку. Међутим, пошто број параметара мреже приближно зависи од квадрата ширине, фактор компресије сада достиже $C = 52,28$. Овако висока компресија оставља мрежи знатно мањи број параметара за репродукцију просторне структуре свих физичких величина.



Слика 4.7: Зависност грешке NRMSE_σ од висине за мрежу ширине $W = 256$ и фактор компресије $C = 52,28$.



Слика 4.8: Поређење предикција мреже и симулираних података у xy пресеку на висини $z = 0$ km за мрежу ширине $W = 256$. У трећој колони приказани су дијаграми корелације, Пирсонов коефицијент корелације и вредност NRMSE $_{\sigma}$ на датом слоју.



Слика 4.9: Поређење предикција мреже и симулираних података у xz пресеку на средишњем y слоју за мрежу ширине $W = 256$. У трећој колони приказани су дијаграми корелације.

Избор ширине мреже

На основу приказаних тестова за наставак рада бирамо мрежу ширине $W = 384$. Ова мрежа представља најповољнији однос квалитета реконструкције, фактора компресије и времена потребног за тренирање. Мрежа ширине $W = 512$ даје мању грешку, али има мањи фактор компресије и захтева више времена за тренирање, док мрежа ширине $W = 256$ не располаже довољним бројем параметара за једнако добру реконструкцију свих просторних структура.

Важно је нагласити да у овом делу физичке једначине још увек нису укључене у функцију губитка, већ се мрежа тренира искључиво на доступним подацима. Ипак, овај корак је неопходан јер нам омогућава да проценимо сложеност мреже потребну за реконструкцију атмосфере и одаберемо архитектуру која уз разуман број параметара задржава задовољавајућу тачност.

Такође уочавамо систематско одступање у реконструкцији температуре код мрежа ширине $W = 256$ и $W = 384$, које се на дијаграмима корелације испољава у облику слова „S”. Ово одступање знатно је израженије код мреже ширине $W = 256$. Сматрамо да је оно последица већ поменутог оштрог температурног градијента у близини фотосфере. Поред тога, прелазак са претежно конвективног на радијативни пренос енергије праћен је променом просторне структуре температуре у xu равни, што мрежи додатно отежава њену реконструкцију.

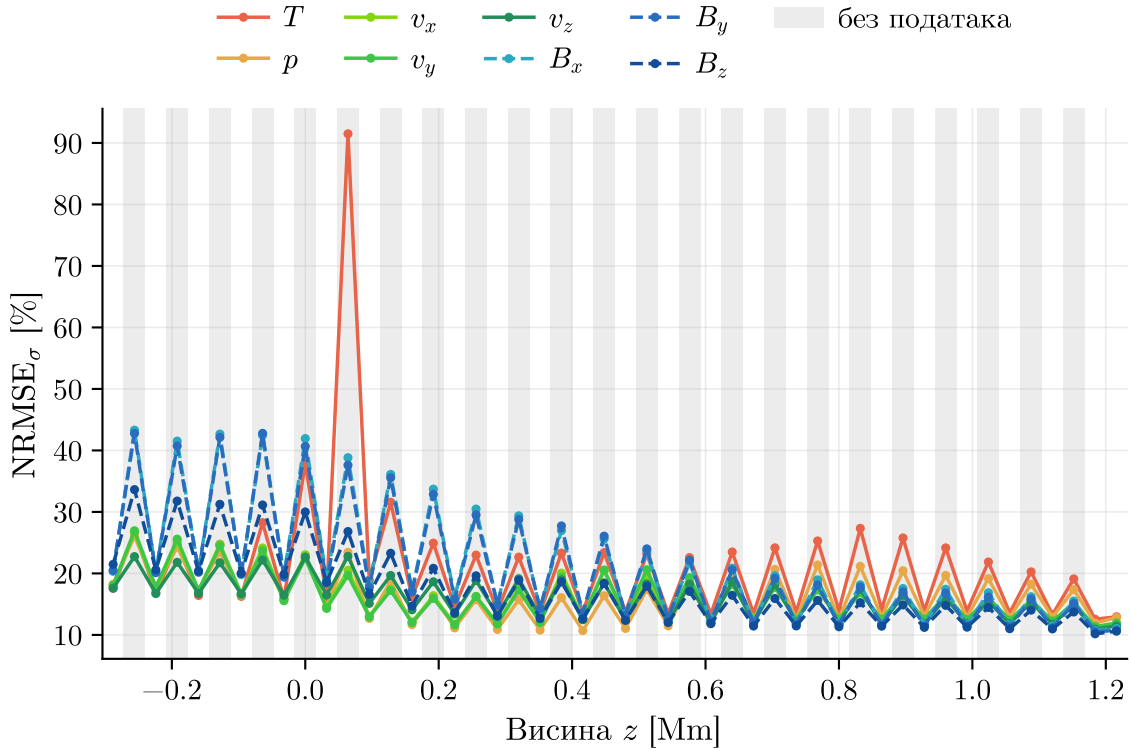
4.2 Уклањање слојева

Сада приступамо проблему уклањања слојева. Циљ овог одељка јесте да испитамо колико добро мрежа одабрана у одељку (4.1) може да интерполира вредности између слојева доступних током тренирања. Разматрамо конфигурације са $N_{\text{skip}} = 1, 5$ и 7 , где N_{skip} представља број узастопних уклањених слојева између два доступна слоја. На крају разматрамо и нехомогено уклањање, при коме слојеве задржавамо гушће у областима са израженијим вертикалним градијентима. Овакав избор има и посматрачку подршку, будући да највише информација о различитим висинама добијамо управо у фотосфери.

4.2.1 Хомогено уклањање слојева

Избацивање сваког другог слоја ($N_{\text{skip}} = 1$)

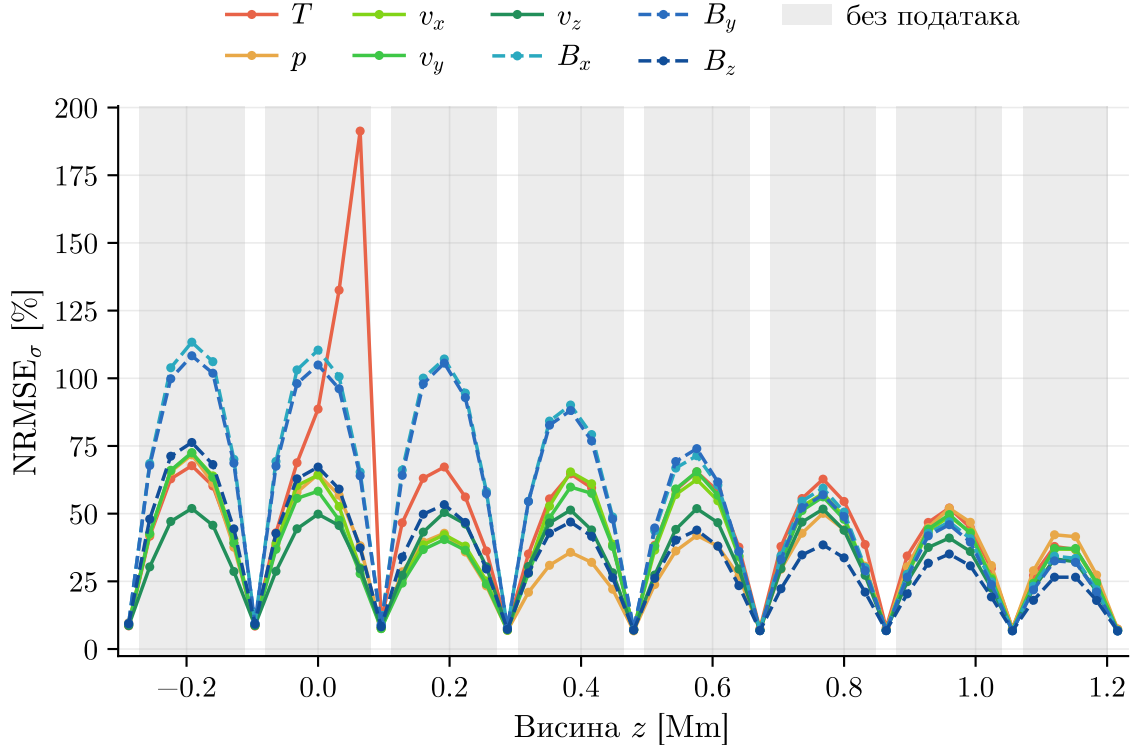
У овој конфигурацији уклањамо сваки други слој по висини, тако да је између свака два слоја доступна током тренирања један слој који мрежа треба да реконструише.



Слика 4.10: Зависност грешке NRMSE_σ од висине при уклањању сваког другог слоја.

Избацивање пет слојева ($N_{\text{skip}} = 5$)

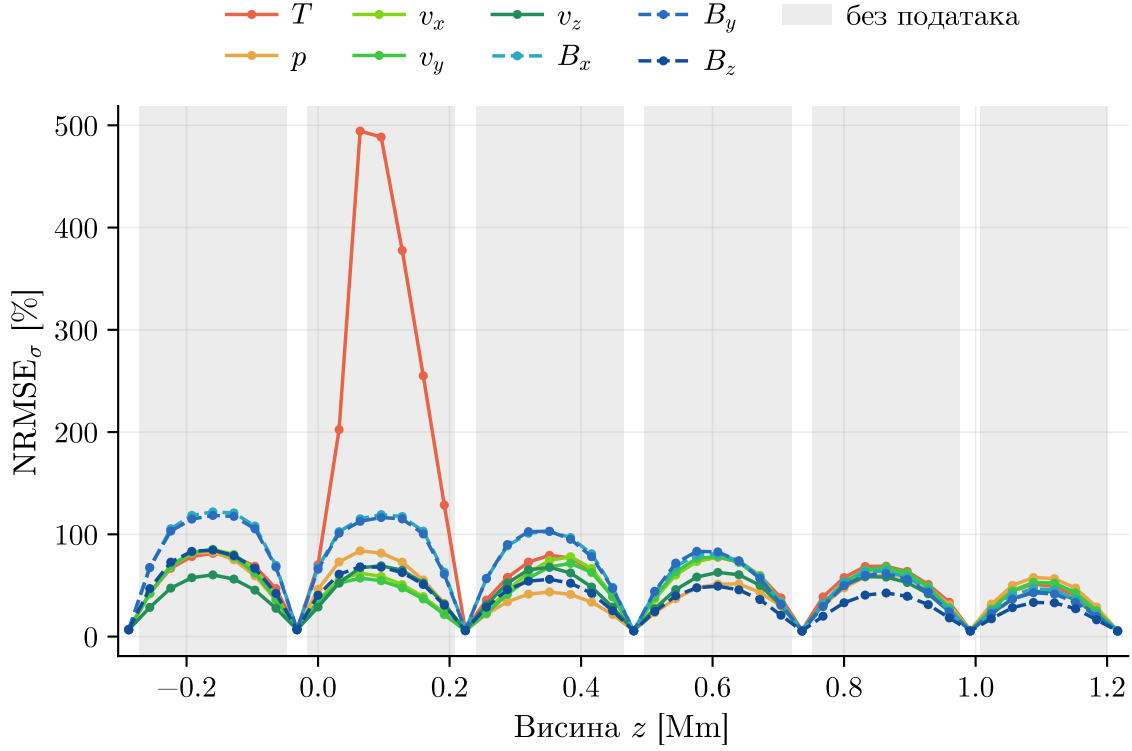
У овој конфигурацији између два слоја доступна током тренирања уклањамо пет узастопних слојева.



Слика 4.11: Зависност грешке NRMSE_σ од висине при уклањању $N_{\text{skip}} = 5$ узастопних слојева.

Избацивање седам слојева ($N_{\text{skip}} = 7$)

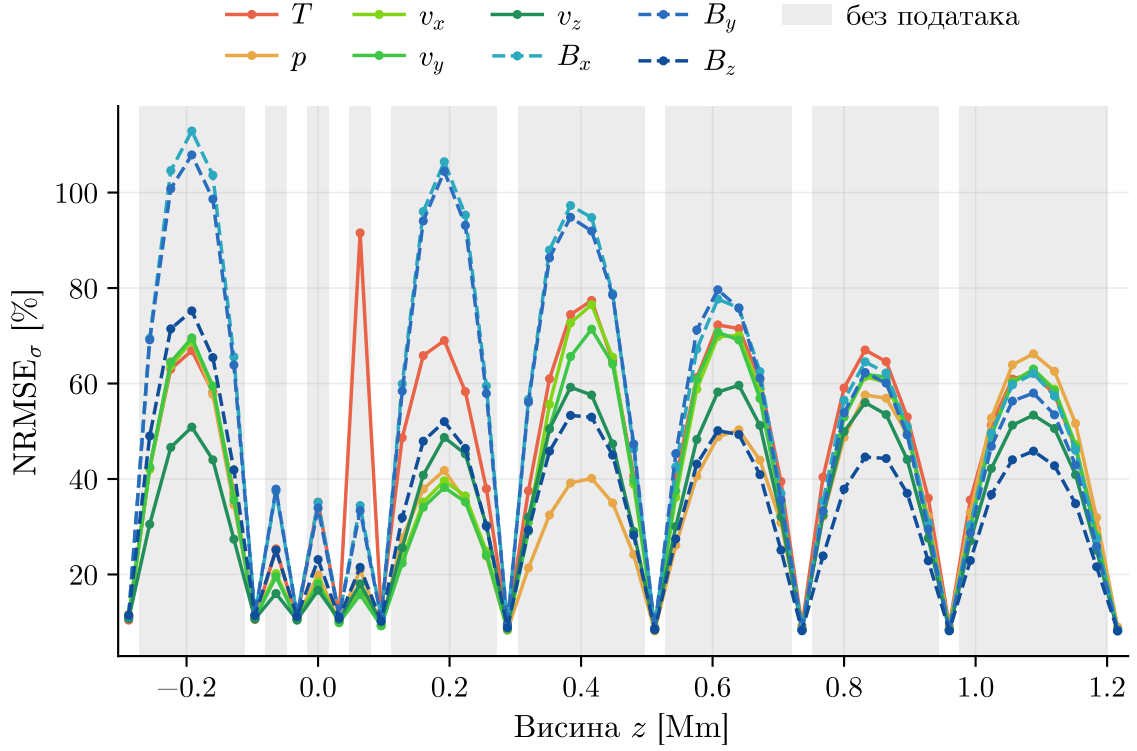
У овој конфигурацији између два слоја доступна током тренирања уклањамо седам узастопних слојева. Ово је најређе хомогено узорковање атмосфере које разматрамо.



Слика 4.12: Зависност грешке NRMSE_σ од висине при уклањању $N_{\text{skip}} = 7$ узастопних слојева.

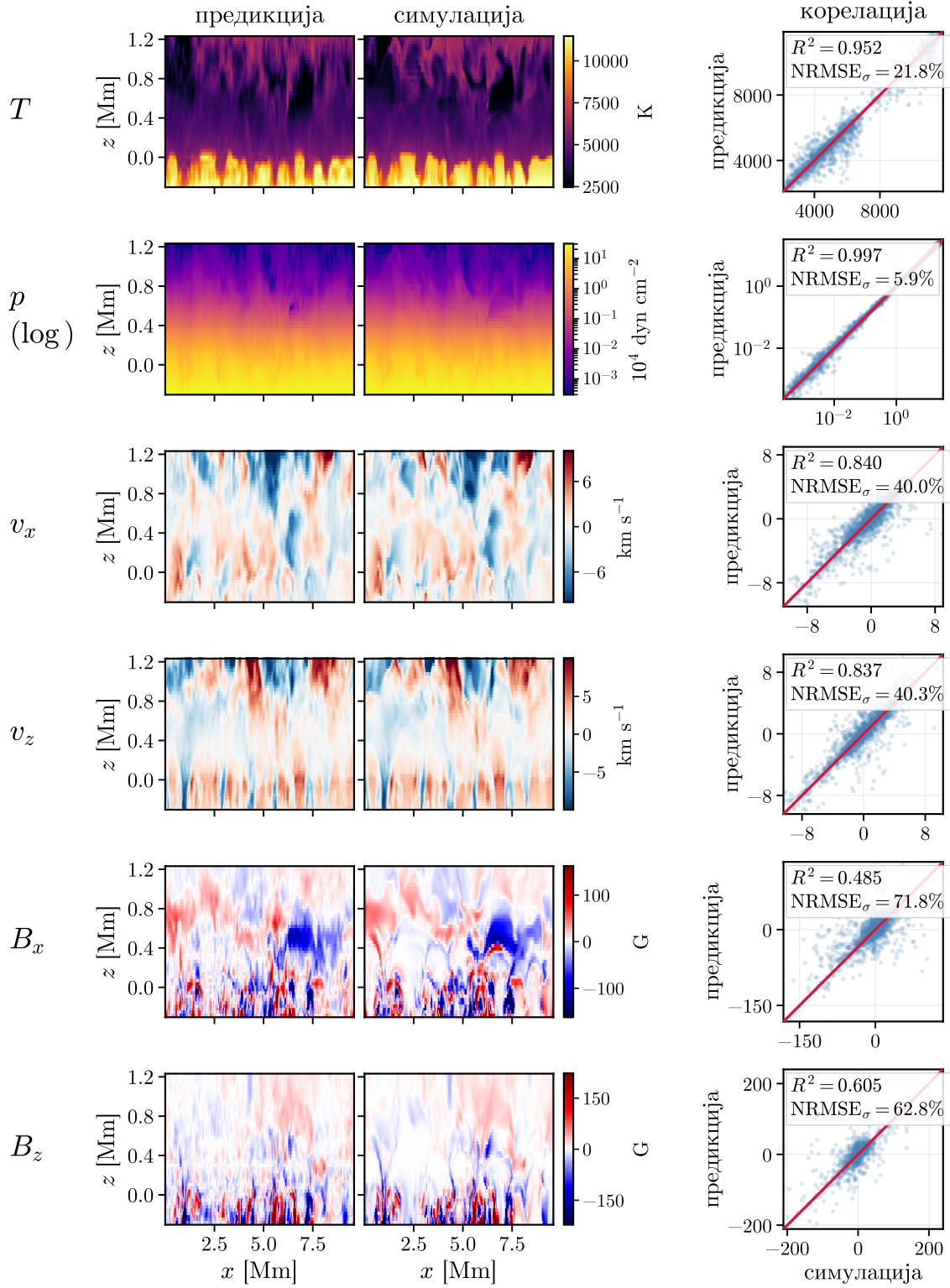
4.2.2 Нехомогено уклањање слојева

Код нехомогеног уклањања размак између доступних слојева није исти кроз цео домен. Слојеве задржавамо гушће у близини фотосфере, где су вертикални градијенти најизраженији, док у областима у којима се физичке величине спорије мењају остављамо већи размак између доступних слојева.



Слика 4.13: Зависност грешке NRMSE_σ од висине при нехомогеном уклањању слојева.

Ову поставку користићемо као референтну за све наредне физички информисане тренинге, при чему ћемо испитати колико физичка ограничења могу да побољшају реконструкцију слојева за које мрежа нема податке.



Слика 4.14: Поређење предикција мреже и симулираних података у xz пре-секу на средишњем y слоју при нехомогеном уклањању слојева.

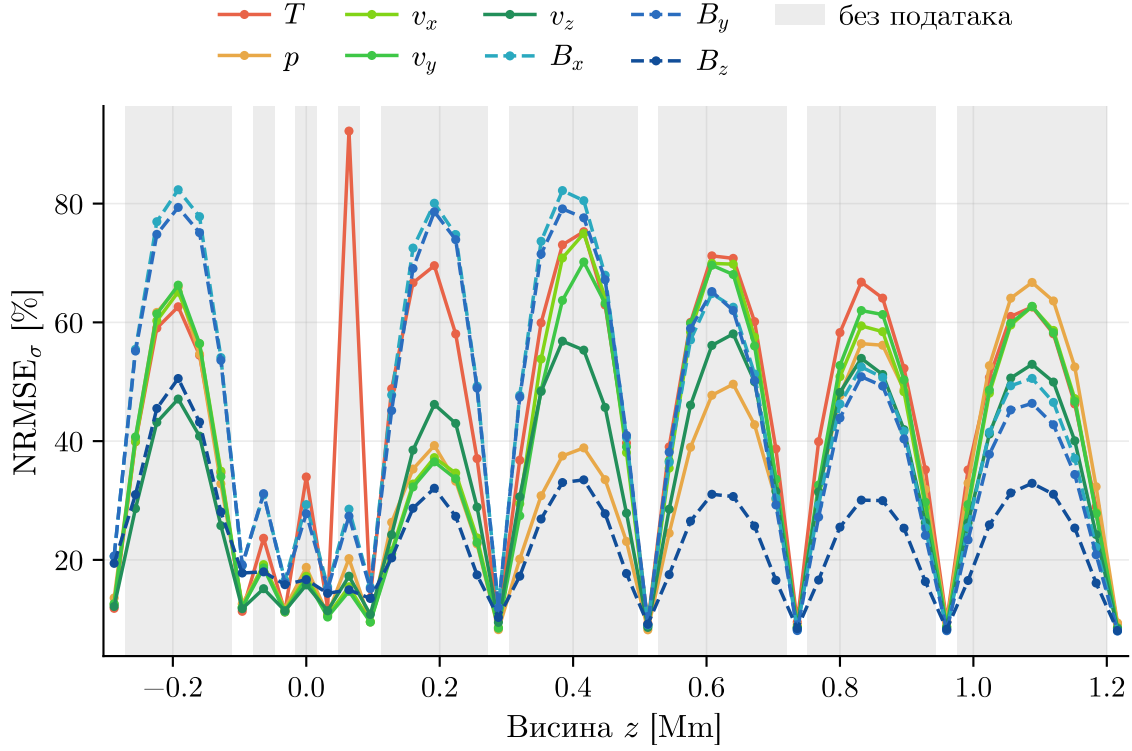
4.3 Додавање физичких једначина

У одељку (4.2) видели смо да је мрежа способна да интерполира између доступних слојева, али уз значајно повећање грешке у уклоњеним слојевима. Сада у функцију губитка уводимо диференцијалне једначине описане у одељку (3.3). Резултате поредимо са мрежом тренираном на истом скупу слојева, али без физичких ограничења, како бисмо издвојили њихов утицај на реконструкцију. Прво разматрамо услов соленоидалности магнетног поља, а затим стационарну једначину континуитета и њихову комбинацију.

У наредним табелама израз „без физике” означава референтни модел без физичких чланова у функцији губитка, док „са физиком” означава модел са наведеним физичким ограничењем. Побољшање је израчунато у односу на референтни модел, па позитивна вредност означава смањење грешке, а негативна њено повећање.

Додавање услова $\nabla \cdot \vec{B} = 0$

Функцији губитка додајемо услов соленоидалности магнетног поља, описан једначином (3.6), са тежином $\mathcal{W}_{\text{div}B} = 1$. На овај начин од мреже, поред слагања са доступним подацима, захтевамо и да реконструисано магнетно поље има што мању дивергенцију.



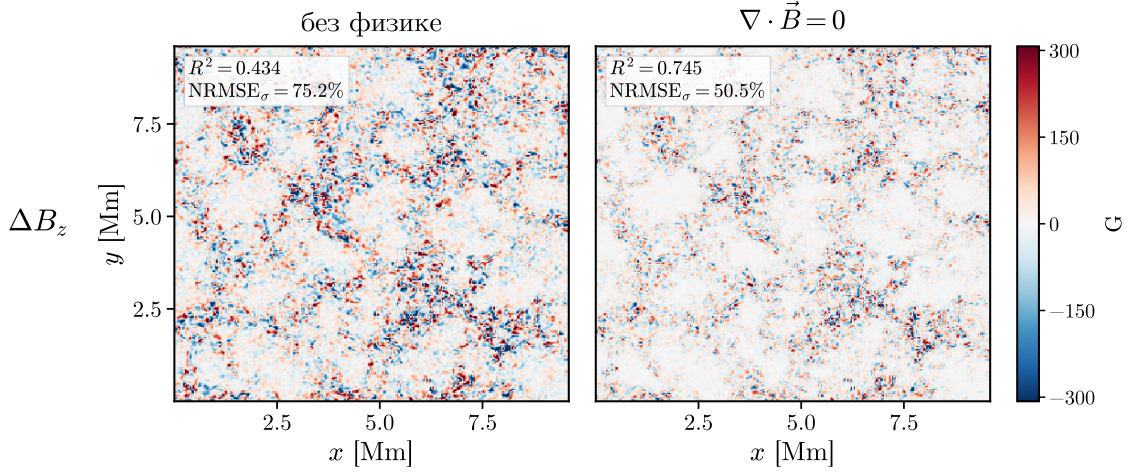
Слика 4.15: Зависност грешке NRMSE_σ од висине за мрежу тренирану уз услов $\nabla \cdot \vec{B} = 0$.

Табела 4.1: Поређење вредности NRMSE_σ за референтни модел без физике (А) и модел са условом $\nabla \cdot \vec{B} = 0$ (В).

Величина	Слојеви са подацима			Слојеви без података		
	А [%]	В [%]	Поб. [%]	А [%]	В [%]	Поб. [%]
T	10.07	10.76	-6.82	54.31	53.35	+1.77
p	9.63	10.26	-6.60	41.25	40.13	+2.73
v_x	9.42	9.96	-5.82	46.79	45.38	+3.01
v_y	9.40	9.92	-5.51	46.16	45.09	+2.32
v_z	9.49	10.10	-6.40	40.78	39.15	+4.00
B_x	10.00	13.25	-32.53	64.16	51.93	+19.06
B_y	9.96	13.13	-31.82	62.57	50.37	+19.49
B_z	9.69	12.52	-29.27	40.29	26.34	+34.62
просек	9.71	11.24	-15.78	49.54	43.97	+11.25

У слојевима без података услов соленоидалности смањује просечну грешку

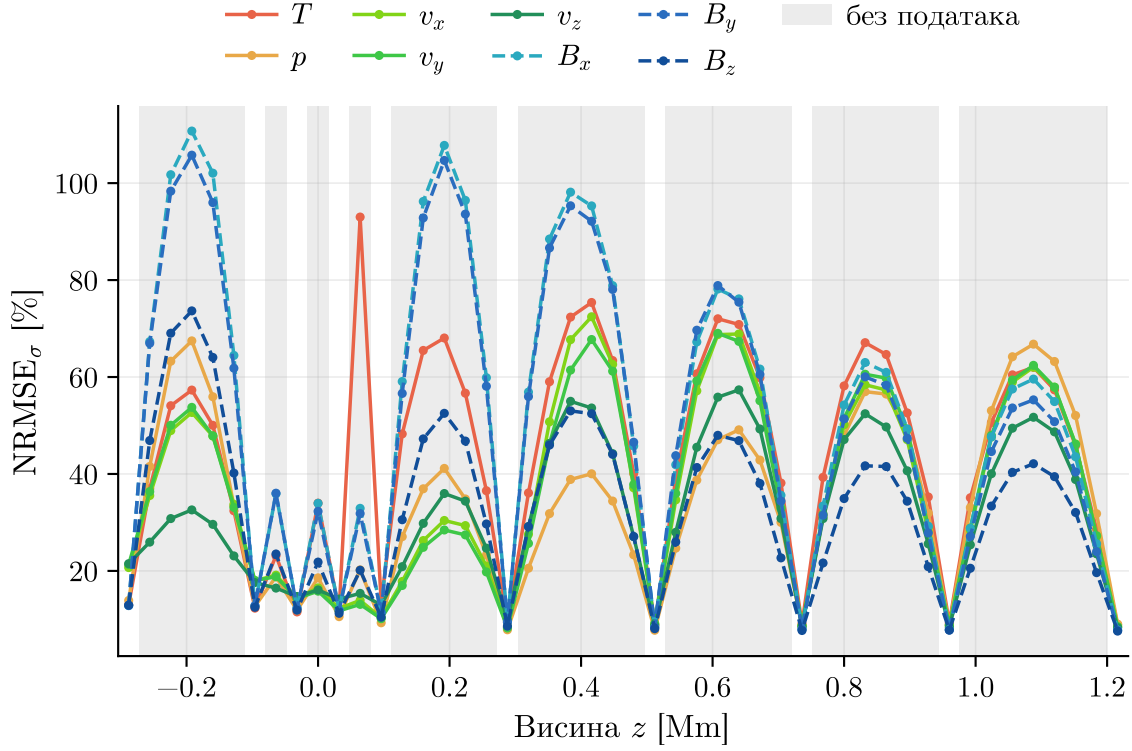
за 11,25%, при чему је највеће побољшање остварено за компоненте магнетног поља. Истовремено се грешка у слојевима са подацима повећава, што указује на компромис између слагања са подацима и задовољавања физичког ограничења.



Слика 4.16: Разлика вредности B_z наспрам симулиране, на висини од $-0,19$ Мм. Поређење за случај са и без физички информисаном функцијом губитка.

Додавање услова $\nabla \cdot (\rho \vec{v}) = 0$

У овом примеру функцији губитка додајемо стационарну једначину континуитета, описану једначином (3.8). Једначину примењујемо на целом висинском домену са тежином $\mathcal{W}_{\text{cont}} = 0,3$. Овим условом повезујемо реконструисана поља густине и брзине и захтевамо да дивергенција масеног флукса буде што ближа нули.



Слика 4.17: Зависност грешке NRMSE_σ од висине за мрежу тренирану уз услов $\nabla \cdot (\rho \vec{v}) = 0$.

Табела 4.2: Поређење вредности NRMSE_σ за референтни модел без физике (А) и модел са условом $\nabla \cdot (\rho \vec{v}) = 0$ (В).

Величина	Слојеви са подацима			Слојеви без података		
	А [%]	В [%]	Поб. [%]	А [%]	В [%]	Поб. [%]
T	10.07	10.79	-7.16	54.31	52.63	+3.10
p	9.63	10.04	-4.27	41.25	40.58	+1.64
v_x	9.42	11.91	-26.50	46.79	42.93	+8.26
v_y	9.40	11.90	-26.60	46.16	42.57	+7.78
v_z	9.49	12.46	-31.22	40.78	36.36	+10.85
B_x	10.00	10.29	-2.93	64.16	63.16	+1.55
B_y	9.96	10.24	-2.81	62.57	61.17	+2.23
B_z	9.69	9.88	-2.02	40.29	38.50	+4.43
просек	9.71	10.94	-12.70	49.54	47.24	+4.65

У слојевима без података стационарна једначина континуитета смањује

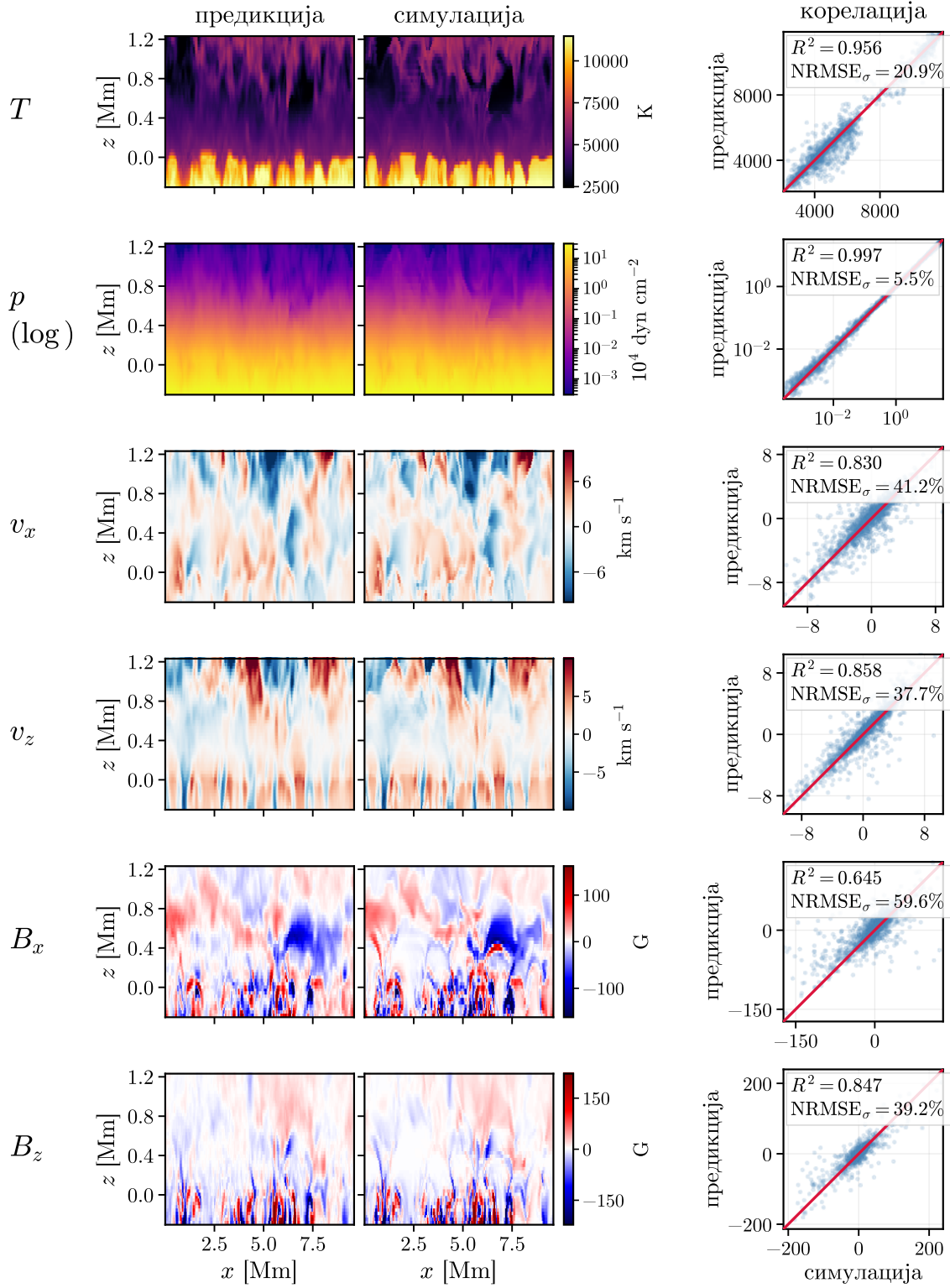
просечну грешку за 4,65%. Највеће побољшање добијено је за компоненте брзине, које ова једначина непосредно повезује са густином, док се у слојевима са подацима њихова грешка повећава.

Комбиновање услова $\nabla \cdot \vec{B} = 0$ и $\nabla \cdot (\rho \vec{v}) = 0$

У овој поставци оба услова примењујемо истовремено, са тежинама $\mathcal{W}_{\text{div}B} = 1$ и $\mathcal{W}_{\text{cont}} = 0,3$. На тај начин могуће је истовремено побољшати реконструкцију магнетног поља и брзине у слојевима без података. Просечно побољшање у овим слојевима износи 12,96%, а највећи добитак поново се јавља код компоненти магнетног поља и брзине. Мање побољшање уочава се и код температуре и притиска, иако ове величине нису непосредно повезане условом соленидалности.

Табела 4.3: Поређење вредности NRMSE_σ за референтни модел без физике (A) и модел са оба физичка услова (B).

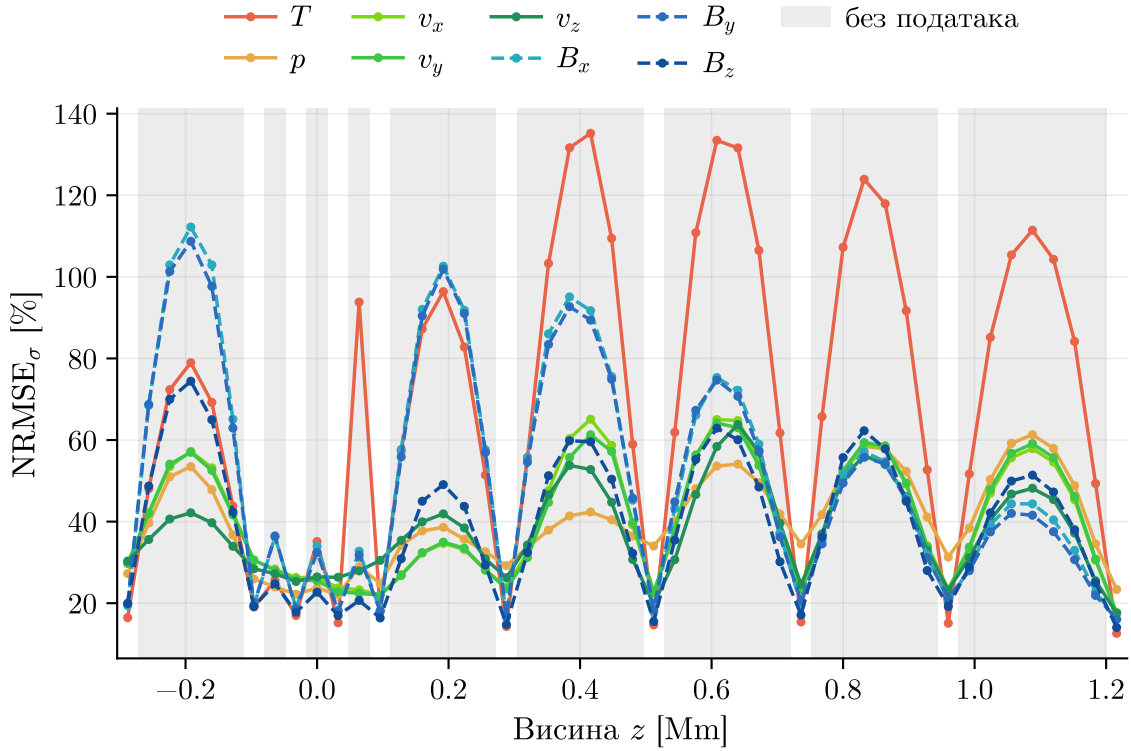
Величина	Слојеви са подацима			Слојеви без података		
	A [%]	B [%]	Поб. [%]	A [%]	B [%]	Поб. [%]
T	10.07	11.62	-15.40	54.31	52.75	+2.88
p	9.63	10.67	-10.81	41.25	40.01	+3.00
v_x	9.42	12.53	-33.11	46.79	42.77	+8.59
v_y	9.40	12.52	-33.23	46.16	42.57	+7.76
v_z	9.49	13.12	-38.15	40.78	36.38	+10.80
B_x	10.00	13.04	-30.44	64.16	52.87	+17.59
B_y	9.96	12.95	-30.03	62.57	50.94	+18.58
B_z	9.69	12.24	-26.41	40.29	26.65	+33.85
просек	9.71	12.34	-27.10	49.54	43.12	+12.96



Слика 4.18: Поређење предикција мреже и симулираних података у xz пре-секу на средишњем y слоју. У трећој колони приказани су дијаграми корелације.

Додавање стационарне једначине количине кретања

Провера резидуала на симулираним подацима показала је да стационарна једначина количине кретања није довољно добро задовољена. Ипак, увели смо је са тежином $\mathcal{W}_{\text{mom}} = 1$ као контролни тест утицаја снажног, али недо-вољно оправданог физичког ограничења. Резултати показују да овако поста-вљен услов повећава грешку и у слојевима са подацима и у слојевима без њих. Мања тежина могла би да ублажи овај ефекат, али би њен избор захтевао додатно испитивање.

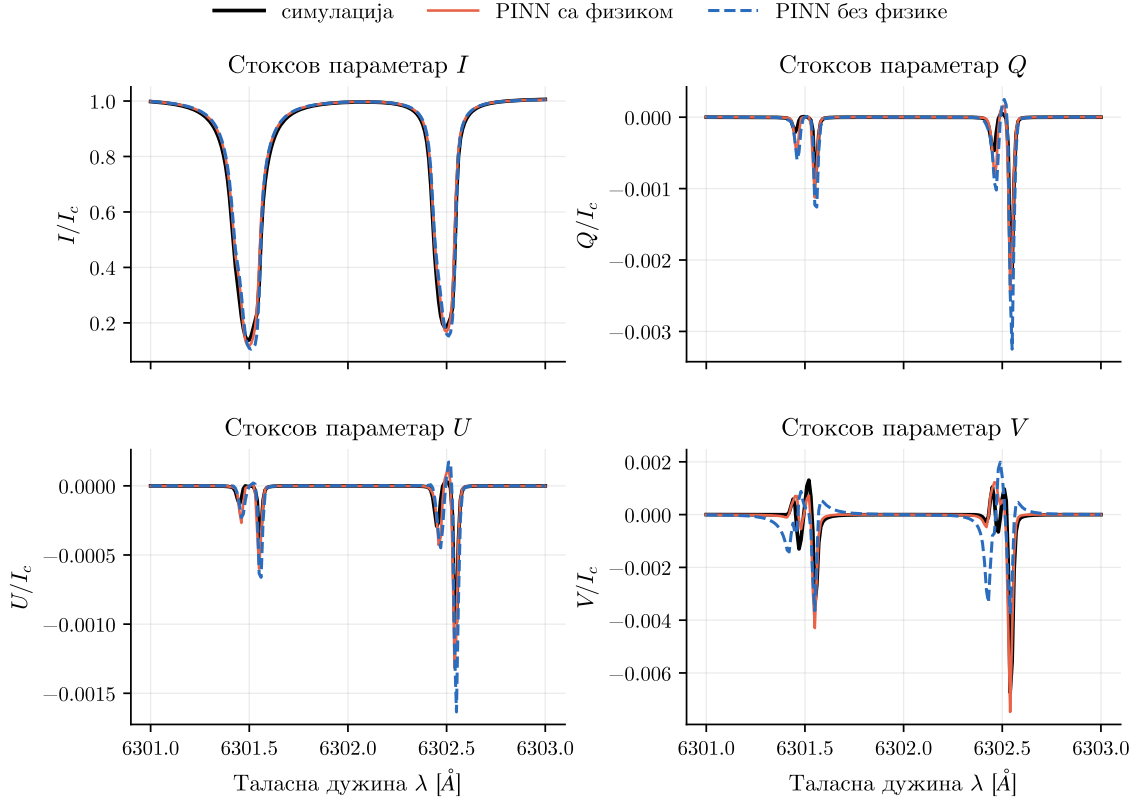


Слика 4.19: Зависност грешке $\text{NRMSE}\sigma$ од висине за мрежу тренирану уз стационарну једначину количине кретања са тежином $\mathcal{W}_{\text{MHS}} = 1$.

4.3.1 Поређење синтетисаних спектра

За додатну проверу реконструисаних атмосфера синтетисали смо профиле Стоксових параметара помоћу програмског оквира *Lightweaver* (Osborne & Milić, 2021). Поредимо спектре добијене из симулиране атмосфере, рекон-

струкције без физичких ограничења и реконструкције са комбинованим физичким условима.



Слика 4.20: Поређење синтетисаних профила Стоксових параметара за симулирану атмосферу и реконструкције са физичким ограничењима и без њих.

На слици (4.20) уочава се да профили добијени из реконструкције са физичким ограничењима показују боље слагање са симулираним профилима за сва четири Стоксова параметра. Овај резултат показује да побољшање реконструисаних атмосферских величина доводи и до верније синтезе спектралних профила, који су непосредно упоредиви са посматрањима.

Глава 5

Закључак и даљи рад

Машинско учење отвара низ могућности за реконструкцију Сунчеве атмосфере, а у овом раду испитујемо да ли се физичке једначине могу искористити да надоместе податке који недостају. Реч је о провери концепта на симулираним подацима, а не о решењу спремном за примену на посматрања.

Најпре је показано да неуронска мрежа са енкодером Фуријеових карактеристика може да реконструише ову врсту података, при чему је испитан однос између броја параметара мреже и тачности реконструкције. Сложенија мрежа боље репродукује структуре у подацима, али добитак у тачности опада са повећањем броја параметара, док време тренирања расте. Зато смо изабрали конфигурацију која представља најбољи компромис између ових захтева и која за наше податке одговара фактору компресије од 25.

Након избора одговарајуће конфигурације, поједини слојеви константне висине уклањани су из скупа за тренирање како бисмо испитали да ли физичка ограничења могу да побољшају реконструкцију недоступних слојева. Испитана су хомогена уклањања са $N_{\text{skip}} = 1, 5$ и 7 , као и нехомогено уклањање, при коме се слојеви задржавају гушће у фотосфери, а ређе изнад и испод ње. Нехомогена конфигурација изабрана је као референтна за све физички информисане тренинге, будући да одговара и посматрачкој ситуацији, у којој највише информација о различитим висинама добијамо управо из фотосфере.

Увођењем услова $\nabla \cdot \vec{B} = 0$ постигнуто је просечно побољшање од 11,3% у слојевима без података, при чему је грешка реконструкције компоненти магнетног поља смањена за 19 до 35%. Стационарна једначина континуитета

дала је просечно побољшање од 4,7%, са највећим добитком код компоненти брзине, до 10,9% за вертикалну компоненту. Иако ова једначина на симулираним подацима није задовољена у истој мери као услов соленоидалности, показала се корисном управо у слојевима без података. Комбиновањем оба услова постигнуто је највеће просечно побољшање, од 13,0% у слојевима без података.

Заједничко за ове случајеве јесте да је побољшање у слојевима без података праћено извесним погоршањем у слојевима у којима подаци постоје. Ово није нужно недостатак методе, већ последица тога што решење представља компромис између слагања са подацима и задовољавања физичког модела. Исто понашање описано је и при екстраполацији короналног магнетног поља (Jarolim et al., 2023). Значајно је и то што свака једначина највише поправља управо оне величине које експлицитно повезује, што указује на то да мрежа користи структуру наметнутог ограничења, а не само ефекат додатне регуларизације. Побољшање реконструисане атмосфере одразило се и на боље слагање синтетисаних профила Стоксових параметара са профилима добијеним из симулације.

Испитана је и стационарна једначина количине кретања, која на нашим подацима није задовољена у мери која би је чинила поузданим ограничењем. Њено увођење са тежином $\mathcal{W}_{\text{mom}} = 1$ повећало је грешку, па је ова поставка задржана првенствено као контролни експеримент.

Даљи рад води у неколико праваца.

1. **Слојеви константне оптичке дубине.** У овом раду слојеви су уклањани по константној геометријској висини, будући да су подаци тако задати. Инверзије спектрополариметријских посматрања, међутим, најчешће дају атмосферске величине на скали оптичке дубине τ , чији слојеви не одговарају равнима константног z . Прелазак на такву поставку неопходан је корак ка примени методе на стварна посматрања.
2. **Једначине са временским изводима.** У овом раду разматрани су услов соленоидалности, који не садржи временски извод, и стационарни облици осталих једначина. Пуна индукциона једначина и једначина континуитета садрже временски извод, па захтевају да мрежа обухвати и временску димензију, односно да истовремено учи више узастопних

снимака. Овај корак би могао додатно да побољша реконструкцију, али захтева знатно већу количину података и рачунарског времена, што тренутно превазилази расположиве ресурсе.

Библиографија

- Asensio Ramos, A., Cheung, M. C. M., Chifu, I., & Gafeira, R. 2023, *Living Reviews in Solar Physics*, 20, 4, doi: 10.1007/s41116-023-00038-x
- Baydin, A. G., Pearlmutter, B. A., Radul, A. A., & Siskind, J. M. 2018, *Journal of Machine Learning Research*, 18, 1
- Bishop, C. M. 2006, *Pattern Recognition and Machine Learning* (New York: Springer)
- Bodenheimer, P., Laughlin, G. P., Różyczka, M., & Yorke, H. W. 2007, *Numerical Methods in Astrophysics: An Introduction* (New York: Taylor & Francis)
- Foukal, P. V. 2004, *Solar Astrophysics*, 2nd edn. (Weinheim: Wiley-VCH)
- Glorot, X., & Bengio, Y. 2010, in *Proceedings of Machine Learning Research*, Vol. 9, *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 249–256
- Hastie, T., Tibshirani, R., & Friedman, J. 2009, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd edn. (New York: Springer), doi: 10.1007/978-0-387-84858-7
- Jarolim, R., Thalmann, J. K., Veronig, A. M., & Podladchikova, T. 2023, *Nature Astronomy*, 7, 1171, doi: 10.1038/s41550-023-02030-9
- Lagaris, I. E., Likas, A., & Fotiadis, D. I. 1998, *IEEE Transactions on Neural Networks*, 9, 987, doi: 10.1109/72.712178
- LeCun, Y., Bottou, L., Orr, G. B., & Müller, K.-R. 1998, in *Lecture Notes in Computer Science*, Vol. 1524, *Neural Networks: Tricks of the Trade* (Springer), 9–50, doi: 10.1007/3-540-49430-8_2

- Mildenhall, B., Srinivasan, P. P., Tancik, M., et al. 2020, in European Conference on Computer Vision (ECCV), 405–421, doi: 10.1007/978-3-030-58452-8_24
- Norris, R. P. 2022, in *Multimessenger Astronomy in Practice* (IOP Publishing), doi: 10.1088/2514-3433/ac2256ch12
- Osborne, C. M. J., & Milić, I. 2021, *The Astrophysical Journal*, 917, 14, doi: 10.3847/1538-4357/ac02be
- Park, J. J., Florence, P., Straub, J., Newcombe, R., & Lovegrove, S. 2019, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 165–174, doi: 10.1109/CVPR.2019.00025
- Pasachoff, J. M., & Suer, T.-A. 2010, *Journal of Astronomical History and Heritage*, 13, 120, doi: 10.3724/SP.J.1440-2807.2010.02.07
- Priest, E. 2014, *Magnetohydrodynamics of the Sun* (Cambridge: Cambridge University Press), doi: 10.1017/CB09781139020732
- Rahaman, N., Baratin, A., Arpit, D., et al. 2019, in *Proceedings of Machine Learning Research*, Vol. 97, *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 5301–5310
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. 2019, *Journal of Computational Physics*, 378, 686, doi: 10.1016/j.jcp.2018.10.045
- Rempel, M. 2014, *The Astrophysical Journal*, 789, 132, doi: 10.1088/0004-637X/789/2/132
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. 1986, *Nature*, 323, 533, doi: 10.1038/323533a0
- Schüssler, M., & Knölker, M. 2001, in *ASP Conference Series*, Vol. 236, *Advanced Solar Polarimetry – Theory, Observation, and Instrumentation*, ed. M. Sigwarth (Astronomical Society of the Pacific), 295
- Stix, M. 2002, *The Sun: An Introduction*, 2nd edn., *Astronomy and Astrophysics Library* (Berlin, Heidelberg: Springer), doi: 10.1007/978-3-642-56042-2
- Tancik, M., Srinivasan, P. P., Mildenhall, B., et al. 2020, in *Advances in Neural Information Processing Systems*, Vol. 33, 7537–7547

БИБЛИОГРАФИЈА

Vögler, A., Shelyag, S., Schüssler, M., et al. 2005, *Astronomy & Astrophysics*, 429, 335, doi: 10.1051/0004-6361:20041507

Wiegmann, T., Thalmann, J. K., & Solanki, S. K. 2014, *The Astronomy and Astrophysics Review*, 22, 78, doi: 10.1007/s00159-014-0078-7

Биографија аутора

Давид Живковић (*Веџе (Vevey)*, 7. мај 2001.) завршио је основну школу „Љуба Нешић” и средњу Техничку школу у Зајечару. Уписао је основне академске студије студијског програма Астрономија и астрофизика (модул Астрофизика) школске 2020/21. године на Математичком факултету Универзитета у Београду. Дипломирао је 2025. године и исте године уписао мастер студије астрофизике. Од 2023. године је, упоредо са студијама, активан у Истраживачкој станици Петница као сарадник на семинарима астрономије. Поред тога, био је привремено ангажован као професор физике у Филолошкој гимназији у Београду.