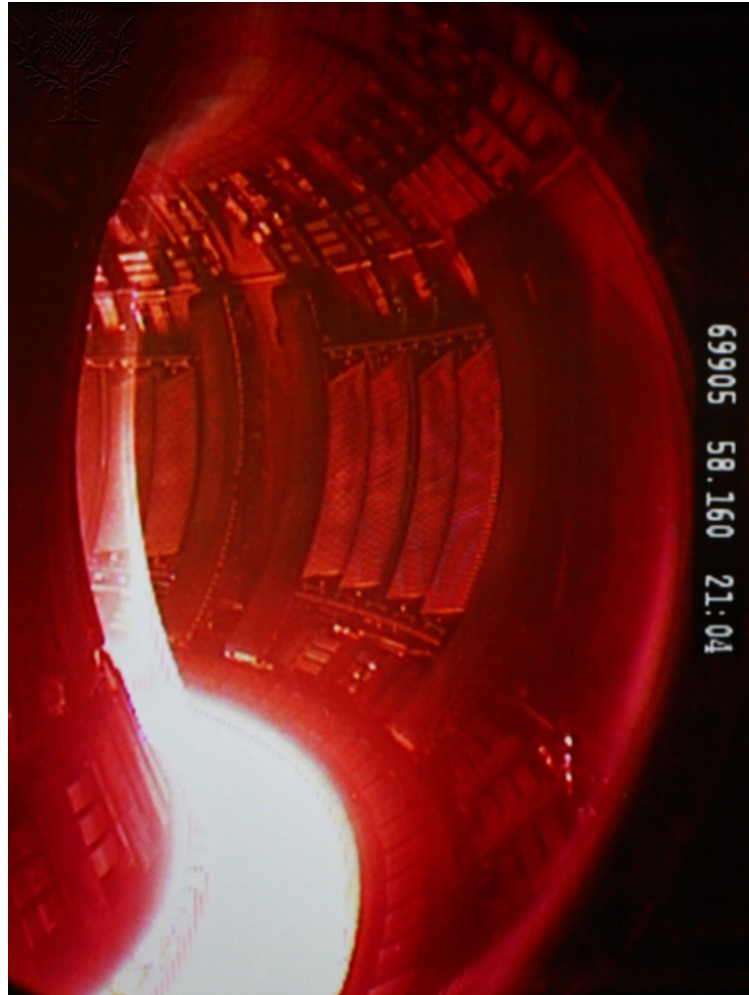




CHALMERS



Surrogatmodell av QuaLiKiz för turbulenta instabiliteter i en tokamak

Kandidatarbete inom programmen Teknisk fysik, Elektroteknik och Maskinteknik

Jonathan Borsander, Oscar Lindberg, Walter Rieck och Malte Åqvist

INSTITUTIONEN FÖR RYMD-, GEO- OCH MILJÖVETENSKAP

CHALMERS TEKNISKA HÖGSKOLA

Göteborg 2023

www.chalmers.se

KANDIDATARBETE 2023

**Surrogatmodell av QuaLiKiz för
turbulenta instabiliteter i en tokamak**

JONATHAN BORSANDER, OSCAR LINDBERG,
WALTER RIECK och MALTE ÅQVIST



CHALMERS

Institutionen för Rymd-, geo- och miljövetenskap
CHALMERS TEKNISKA HÖGSKOLA
Göteborg 2023

Surrogatmodell av QuaLiKiz för turbulenta instabiliteter i en tokamak
JONATHAN BORSANDER, OSCAR LINDBERG, WALTER RIECK
och MALTE ÅQVIST

© JONATHAN BORSANDER, OSCAR LINDBERG, WALTER RIECK
och MALTE ÅQVIST, 2023.

Handledare: Emil Fransson och Andreas Gillgren
Examinator: Pär Strand

Kandidatarbete 2023
Institutionen för Rymd-, geo- och miljövetenskap
Chalmers Tekniska Högskola
SE-412 96 Göteborg
Telefon +46 31 772 1000

Omslagsbild: En superhet fusionsplasma inuti JET (Joint European Torus), belägen
vid Culham Science Centre, Oxfordshire, Storbritannien, Europas största anläggning
för fusionsforskning (Encyclopædia Britannica ImageQuest, [2023](#)).

Skriven i L^AT_EX
Göteborg 2023

Surrogatmodell av QuaLiKiz för turbulenta instabiliteter i en tokamak
JONATHAN BORSANDER, OSCAR LINDBERG, WALTER RIECK
och MALTE ÅQVIST

Institutionen för Rymd-, geo- och miljövetenskap
Chalmers Tekniska Högskola

Sammandrag

Fusion är en energikälla med enorm potential, eftersom den endast kräver väte och inte producerar farligt avfall. Under de senaste åren har det skett betydande framsteg inom utvecklingen av fusionsenergi. En av de mest lovande metoderna för praktisk fusionsenergi är magnetiskt innesluten fusion i en tokamak. I en sådan fusionsanordning uppkommer turbulens som är viktig för inneslutning av energi och partiklar. Givet dess betydelse är kunskap om turbulensen värdefull. Dessvärre kräver de existerande modellerna mycket datorkraft för att simulera och analysera turbulensen, vilket innebär att det finns en motivation till att skapa mindre beräkningstunga modeller, till exempel genom maskininlärning. I detta projekt har ett antal surrogatmodeller av QuaLiKiz utvecklats med syftet att prediktera hastigheten som turbulenta instabiliteter växer med i fusionsplasma (tillväxthastighet, jfr. en. "growth rate") och associerade realfrekvenser. Resultaten indikerar att en optimal modell har ett neuralt nätverk per prediktionsvariabel och tränas på ett balanserat dataset med lika andelar stabila och instabila lösningar. I arbetet undersöks också huruvida ensembler-nätverk kan användas för identifiera kraschade simuleringar i QuaLiKiz-databasen.

Nyckelord: fusionsplasma, tokamak, qualikiz, turbulenta instabiliteter, maskininlärning, neurala nätverk.

Innehåll

1	Inledning	1
1.1	Bakgrund	1
1.2	Syfte	2
1.3	Mål	2
1.4	Avgränsningar	2
2	Teori	5
2.1	Plasma	5
2.2	Tokamaker	5
2.3	Gyrorörelse	6
2.4	Drift i plasma / Partikelrörelse i magnetiska fält	7
2.5	Transport	8
2.6	Instabiliteter	9
2.7	QuaLiKiz	9
2.8	Maskininlärning	9
3	Metod	13
3.1	Dataförståelse	13
3.2	Bearbetning av QuaLiKiz-databasen	13
3.2.1	Reduktion av dataset	14
3.2.2	Identifiering och borttagning av avvikande värden	14
3.3	Relevans av subdominant instabilitet	15
3.4	Jämförelse av neurala nätverksstrukturer	15
3.5	Införande av balanserade dataset	16
3.6	Hantering av kraschade simuleringar	17
3.6.1	Ensemble nätverk	17
3.6.2	Vågspektrumens kontinuitet	18
4	Resultat och diskussion	19
4.1	Jämförelse av neurala nätverksstrukturer	19
4.1.1	Nätverksarkitektur	19
4.1.2	Antal nätverk	19
4.2	Införande av balanserade dataset	22
4.3	Hantering av kraschade simuleringar	24
4.3.1	Ensemble nätverk	25
4.3.2	Vågspektrumens kontinuitet	25

5	Slutsats	27
5.1	Framtida projekt	27
	Litteratur	29
A	Träddiagram för balanserad databas	I
B	Spridningsdiagram för modellprediktioner	III

1

Inledning

Fusion är en process där två atomkärnor kombineras för att bilda en tyngre kärna, vilket frigör en enorm mängd energi. I solen och andra stjärnor förekommer fusion naturligt. Forskare har länge försökt återskapa denna process i kontrollerade former här på jorden, eftersom det är en nästan obegränsad och ren energikälla. Under de senaste åren har betydande framsteg gjorts inom utvecklingen av fusionsteknologi, men det finns fortfarande många utmaningar att övervinna för att utveckla ett praktiskt och effektivt fusionskraftverk.

Ett av de mest lovande tillvägagångssätten för att uppnå praktisk fusionsenergi är genom användningen av tokamaker och magnetiskt innesluten fusion. En tokamak är en torusformad konstruktion som använder starka magnetfält för att innesluta ett fusionsbränsle. I tokamaken värms bränslet upp till extrema temperaturer så att det hamnar i plasmatillstånd — ett gasliknande tillstånd av joniserad materia, det vill säga materia som består av fria elektroner och atomkärnor. De extrema temperaturerna ger förutsättningar för atomkärnor att kollidera, kombineras och frigöra energi. För att lyckas uppnå effektiv fusion i tokamaker är det viktigt att förstå plasmats beteende i dessa anordningar.

1.1 Bakgrund

Plasmapartiklar interagerar ständigt med varandra och med omgivningen. Eftersom plasma består av laddade partiklar påverkas det av elektriska och magnetiska fält. Dessutom kan plasma leda elektrisk ström, och därmed ge upphov till magnetiska fält. Dessa egenskaper gör att komplexa beteenden uppstår i bränslet för tokamaker och andra fusionsanordningar.

Det finns flera forskningsanläggningar runt om i världen som använder tokamak-konstruktioner för fusion, varav den största är Joint European Torus (JET) i Culham, Storbritannien. Under åren har JET uppnått flera milstolpar inom fusionsforskning (se bl.a. Gibney, 2022). För framtida fusionsreaktorer måste partikel- och energiinneslutningen, som till stor del bestäms av den turbulenta transporten i plasma, vara god. Ett sätt att beräkna och förstå den turbulenta transporten är genom simuleringar med modellen QuaLiKiz (Bourdelle m. fl., 2015). QuaLiKiz används för att för att beräkna transportflöden (jfr. eng. “transport fluxes”), samt hastigheten

som turbulenta instabiliteter växer med (tillväxthastighet, jfr. eng. “growth rate”) och associerade realfrekvenser. I QuaLiKiz-modellen beräknas tillväxthastigheten först, och därefter transportkoefficienterna från tillväxthastigheten.

Modellering med QuaLiKiz och liknande mjukvara är mycket beräkningstungt, vilket ofta leder till långa simuleringsperioder. En komplicerad och tidskrävande modell kan ibland ersättas av en approximerande modell, en så kallad surrogatmodell, som efterliknar den ursprungliga modellen. Med hjälp av maskininlärning kan en surrogatmodell tränas på ett dataset som representerar den ursprungliga modellen. Därefter kan surrogatmodellen användas för att göra tidseffektiva förutsägelser utifrån helt ny data. Denna modell presterar bäst på data inom samma parameterområde som den tränats i.

I fallet med QuaLiKiz finns en neural nätverksmodell utvecklad av Ho m. fl. (2021) som är utformad för att prediktera transportflöden. Det finns emellertid ingen neural nätverksmodell för QuaLiKiz som predikterar tillväxthastigheten. Fördelen med en surrogatmodell för tillväxthastigheten är att en sådan modell har en mer solid fysik bakom sig och lättare kan extrapoleras till framtida, större tokamaker. Arbetet med att utforma en modell för tillväxthastigheten skulle kunna ge viktiga insikter för framtida projekt — inte bara för QuaLiKiz-relaterade arbeten, utan också för andra modeller där tillväxthastigheterna beräknas. Modellen skulle även vara ett bidrag mot målet om snabbare och enklare fusionsforskning.

1.2 Syfte

Syftet med projektet är att utveckla en surrogatmodell av QuaLiKiz som predikterar hastigheten som turbulenta instabiliteter växer med i fusionsplasma och associerade realfrekvenser.

1.3 Mål

De huvudsakliga målen för projektet är följande:

- (i) Presentera ett eller flera förslag på surrogatmodell av QuaLiKiz som tidseffektivt predikterar tillväxthastigheter och associerade realfrekvenser.
- (ii) Presentera resonemang kring de aspekter och överväganden som påverkar surrogatmodellens prestanda.

1.4 Avgränsningar

Projektet avgränsas genom att använda relativt enkla neurala nätverk samt manuell hyperparameteroptimering. Vidare avgränsas modellen till att enbart prediktera tillväxthastigheter och associerade realfrekvenser, och inte transportflöden, eftersom detta redan har genomförts med en annan modell. I projektet används också redan

existerande QuaLiKiz-data, och därför kommer ingen ny test- eller träningsdata att genereras.

2

Teori

Fusion är namnet på processen som sker då två atomkärnor sammanslås och bildar en större kärna. Denna process har möjligheten att omvandla enorma mängder energi enligt $E = mc^2$, och är grundprincipen bakom hur alla stjärnor i universum utstrålar energi. Problemet med att återskapa och nyttja denna process är att två atomkärnor med samma laddning repellerar varandra långt innan de är nära nog för att smälta samman. Alltså måste de två atomkärnorna ha nog med rörelseenergi för att övervinna Coulomb-repulsionen tills den starka kärnkraften börjar verka. Eftersom rörelseenergi och temperatur är kopplade till varandra enligt ekvipartitionsteoremet leder detta till en extremt hög temperatur.

2.1 Plasma

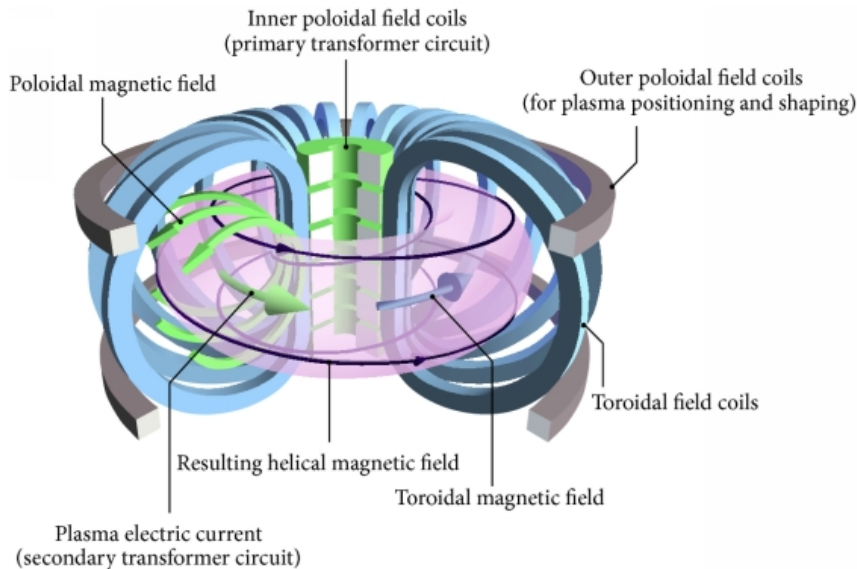
För att fusion ska ske krävs extremt höga temperaturer, vilket tvingar bränslet in i ett plasmatillstånd. Detta innebär att bränslet blir till en varm joniserad gas, där elektroner och atomkärnor inte längre är kopplade till varandra. Plasman är så kallad kvasineutral därför att trots att antalet positiva och negativa laddningar är samma så kan mycket små lokala elektriska fält uppstå i plasman då partiklarna inte är bundna till varandra. Att alla enskilda partiklar har en laddning leder till att plasman som helhet kan manipuleras av magnetiska fält (Dan Anderson och Marklund, 2003).

Plasma skapas vid en kombination av hög temperatur samt låg densitet. Detta betyder att plasma finns mellan stjärnorna där temperaturen är väldigt låg men också densiteten väldigt låg, men också i fusionsreaktorer med hög temperatur och relativt hög densitet. Plasma upptar cirka 99 % av normal materia i universum.

2.2 Tokamaker

Tokamak är en akronym för den ryska översättningen av “toroidal kammare i magnetiska spolar”. Tokamaken kännetecknas av sin torusformade, eller “donutformade”, geometri samt de starka magneterna runt anordningen. Tokamaken utnyttjar att de extremt höga temperaturerna som krävs för fusion tvingar bränslet in i plasmatillstånd som sedan, med hjälp av magneter, kan styras runt i toroiden. En schematisk

bild för hur en tokamak ser ut visas i figur 2.1 (S. Li och Xu, 2014). Det toroidformade lila området i figuren är plasmat som innesluts och leds av magnetfältet som genereras av de blå poloidala magneterna.



Figur 2.1: Skiss av tokamakens struktur samt de magnetiska fältlinjer som manipulerar plasmat.

2.3 Gyrorörelse

Laddade partiklar som rör sig i ett magnetfält upplever en kraft som kallas för Lorentzkraften, vilken ges av

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}), \quad (2.1)$$

där q är partikelns laddning, \vec{E} är det elektriska fältet, \vec{v} är partikelns hastighet och \vec{B} är den magnetiska flödestätheten. Eftersom att $\vec{E} = 0$ i en tokamak, om man bortser från bidraget från små störningar, kan Lorentzkraften uttryckas enligt

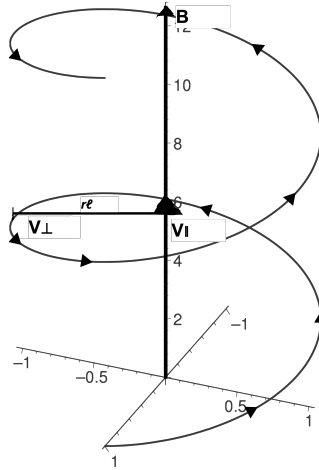
$$\vec{F} = q\vec{v} \times \vec{B}. \quad (2.2)$$

Den resulterande kraften accelererar partikeln vinkelrätt mot både rörelseriktningen och det magnetiska fältet. Följaktligen är hastigheten parallell med magnetfältet oförändrad. Detta leder till att partikeln förflyttas i en spiralbana längs med fältlinjerna. Radien som partikeln roterar kring fältlinjen med kallas för Larmor-radien och definieras enligt

$$r_L = \frac{mv_{\perp}}{|q|B}. \quad (2.3)$$

Denna Larmor-radie är direkt kopplad till storleken av instabiliteterna som uppstår i en tokamak. Eftersom elektroner har mindre massa än joner så har de en mindre

Larmor-radie, vilket betyder att instabiliteter associerade med elektroner är mindre och därför oftast inte lika viktiga (Jenko m. fl., 2000). En bild på hur gyrorörelse ser ut kan ses i figur 2.2.



Figur 2.2: Den helixformade rörelse en laddad partikel färdas genom tokamaken med. v_{\perp} är den rörelse är vinkelrät mot magnetfältet B och därmed den rörelse som ger upphov till Lorentzkraften, medan v_{\parallel} är den rörelse som är parallell med magnetfältet och r_l är Larmorradien.

2.4 Drift i plasma / Partikelrörelse i magnetiska fält

För att upprätthålla stabiliteten i tokamaken är det viktigt att veta vilka faktorer som skulle kunna störa stabiliteten och hur dessa kan minimeras. Eftersom tokamaken använder sig av magnetiska fält kommer partiklar uppleva olika sorters drifter som kan bidra till en asymmetri i plasmat. Drift innebär partikelrörelse i en riktning som skiljer sig från de magnetiska fältlinjerna. De tre drifter som är mest relevanta för plasma i en tokamak är $\vec{E} \times \vec{B}$ -drift, $\nabla \vec{B}$ -drift och kurvaturdrift. Den förstnämnda driften ges av:

$$\vec{v}_{\vec{E} \times \vec{B}} = \frac{\vec{E} \times \vec{B}}{B^2}. \quad (2.4)$$

Eftersom hastigheten är oberoende av partikelns laddning påverkas både negativt- och positivt laddade partiklar med en kraft som verkar i samma riktning. Detta gör att $\vec{E} \times \vec{B}$ -driften kan orsaka makroskopisk rörelse av plasmat.

Magnetfältet i tokamaken har en gradient, det vill säga att magnetfältet varierar över rummet, och därför kommer även en ∇B -drift att påverka partiklarna. Partiklar i plasmat upplever då en kraft som är proportionell mot gradienten av det magnetiska fältet enligt

$$v_{\nabla B} = \frac{\mu \vec{B} \times \nabla B}{q B^2}, \quad (2.5)$$

där μ är det magnetiska momentet. Eftersom det finns ett laddningsberoende, q , kommer negativa och positiva joner känna av denna kraft på olika sätt, vilket leder till drifter åt olika håll. Detta kan skapa en laddningsseparation i plasmat och således påverka plasmats beteende.

Kurvaturdrift uppstår när det magnetiska fältet är vridet, det vill säga när det finns en krökning eller kurvatur längs med fältlinjerna, vilket är fallet i en tokamak. Partiklar som rör sig längs med dessa fältlinjer upplever då en kraft som är proportionell mot krökningen av fältlinjerna. Denna kraft kan orsaka att partiklar av olika hastighet eller laddning förflyttas i olika riktningar. Sammantaget är dessa tre drifter särskilt viktiga för uppkomsten av turbulenta instabiliteter, som diskuteras djupare i avsnitt 2.6.

2.5 Transport

Inom fusion och tokamak-forskning är transport mot väggen av tokamaken en viktig faktor som påverkar prestandan och stabiliteten hos plasmat. Transport handlar om hur partiklar och energi rör sig i plasman och kan delas upp i flera olika kategorier. Partiklarna är bundna till magnetfältlinjerna enligt ekvation 2.1, vilket gör att partiklarna rör sig fritt längs med fältlinjerna. Transport ut från tokamaken, mot väggarna, är mycket långsammare men viktigare eftersom den bestämmer inneslutningen av partiklar och energi. De tre mest relevanta transporttyperna är: klassisk-, neoklassisk- och turbulent transport.

Den klassiska och neoklassiska transporten är två olika typer av kollisionell transport, vilket innebär att transporten drivs av kollisioner mellan partiklar. En kollision av två partiklar av samma typ ger inget överflödigt partikelflöde på grund av rörelsemängdsbevaring. Att nettoflödet av partiklar är oändrat leder till partikeltransporten för elektroner och joner är densamma. En viktig del av den klassiska transporten är "random walk" (jfr. sv. slumpmässig rörelse), vilket är benämningen på den process där en partikel rör sig vinkelrätt mot fältlinjerna och således byter omloppsbana. Steglängden på dessa drifter bestäms av gyrorörelsen och därmed Larmor-radien.

Neoklassisk transport tar hänsyn till det magnetiska fältet, partiklarnas egenskaper, samt hur partiklarna rör sig i geometrin, i detta fall en tokamak. Neoklassisk transport uppkommer huvudsakligen från fångade partiklar som har fastnat på utsidan av tokamaken där magnetfältet är svagare. En partikel som är fast i detta område har inte nog med hastighet parallell med magnetfältet för att åka runt hela torusen. Istället vänder partikeln när magnetfältet blir för starkt och får därmed en bananformad rörelse i tokamakens ytterkant. Bredden på den bana som de fångade partiklarna tar runt tokamaken är vad som bestämmer steglängden på neoklassisk transport. Denna bredd är mycket större än Larmor-radien i klassisk transport, vilket leder till att steglängden för neoklassisk transport är större.

Den typ av transport som står för det mesta av värme- och partikeltransporten i

en tokamak är turbulent transport, som orsakas av mikroinstabiliteter i plasmat. Dessa mikroinstabiliteter kan uppstå till följd av den radiella rörelsen som genereras av $\vec{E} \times \vec{B}$ -driften (Nordman och Weiland, 1989).

2.6 Instabiliteter

De tre dominanta mikroinstabiliteterna i turbulent transport är: Ion temperature gradient (ITG), electron temperature gradient (ETG) och trapped electron mode (TEM) (Liu, 1969). ITG-instabilitet uppstår eftersom fusion kräver att jonerna i plasman har mycket hög temperatur. Detta gör att det naturligt förekommer stora temperaturgradienter i plasmat. Om temperaturgradienten är tillräckligt stor kan den orsaka störningar i plasmans elektromagnetiska fält. Detta kan leda till att energi och partiklar transporteras ut från plasmat mot väggen. De höga temperaturgradienterna påverkar även elektronerna i plasmat och leder till en motsvarande instabilitet för elektronerna — ETG-moden. Då ETG-instabiliteten verkar vid den mindre elektronlängdskalan är den vanligtvis inte är lika viktig som ITG-instabiliteten. TEM-instabiliteten är associerad till fångade partiklar i den yttre delen av tokamaken. Den drivs av densitetsgradienten samt elektrontemperaturgradienten. TEM sker på samma längdskala som ITG.

2.7 QuaLiKiz

QuaLiKiz är en öppen källkod utformad för att modellera den turbulenta transporten hos plasma i tokamaker (Bourdelle m. fl., 2015), vilket inkluderar ITG, ETG samt TEM. Genom att beräkna styrkan för de turbulenta instabiliteterna och kombinera det med olika plasmametrar, får QuaLiKiz fram den turbulenta transporten. I detta projekt kommer vi fokusera på styrkan eller hastigheten för de turbulenta instabiliteterna vid olika längdskalor.

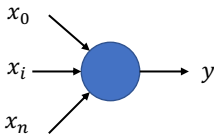
QuaLiKiz beräknar den turbulenta transporten för ett antal olika längdskalor från större jonskalor till mindre elektronskalor. QuaLiKiz använder en fourierrepresentation så att längdskalorna representeras av 18 olika index för $k\rho$, där k är vågnumret och ρ är en normaliseringsfaktor som är av samma storleksordning som Larmorradien för joner. Varje körning av QuaLiKiz har som output den dominanta och subdominanta instabiliteten för tillväxthastigheten och realfrekvens per $k\rho$ -index.

2.8 Maskininlärning

Maskininlärning är en gren inom artificiell intelligens som syftar till att utveckla algoritmer och system som kan lära och förbättra sig själva genom att analysera data och erfarenheter. Neurala nätverk är en del av maskininlärning, de är universella approximatorer vilket gör dem till ett mycket kraftfullt verktyg (Goodfellow m. fl., 2016).

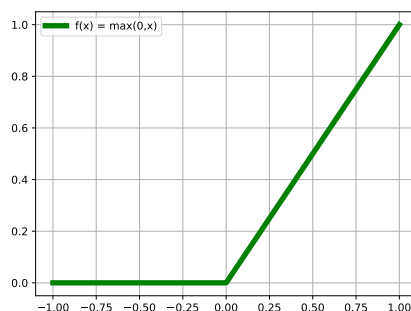
I det här arbetet används fullständigt sammankopplade neurala nätverk för att ap-

proximera hur output-variabler beror på input-variabler. I ett fullständigt sammankopplat neuralt nätverk är neuronerna och aktiveringsfunktionen de grundläggande byggstenarna, se figur 2.3.

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right)$$


Figur 2.3: Matematisk och schematisk representation av en neuron. Neuronerna tar flera inputvariabler, multiplicerar varje input, x_i , med en given vikt, w_i , summerar resultatet, lägger till en tröskelparameter, b , och applicerar en aktiveringsfunktion, f .

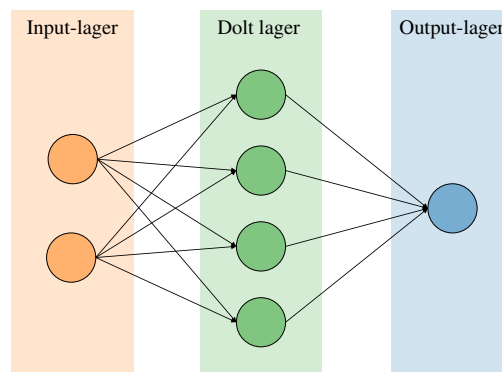
Neuronerna är noderna i nätverket som utför beräkningar och skickar information mellan varandra. Aktiveringsfunktionerna gör att neuralt nätverk kan lära sig att känna igen mönster och samband i data och göra förutsägelser eller ta beslut baserat på denna information. En av de vanligaste aktiveringsfunktionerna är ReLU, figur 2.4, outputen av en ReLU blir noll om inputen är negativ, samt att outputen av en ReLU är samma som inputen om inputen är positiv. ReLU är en icke-linjär funktion med vilket medöljer att nätverket då kan lära sig approximera icke-linjära funktioner.



Figur 2.4: Aktiveringsfunktionen ReLU.

Tröskelparametrar (jfr. en. “bias”) är värden som läggs till i varje neuron i nätverket vilket fungerar som en förskjutning av aktiveringen för neuronerna. Neuronerna är fördelade i dolda lager, där varje neuron tar input-data från föregående lager, se figur 2.5.

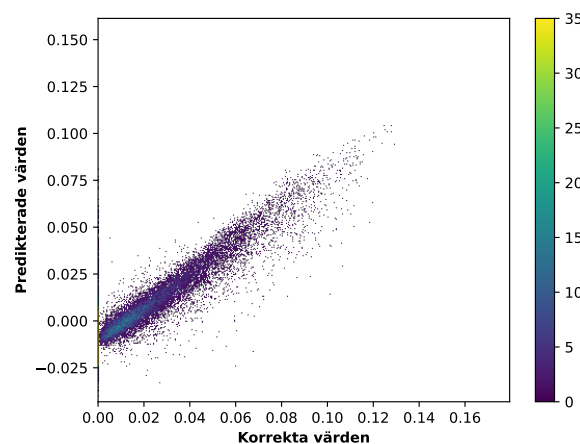
En kostnadsfunktion (jfr. en. “cost function”), är en funktion som används inom maskininlärning för att mäta hur bra ett neuralt nätverk presterar för en given uppgift. I korthet beräknar kostnadsfunktionen skillnaden mellan de förutsagda resultaten från nätverket och de faktiska resultaten. Syftet med kostnadsfunktionen är att optimera nätverkets vikter och tröskelparametrar för att minska kostnaden och förbättra nätverkets prestanda på uppgiften. Ju lägre kostnad, desto bättre prestanda från nätverket (Mehlig, 2022).



Figur 2.5: Visualisering av neuralt nätverk med input-lager, dolt lager och output-lager, och neuroner.

Genom att justera vikterna mellan neuroner under träningen kan nätverket förbättra sin förmåga att utföra specifika uppgifter. I det här arbetet används återpropagering (jfr. en. “backpropagation”) som inlärningsmetod. Återpropagering innebär att man kan derivera sin kostnadsfunktion som en funktion av nätverkets vikter och tröskelparametrar. Då erhålls ett optimeringsproblem där man gradvis går mot lokala minimum av kostnadsfunktionen — vikterna justeras så att kostnaden minimeras.

Ett tränat nätverk testas genom att göra prediktioner ett testset, som innehåller data som inte tidigare använts. I figur 2.6 visas ett exempel på hur ett neuralt nätverk har predikerat på ett testset genom ett spridningsdiagram med predikerade mot verkliga värden. Om nätverket predikterar ett korrekt värde exakt, hamnar motsvarande datapunkt på en tänkt rät linje genom origo med lutningskoefficienten ett. Hur väl nätverket lyckats prediktera kan mätas på fler olika vis, till exempel genom anpassningsmättet R^2 .



Figur 2.6: Ett spridningsdiagram för korrekta jämfört med predikerade värden av ett neuralt nätverk. Varje punkt i diagrammet motsvarar en prediktion av nätverket.

3

Metod

I detta kapitel presenteras metoden för studien. Metoden delas upp i tre huvudsakliga delar: jämförelse av neurala nätverksstrukturer, införande av balanserade dataset och hantering av kraschade simuleringar. I metoden ingår även två inledande delar kring dataförståelse och bearbetning av QuaLiKiz-databasen. Samtliga datorberäkningar i projektet genomfördes i Python, och all maskininlärning med Python-biblioteket Keras från TensorFlow. Koden för detta finns i mappen [länkad här](#).

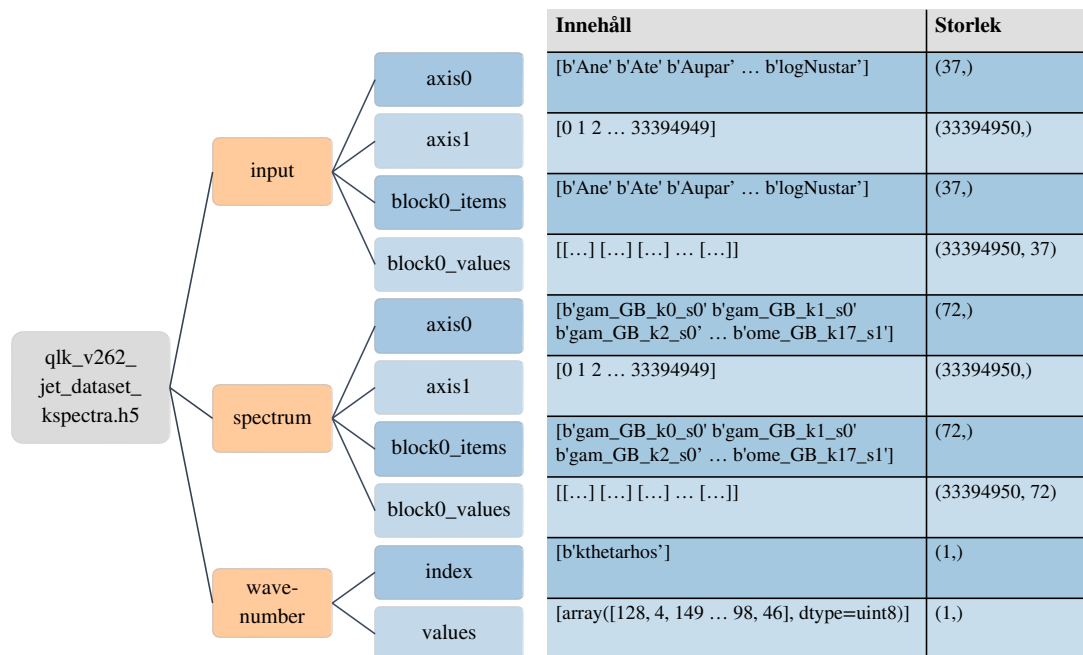
3.1 Dataförståelse

Innan maskininlärningsarbetet påbörjades genomfördes en grundlig explorativ dataanalys. Databasen (`qlk_v262_jet_dataset_kspectra.h5`), beskriven i Ho m. fl. (2021), laddades ner och extraherades med hjälp av Python. Databasen består av cirka 33 miljoner simuleringar, var och en utförd med en standardvektor bestående av 18 vågnummer. Dess struktur och innehåll beskrivs närmare nedan. Ett träd-diagram av den övergripande databasstrukturen visas i figur 3.1.

Datan är sparad i HDF5-format under tre nycklar: “input”, “spectrum” och “wavenumber”. Nyckeln “input” innehåller de ingångar som använts för QuaLiKiz-utvärderingarna. Där ingår lokala plasmaparametrar som erhållits från experimentella mätningar på JET (Joint European Torus) i Storbritannien, samt variationer av utvalda parametrar enligt experimentella osäkerheter. Nyckeln “spectrum” innehåller de linjära tillväxthastigheter och frekvensspektra som svarar mot de två mest dominerande mikroinstabiliteterna, där “s0” betecknar första ordningens instabilitet (dominant) och “s1” andra ordningens instabilitet (subdominant). Nyckeln “wavenumber” innehåller en matris med en standarduppsättning av 18 vågnummer ($k_y \rho_s$), som använts för att generera spektrumet, där “k0” är det lägsta vågnumret och “k17” det högsta.

3.2 Bearbetning av QuaLiKiz-databasen

Den ursprungliga QuaLiKiz-databasen bearbetades genom att reducera antalet simuleringar och ta bort vissa extremvärden.



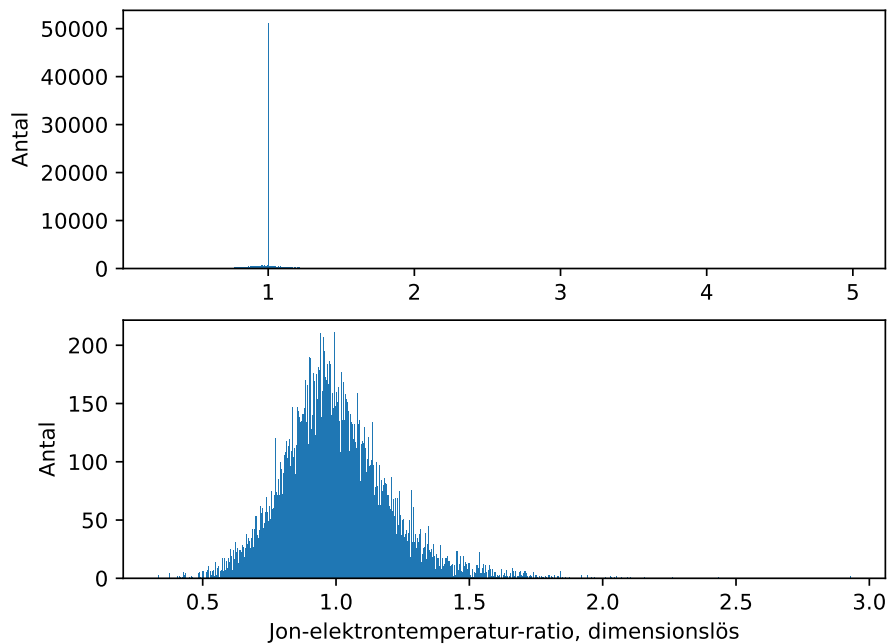
Figur 3.1: Träddiagram över databasens innehåll. Databasen (qlk_v262_jet_dataset_kspectra.h5) består av tre grupper: “input”, “spectrum” och “wave-number”. Varje grupp innehåller ett antal dataset, till exempel “axis0”, “axis1”, “block0_items” och “block0_values”, vars innehåll och storlek beskrivs av motsvarande rad i den intilliggande tabellen. Se tabell 3.1 för en beskrivning av variabelnamnen.

3.2.1 Reduktion av dataset

Datan i QuaLiKiz har 37 indatavariabler enligt figur 3.1. Av dessa valdes 15 som de viktigaste variablerna, till exempel valdes “Ati0” men inte “Ati1” och “Ati2” eftersom dessa variabler är mycket lika sett till distributionen. Dessa 15 användes också i ett liknande projekt som skapade ett neuralt nät för beräkning av transportflöden (Ho m. fl., 2021). I ett neuralt nätverk behöver inte hela databasen användas för att skapa en bra förutsägelse. Därför bestämdes det att träningsset och testset med 50000 exempel var skulle skapas så att träning av modellerna inte skulle ta för lång tid.

3.2.2 Identifiering och borttagning av avvikande värden

I datan finns flera simuleringar där värden som inte borde kunna existera finns. För att förhindra att dessa värden påverkar modellen raderades de simuleringar som innehåller “outliers” och värden som inte är fysikaliskt möjliga. Extremvärdena bestämdes genom inputvariablernas histogram. Ett exempel på hur borttagandet av vissa värden för inputvariabeln för jon-elektrontemperatur-ratio kan förbättra modellen ses i figur 3.2. I tabell 3.1 presenteras de bestämda värdena.



Figur 3.2: Histogram för jon-elektrontemperatur-ratio som visar spik vid 1 i övre bild samt histogram för samma variabel då denna spik är borttagen i undre bild.

3.3 Relevans av subdominant instabilitet

I figur 3.1 kan man se att för varje ingångsmatrix finns det totalt 72 outputs. De sista 36 av dessa är tillväxthastigheter och realfrekvenser för den subdominanta instabiliteten. Databasen för dessa outputs är nästan uteslutande, minst $\sim 98\%$ per $k\rho$ -index, nollvärden. Dessa nollvärden indikerar stabila eller kraschade simuleringar, och på grund datans enformighet så begränsas nyttan av en prediktiv modell. Eftersom relevansen av den subdominanta instabiliteten i sammanhanget är låg utan vidare korrektioner eller viktning, avgränsas den fortsatta analysen till de 36 output-variabler som tillhör den dominanta instabiliteten.

3.4 Jämförelse av neurala nätverksstrukturer

De 36 outputs som anses vara nyttiga att prediktera med hjälp av neurala nätverk är tillväxthastigheten och realfrekvensen för 18 stycken $k\rho$ -index för den dominanta instabiliteten. Det finns flera olika sätt att kategorisera hur dessa outputs skall hanteras och hur mycket varje enskilt nätverk ska belastas med att prediktera. Till exempel kan ett nätverk förutsäga alla 36 värden eller så kan 36 nätverk skapas där varje nätverk förutsäger ett värde. För att undersöka vilken nätverksstruktur som fungerar bäst för denna uppgift skapades fyra olika nätverksstrukturer som jämfördes. Dessa visas i tabell 3.2.

Tabell 3.1: Tabell av dimensionslösa parametrar som används i de neurala nätverken, dess min, max samt värden som ska tas bort i databasen, “-” betyder att värdet är det originella från databasen.

Ingångsvariabel	Namn	Min	Max	Borttaget värde
Tryckgradient, norm.	alpha	-0,3	1,5	-
Elektrondensitetsgradient, norm.	Ane	-7	15	-
Jondensitetgradient nr 1, norm.	Ani1	-5	10	0
Elektrontemperaturgradient, norm.	Ate	-5	25	-
Jontemperaturgradient nr 0, norm.	Ati0	-5	25	-
Rotationsgradient, norm.	Autor	-	-	0
ExB shearings hastighet	gammaE	-	-	0
Kollisionalitet	logNustar	-2	2	-
Toroidalt machtal	Machtor	-	-	0
Jondensitet nr 1, norm.	normni1	0	0,05	-
Säkerhetsfaktor	q	0	8	-
Magnetiskt shear	smag	-0,2	4	-
Jon-elektrontemperatur-ratio	Ti_Te0	-	-	1
Normaliserad radiell position	x	0	1	-
Föreningssinnehåll	Zeff	0	4	-

Tabell 3.2: Beskrivning av de nätverksstrukturer som testas i projektet.

Antal nätverk	Prediktion
1	Alla 36 outputvariabler
2	Ett för tillväxthastigheter, ett för realfrekvenser
18	Ett för varje $k\rho$ -index (tillväxthastighet och realfrekvens)
36	Ett för varje output

Det är dock inte endast olika mängder av nätverk som kan påverka hur modellen förutsäger, även nätverksarkitekturen kan påverka resultaten. För att bestämma vilken arkitektur som fungerar bäst gjordes försök med mellan 50 och 500 noder samt med mellan ett och tre lager.

3.5 Införande av balanserade dataset

I de slumpmässigt skapade dataseten har vissa $k\rho$ mycket fler stabila lösningar, där QuaLiKiz ger värde noll, än instabila lösningar. Modellen kan då ha svårt att förut säga instabila lösningar eftersom modellen blir partisk mot att prediktera nollvärden då felet för modellen blir lägre, detta gäller speciellt för $k\rho$ -index 8 och 9 då dessa har cirka 99 % nollor. Därför skapades balanserade dataset med 50 % stabila och 50 % instabila lösningar. Detta gjordes genom att för varje $k\rho$ dela upp all data i stabila och ickestabila lösningar och sedan slumpmässigt välja 25000 av varje för att då skapa en ingångsmatrix samt utgångsvektor för varje $k\rho$. Databasens struktur

presenteras i bilaga A. En modell skapades sedan med dessa dataset med 250 noder och ett lager samt ett nätverk per utgångsvärde.

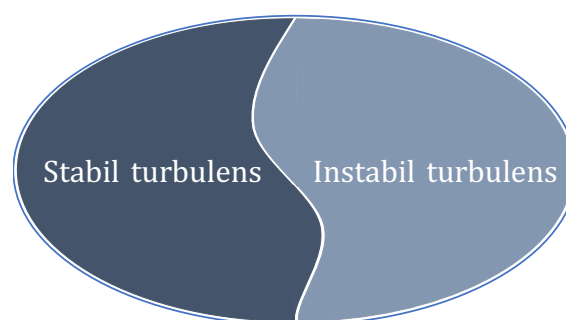
3.6 Hantering av kraschade simuleringar

QuaLiKiz kan för vissa simuleringar krascha och ger då samma resultat som om input-variablerna producerade en stabil lösning. QuaLiKiz ger output noll för krascher och stabila lösningar. Om träning görs på dessa simuleringar kan nätverket tro att en del av den instabila domänen är stabil och därför påverka resultatet.

3.6.1 Ensemblenätverk

Att träna flera olika modeller på samma data och sedan studera skillnaderna i hur de förutspår kallas för ensembleträning och kan ge en inblick i datan. Ensembleträning är den metod som valdes för att lokalisera krascher. Neurala nätverk är grunden till ensemblenätverk, men flera stycken används för att se saker ett ensamt nätverk inte kan. Skillnaden mellan nätverken i en ensembleträning kommer från att vikter och tröskelparametrar initialt valdes slumpmässigt.

De neurala nätverken tränades på data inom domän för instabil turbulens, figur 3.3, vilket är output-variabler skilda från noll. Sedan fick nätverken prediktera på data vars output-variabler var noll, vilket är stabil turbulens och krascher. Nätverken ska vara osäkra på att prediktera stabil turbulens då de inte är tränade för det. Nätverken ger en distribution av output-värden, det är variansen i denna distribution som indikerar ifall nätverken tillsammans är säkra på prediktionen. Om nätverken är väldigt säkra tyder det på att det är data som den är tränad för att prediktera och bör därför tillhöra domänen för instabil turbulens, alltså har simuleringen kraschat här.



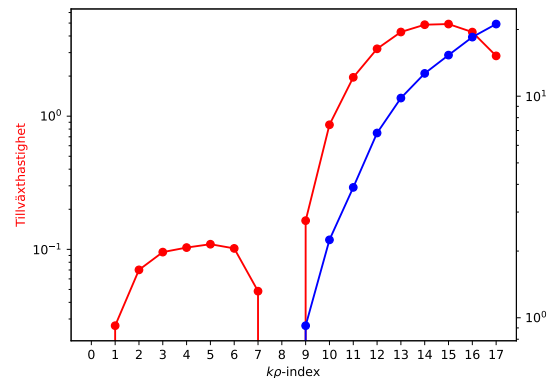
Figur 3.3: Hela domänen för datan från QuaLiKiz innefattar stabil turbulens, instabil turbulens och i domänen för stabil turbulens finns kraschade simuleringar av instabil turbulens.

Ensemblenätverken använde sig liksom tidigare delar av arbetet av fullständigt sammankopplade neurala nätverk. Arkitekturen för de neurala nätverken i ensembleträningen var samma för alla tio nätverk, 15 input-noder, ett dolt lager med 150 noder och en output-nod.

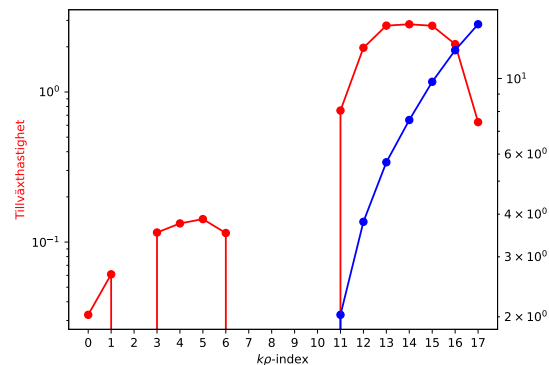
3.6.2 Vågspektrumens kontinuitet

500 slumpmässigt utvalda spektrum från QuaLiKiz togs fram för att få ett mått på hur stor andel av simuleringarna som kraschar. Därefter valdes slumpmässigt 500 simuleringar ut från när modellen predikterar att simuleringen är en krasch. Det gjordes för de lägsta 50 %, 30 % och 10 % av variansen. Hypotesen var att andelen identifierade krascher skulle vara fler när modellen tagit fram det den identifierat som krasch än de slumpmässigt utvalda. Samt skulle andelen öka då gränsen för variationen minskade.

Varje spektrum ritades upp med ett datorprogram för att sedan studeras visuellt. Figur 3.4 och 3.5 är exempel på vågspektrum som har analyserats, varje figur har två spektrum första spektrumet är upp till $k\rho$ -index 9 och andra spektrumet följer det. I figur 3.4 syns två kontinuerliga spektrum i tillväxthastighet och ett kontinuerligt i realfrekvens. Det syns inget spektrum för lägre $k\rho$ -index då y-axeln är logariterad och realfrekvensen ibland är negativ för mindre $k\rho$ -index. I figur 3.5 saknas det en tillväxthastighet vid $k\rho$ -index 2 för att göra spektrumet kontinuerligt vid låga $k\rho$ -index. Avsaknaden av ett värde visar diskontinuitet och simuleringen anses här ha kraschat.



Figur 3.4: Kontinuerliga vågspektrum för tillväxthastighet och realfrekvens, tillväxthastighet och realfrekvens är färgkodade med röd respektive blå färg.



Figur 3.5: Diskontinuerligt vågspektrum för tillväxthastighet och realfrekvens, tillväxthastighet och realfrekvens är färgkodade med röd respektive blå färg.

4

Resultat och diskussion

I detta kapitel presenteras resultat och diskussion för studiens tre huvudsakliga delar; jämförelse av neurala nätverksstrukturer, införande av balanserade dataset och hantering av kraschade simuleringar.

4.1 Jämförelse av neurala nätverksstrukturer

Vid jämförelse av olika nätverksstrukturer för surrogatmodellen av QuaLiKiz beaktas dels nätverksarkitekturen, dels antal nätverk.

4.1.1 Nätverksarkitektur

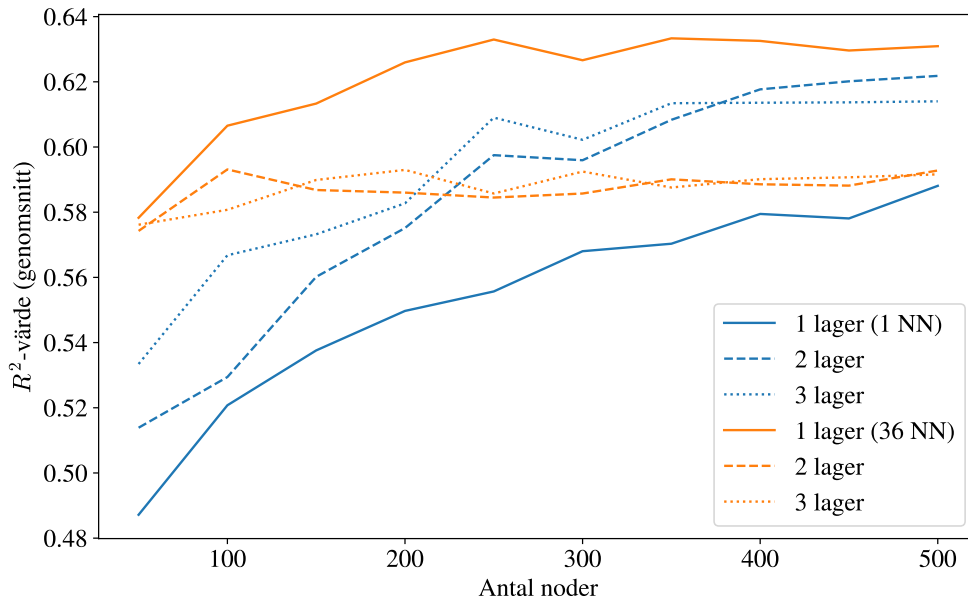
De utvalda arkitekturernas prestanda presenteras i figur 4.1. Här är det genomsnittliga R^2 -värdet en funktion av antalet noder per dolt lager för ett, två och tre lager. Det genomsnittliga R^2 -värdet korrelerar positivt med antalet lager och antalet noder per dolt lager för modellerna med ett stort nätverket (1 NN). Vid ≥ 400 noder stagnerar utvecklingen något, särskilt för arkitekturen med tre lager.

Modellerna med de mindre nätverken (36 NN) presterar genomgående bättre med ett dolt lager än med två eller tre stycken. Redan vid ~ 50 noder stagnerar R^2 -värdet för två- och trelagersarkitekturen. Att prestandan blir sämre med fler lager kan antagligen förklaras av överträning (jfr. en. “overfitting”), det vill säga att modellerna anpassas till träningsdata i så hög grad att problem uppstår med att extrapolera till testdata.

Baserat på resultaten från figur 4.1 skulle en lämplig arkitektur kunna vara att använda ett dolt lager med 250 noder. Denna arkitektur ger ett förhållandevis högt genomsnittligt R^2 -värde, och verkar erbjuda en god kombination av prestanda och tidseffektivitet. Fortsättningsvis används därför uteslutande arkitekturen med ett dolt lager med 250 noder, om inget annat anges.

4.1.2 Antal nätverk

Det finns flera olika möjliga nätverksstrukturer med avseende på antalet neurala nätverk och därmed antalet outputvariabler som varje nätverk predikterar (se ta-

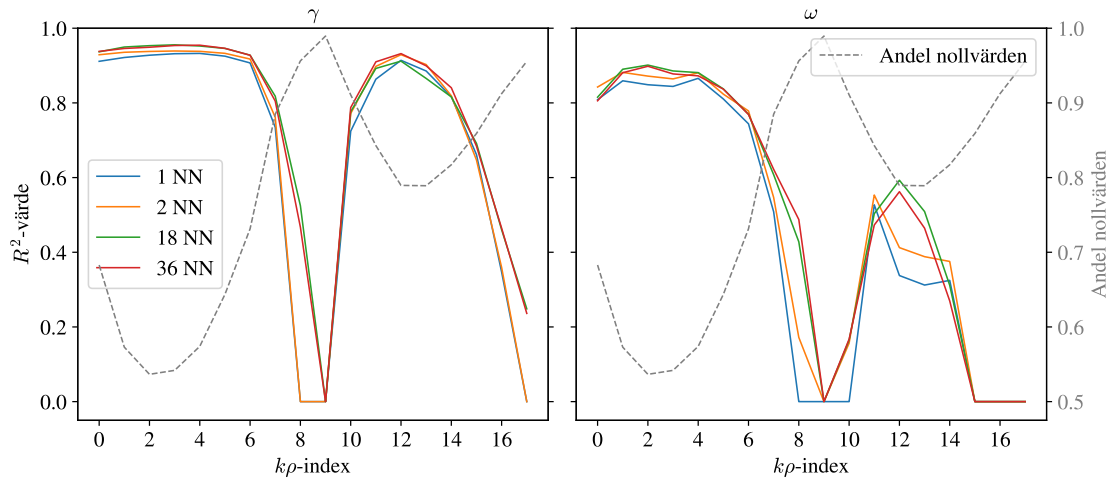


Figur 4.1: Genomsnittligt R^2 -värde som funktion av antalet noder per dolt lager. Graferna för det stora nätverket (1 NN) visas i blått, och graferna för de mindre nätverken (36 NN) i orange. Heldragen, streckad och prickad linje indikerar ett, två respektive tre dolda lager.

bell 3.2). De olika nätverksstrukturernas prestanda, mätt genom anpassningsmättet R^2 som funktion av $k\rho$ -index, presenteras i två bilder i figur 4.2. Diagrammen är uppdelade efter tillväxthastighet, γ , och realfrekvens, ω . Överlag har de mindre nätverken (18 NN och 36 NN) högre R^2 -värden än de större nätverken (1 NN och 2 NN), och visar därmed på bättre prestanda (se även tabell 4.1). Även de mindre nätverken presterar emellertid mycket dåligt ($R^2 < 0,25$) för tillväxthastighetens $k\rho$ -index 9 och 17, och för realfrekvensens $k\rho$ -index 9–10 och 15–17.

I figur 4.2 visas också andelen nollvärden i det dataset som modellerna har tränats på. Det verkar finnas en koppling mellan modellernas prestanda och andelen nollor i datasetet, vilket ungefär motsvarar andelen stabila lösningar. Vid $k\rho$ -index med mycket hög andel nollor, är R^2 -värdet generellt lågt (se t.ex. $k\rho$ -index 9). Antagligen beror detta på att modellerna vid dessa index inte har tillräckligt mycket data för instabila lösningar, och därför blir partiska mot att prediktera noll.

En möjlig förklaring till att de modeller som har fler nätverk (18 eller 36 NN) ger bättre prediktioner är att varje nätverk är mer specialiserat per $k\rho$ -index. Om en modell tränas på samtliga $k\rho$ -index är det möjligt att de index som innehåller en stor andel nollvärden korrumpierar träningen och således försämrar prediktionsförmågan överlag.



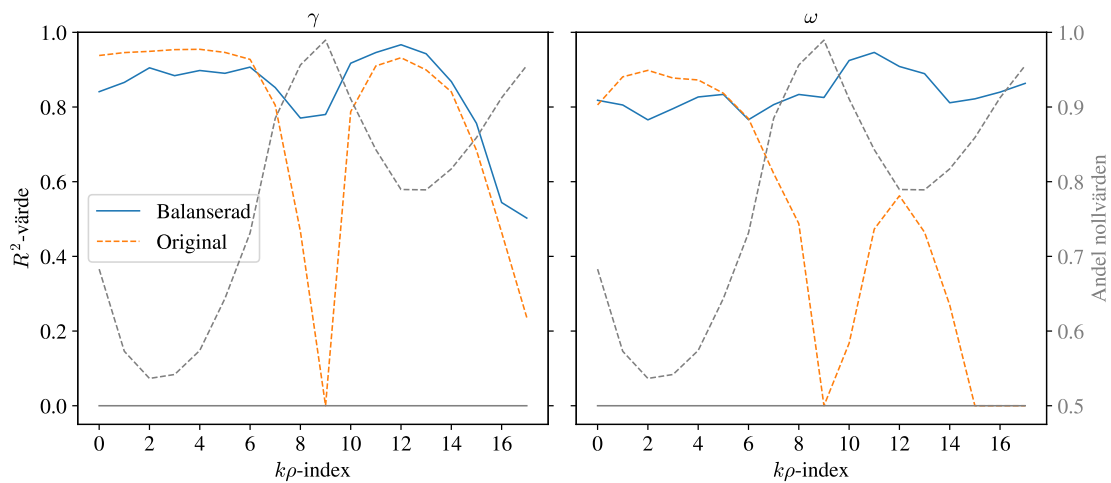
Figur 4.2: R^2 -värde som funktion av $k\rho$ -index för 1 (blå), 2 (orange), 18 (grön) och 36 (röd) neurala nätverk. Diagrammen delas in efter tillväxthastighet γ (t.v.) och realfrekvensen ω (t.h.). En sekundär lodrät axel anger andelen nollvärden i databasen genom en gråstreckad linje.

Tabell 4.1: R^2 -värden för tillväxthastigheten (t.v.) och realfrekvensen (t.h.), beräknade på nollskilda värden och med trunkering för $R^2 < 0$ för fyra olika neurala nätverksmodeller. Gråmarkerade celler indikerar högst R^2 -värde för aktuellt $k\rho$ -index.

$k\rho$ -index	Neural nätverksmodell				$k\rho$ -index	Neural nätverksmodell			
	1 NN	2 NN	18 NN	36 NN		1 NN	2 NN	18 NN	36 NN
0	0,911	0,929	0,937	0,938	0	0,808	0,842	0,815	0,806
1	0,922	0,936	0,950	0,946	1	0,859	0,882	0,890	0,881
2	0,928	0,938	0,953	0,949	2	0,849	0,872	0,901	0,898
3	0,932	0,939	0,955	0,954	3	0,844	0,864	0,886	0,878
4	0,933	0,938	0,952	0,955	4	0,866	0,881	0,881	0,872
5	0,925	0,933	0,946	0,946	5	0,810	0,823	0,837	0,837
6	0,907	0,917	0,928	0,928	6	0,744	0,778	0,770	0,769
7	0,735	0,761	0,818	0,805	7	0,508	0,547	0,608	0,622
8	0,0	0,0	0,525	0,467	8	0,0	0,172	0,428	0,488
9	0,0	0,0	0,0	0,0	9	0,0	0,0	0,0	0,0
10	0,724	0,771	0,776	0,788	10	0,0	0,157	0,165	0,167
11	0,864	0,898	0,892	0,910	11	0,527	0,553	0,503	0,473
12	0,914	0,929	0,912	0,932	12	0,338	0,412	0,593	0,562
13	0,885	0,903	0,866	0,899	13	0,312	0,388	0,509	0,464
14	0,817	0,819	0,816	0,841	14	0,325	0,375	0,311	0,270
15	0,662	0,647	0,692	0,684	15	0,0	0,0	0,0	0,0
16	0,347	0,359	0,460	0,465	16	0,0	0,0	0,0	0,0
17	0,0	0,0	0,248	0,236	17	0,0	0,0	0,0	0,0
Medel	0,561	0,588	0,631	0,629	Medel	0,433	0,475	0,505	0,500

4.2 Införande av balanserade dataset

Surrogatmodellens prestanda efter införandet av ett balanserat träningsset, med lika andelar nollor och icke-nollor för varje $k\rho$ -index, presenteras i figur 4.3. I de två diagrammen, som delats upp efter tillväxthastighet och realfrekvens, visas grafer för R^2 -värdet både efter träning med originaldatasetet och efter träning med det balanserade datasetet. Detta R^2 -värde beräknades endast på de instabila lösningarna eftersom R^2 annars skulle kunna ge de $k\rho$ -index med många nollor en stor fördel. Figuren visar att träning med det balanserade datasetet ger högre R^2 -värden än det ursprungliga datasetet för $k\rho$ -index ≥ 7 , både för tillväxthastigheten och realfrekvensen. Skillnaden är särskilt påtaglig för tillväxthastighetens $k\rho$ -index 8 och 9, och för realfrekvensens $k\rho$ -index 7–17.

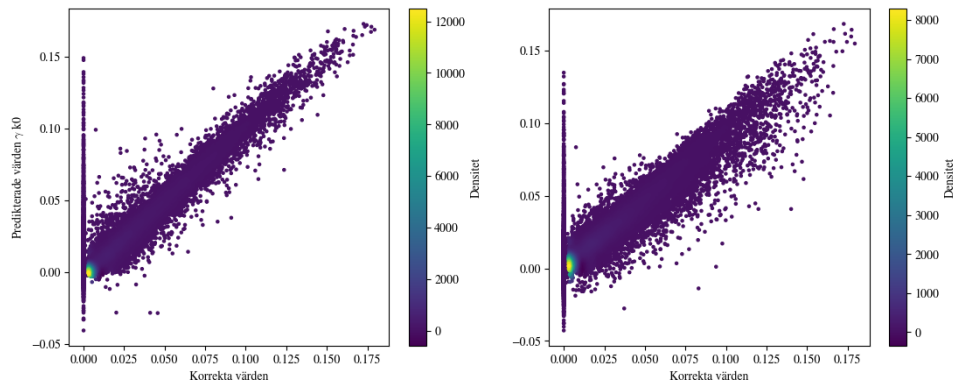


Figur 4.3: R^2 -värde som funktion av $k\rho$ -index för balanserat träningsset (blå) och originaldataset (orange), med 36 neurala nätverk (ett per $k\rho$ -index). Diagrammen delas in efter tillväxthastigheten γ (t.v.) och realfrekvensen ω (t.h.). En sekundär lodrät axel anger andelen nollvärden i dataseten genom en heldragen (balanserat träningsset) respektive streckad (originaldataset) grå linje.

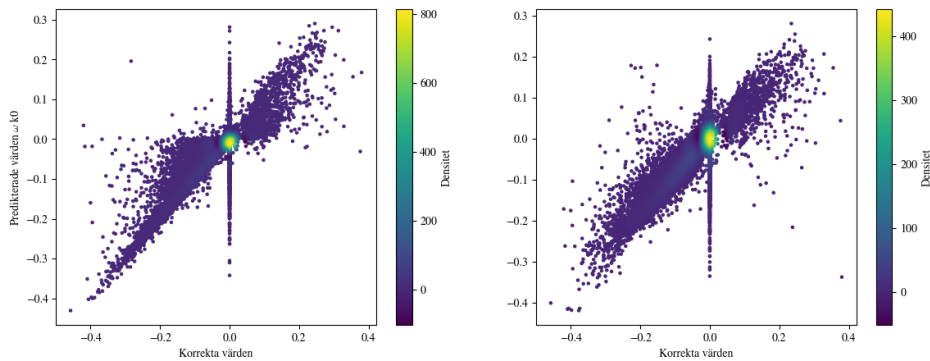
Bland de $k\rho$ -index som förbättrats mest i termer av R^2 , ingår samtliga index som har en mycket hög andel nollvärden i det ursprungliga datasetet. Detta ger stöd åt hypotesen om att träning på en mycket hög andel nollvärden i träningsdatan försämrar modellens prestanda. För andra $k\rho$ -index presterar modellen lite sämre, vilket kan bero på flera orsaker.

Anledningen till att de balanserade dataseten presterar sämre för $k\rho$ -index mindre än 7 skulle kunna vara att hälften stabila och hälften instabila lösningar inte är den mest optimala fördelning av lösningar. Det skulle också kunna bero på slump när det balanserade datasetet skapades. Denna förklaring är dock mindre trolig då varje $k\rho$ -index har ett eget slumpmässigt utvalt dataset och sannolikheten att alla dessa mellan $k\rho$ -index 0 och 7 skulle vara sämre är liten. Att varje $k\rho$ -index har ett eget slumpat dataset är också anledningen till att R^2 -grafnen för det balanserade datasetet är mer oregelbunden.

I figur 4.4 till 4.7 presenteras spridningsdiagram för prediktion av tillväxthastighet och realfrekvens för $k\rho$ -index 0 (se figur 4.4 och 4.5) och 9 (se figur 4.6 och 4.7), efter träning med originaldataset (t.v.) och balanserat dataset (t.h.). Som ses i figur 4.4 och 4.7 är det tydligt att de balanserade dataseten ger mycket bättre resultat än de ursprungliga dataseten. Resterande spridningsdiagram finns samlade i bilaga B.

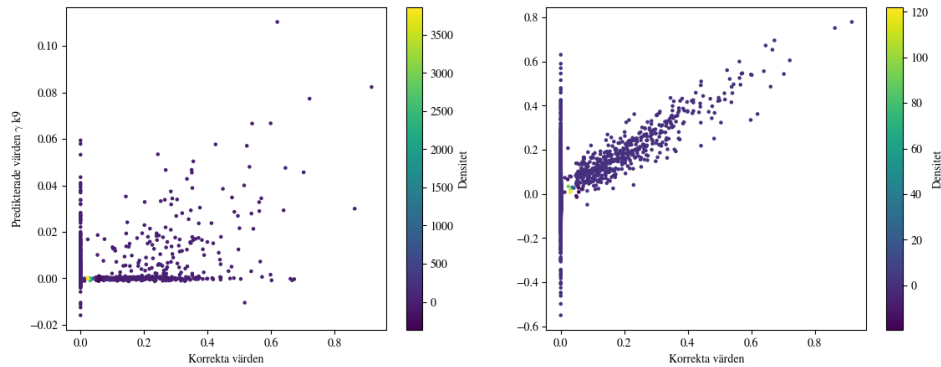


Figur 4.4: Jämförelse mellan spridningsdiagram för $k\rho$ -index 0 för träning med originella samt balanserade träningsset för tillväxthastigheten (γ). De originella dataseten visas i den vänstra figuren och de balanserade visas i den högra.

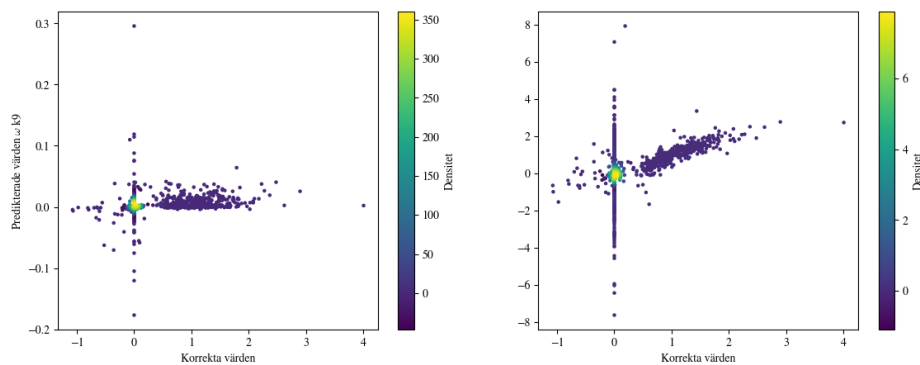


Figur 4.5: Jämförelse mellan spridningsdiagram för $k\rho$ -index 0 för träning med originella samt balanserade träningsset för realfrekvensen (ω). De originella dataseten visas i den vänstra figuren och de balanserade visas i den högra.

Den linje som fanns vid noll för de ursprungliga dataseten finns även här, vilket antingen kan beror på att dessa är krascher av QuaLiKiz som fått ett värde av noll, eller att eftersom det finns en spik för output-distributionen vid noll så kan en sådan linje finnas oavsett krascher eller inte. Det kan också vara en kombination av dessa faktorer.



Figur 4.6: Jämförelse mellan spridningsdiagram för $k\rho$ -index 9 för träning med originella samt balanserade träningsset för tillväxthastigheten (γ). De originella dataseten visas i den vänstra figuren och de balanserade visas i den högra.



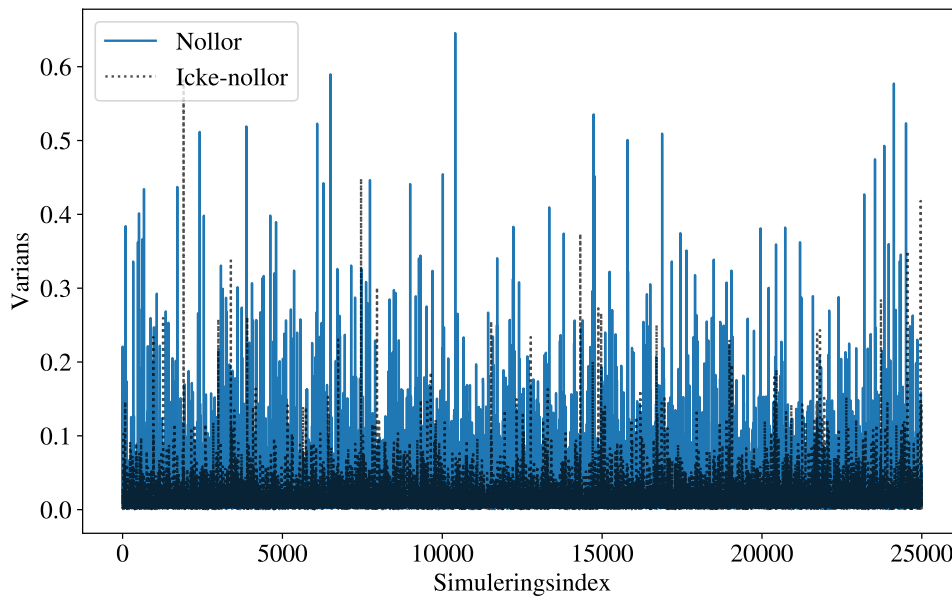
Figur 4.7: Jämförelse mellan spridningsdiagram för $k\rho$ -index 9 för träning med originella samt balanserade träningsset för realfrekvensen (ω). De originella dataseten visas i den vänstra figuren och de balanserade visas i den högra.

4.3 Hantering av kraschade simuleringar

I vissa fall predikterar de neurala nätverken en instabil turbulens (ett nollskilt värde), trots att det korrekta värdet är noll. Sådana avvikelser kan förekomma slumpmässigt i ett begränsat antal prediktioner, men spridningsdiagrammen visar på ett relativt stort antal prediktioner av denna typ (se t.ex. figur 4.4 och axeln med korrekta värden där datapunkterna bildar ett vertikalt streck). Det är möjligt att fenomenet kan förklaras av att det förekommer kraschade simuleringar i databasen. En kraschad simulering tilldelas på samma sätt som en stabil lösning värdet noll av QuaLiKiz. Således skulle de vertikala strecken i spridningsdiagrammen kunna vara spår av att QuaLiKiz ger stabila värden (nollvärden) för kraschade instabila lösningar. Utvecklingarna bakom QuaLiKiz-databasen har dessutom indikerat att dessa krascher kan påverka inlärningen av neurala nätverksmodeller.

4.3.1 Ensemblenätverk

De neurala nätverken i ensemblenätverket predikterade var för sig på QuaLiKiz data och utifrån prediktionerna beräknas variansen. Som ett jämförande mått predikterade nätverken på data den kände igen (instabil turbulens, icke-nollor i figur 4.8), då var medelvariansen låg, 0,0129. När nätverken sedan predikterade på stabil turbulens, vilket för nätverken är okänd data (nollor i figur 4.8) var medelvariansen nästan dubbelt så hög, 0,0245.



Figur 4.8: Varians som funktion av simuleringsindex för prediktion på nollor (blå heldragen linje) respektive icke-nollor (grå prickad linje), efter träning på icke-nollor.

Medelvariansen var betydligt lägre när nätverken predikterade på instabil turbulens, vilket liknade data den var tränad på, än när den predikterade på instabil data som kunde innehålla krascher. Den högre variansen tydde på högre osäkerhet från nätverken vilket skulle varit den stabila datan. Resultatet visar att det går att analysera variansen av ensemblenätverk för att göra en gissning om en simulering är stabil eller inte, givetvis med en viss osäkerhetsmarginal eftersom det exempelvis även finns stabila simuleringar med låg varians.

4.3.2 Vågspektrumens kontinuitet

Analysen av kontinuiteten för ett givet spektrum gav andelen simuleringar som antas vara krascher. Varje spektrum innehåller 18 simuleringar, vilket betyder att med 500 spektrum analyserades 9000 simuleringar. Resultaten från denna analys presenteras i tabell 4.2.

Fördelningen av datapunkter i spridningsdiagrammen (till exempel i figur 4.4), tillsammans med data från tabell 4.2, tyder på att antalet krascher är litet och därför

Tabell 4.2: Tabellen visar hur stor andel av datan som klassificerats som krascher.

Data	Antal spektrum	Antal krascher	Andel krascher (%)
Slumpmässig	500	66	0,73
Lägsta 10 % varians	500	81	0,9
Lägsta 30 % varians	500	75	0,83
Lägsta 50 % varians	500	51	0,57

kan vara svåra att identifiera. Sammantaget är det svårt att se ett tydligt samband mellan varians och andel krascher.

5

Slutsats

I projektet har vi skapat flera olika neurala nätverksmodeller med målet att snabbt prediktera tillväxthastigheten och realfrekvensen för den turbulenta transportmodellen QuaLiKiz. Dessa modeller har haft olika arkitekturer och strukturer och även olika träningsdata och därmed haft olika prestanda. Den modell som gett bäst resultat har ett neuralt nätverk för varje outputvariabel med 250 noder och ett dolt lager, och tränas på data med hälften stabila och hälften instabila lösningar.

I projektet har även krascher i QuaLiKiz-databasen undersökts. För att kunna hitta dessa simuleringar försökte ensemblerätverk användas där variansen av prediktionerna borde vara låg om simuleringen är en krasch. Denna metod gav cirka 23 % fler krascher än att slumpmässigt välja simuleringar med krascher.

5.1 Framtida projekt

I QuaLiKiz är det stor skillnad på en lösning för tillväxthastigheten som har ett värde nära noll och en lösning som har ett värde som är exakt noll, den ena är en mycket svag instabil lösning medan den andra är en stabil lösning. De modeller som har skapats här har inte kunskap om denna skillnad och därför skulle ett extra nätverk som innan bestämmer om det är en stabil eller instabil lösning kunna användas för att sedan endast träna på de instabila lösningarna. Man skulle också kunna sätta en gräns där alla prediktioner under ett visst värde räknas som stabila. En annan möjlighet till förbättring är att dela upp neurala nätverk enligt vilken typ av instabilitet som är dominerande.

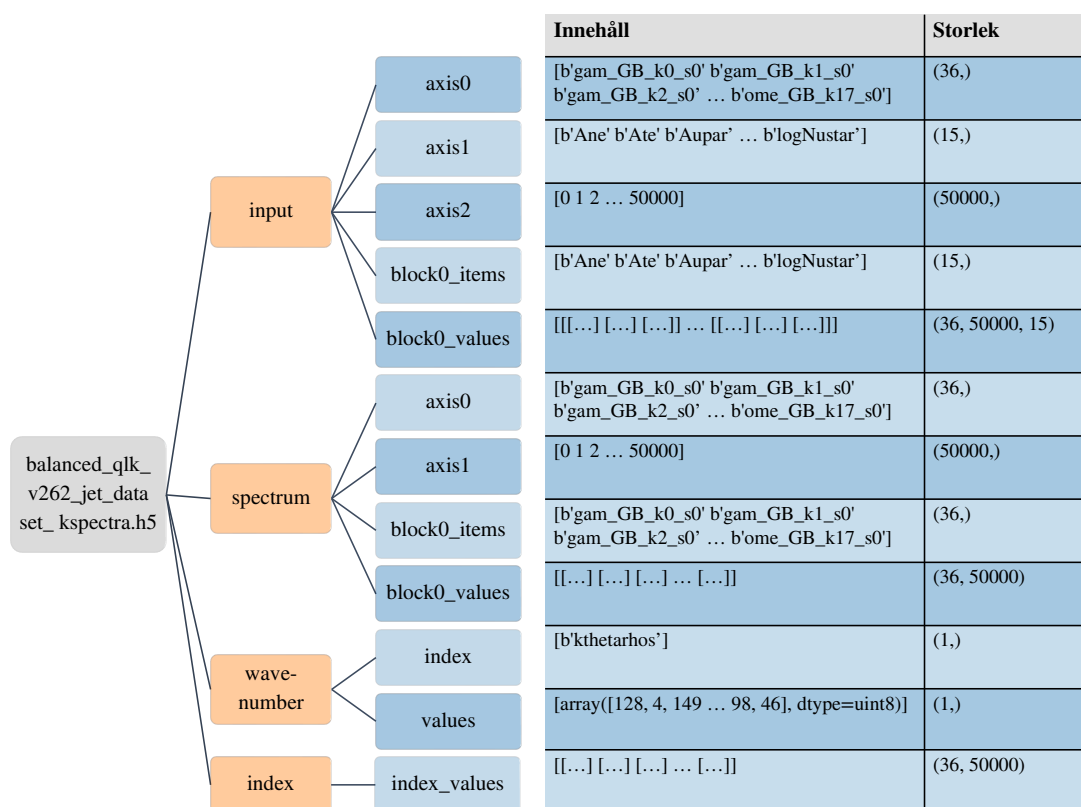
Eftersom tillväxthastigheten i QuaLiKiz inte är negativ skulle en förskjuten ReLU-funktion kunna användas även för output-lagret av den neurala nätverksmodellen. Detta skulle kunna förhindra att den neurala nätverksmodellen predikterar ett negativt värde och därför göra modellen för tillväxthastigheten bättre.

Litteratur

- Bourdelle, C., Citrin, J., Baiocchi, B., Casati, A., Cottier, P., Garbet, X., Imbeaux, F., & Contributors, J. (2015). Core turbulent transport in tokamak plasmas: bridging theory and experiment with QuaLiKiz. *Plasma Physics and Controlled Fusion*, *58*(1), 014036.
- Dan Anderson, P. J., Mietek Lisak, & Marklund, M. (2003). *Basic Plasma Physics Theory and Applications* (tekn. rapport).
- Encyclopædia Britannica ImageQuest. (2023). Superhot Fusion Plasma Inside JET. https://quest.eb.com/search/139_3923673/1/139_3923673/cite (Hämtad: 2023-03-27)
- Gibney, E. (2022). REACTOR SMASHES ENERGY RECORD. *Nature*, *602*, 371.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>. (Hämtad: 2023-05-03)
- Ho, A., Citrin, J., Bourdelle, C., Camenen, Y., Casson, F. J., van de Plassche, K. L., Weisen, H., & Contributors, J. (2021). Neural network surrogate of QuaLiKiz using JET experimental data to populate training space. *Physics of Plasmas*, *28*(3), 032305.
- Jenko, F., Dorland, W., Kotschenreuther, M., & Rogers, B. (2000). Electron temperature gradient driven turbulence. *Physics of plasmas*, *7*(5), 1904–1910.
- Liu, C. S. (1969). Temperature gradient instabilities in axisymmetric systems. *The Physics of Fluids*, *12*(7), 1489–1493.
- Mehlig, B. (2022). *Machine learning with neural networks: An introduction for scientists and Engineers*. Cambridge University Press.
- Nordman, H., & Weiland, J. (1989). Transport due to toroidal η_i mode turbulence in tokamaks. *Nuclear Fusion*, *29*(2), 251.
- S. Li, Z. R., H. Jiang, & Xu, C. (2014). Optimal Tracking for a Divergent-Type Parabolic PDE System in Current Profile Contro. [doi:10.1155/2014/940965](https://doi.org/10.1155/2014/940965)

A

Träddiagram för balanserad databas

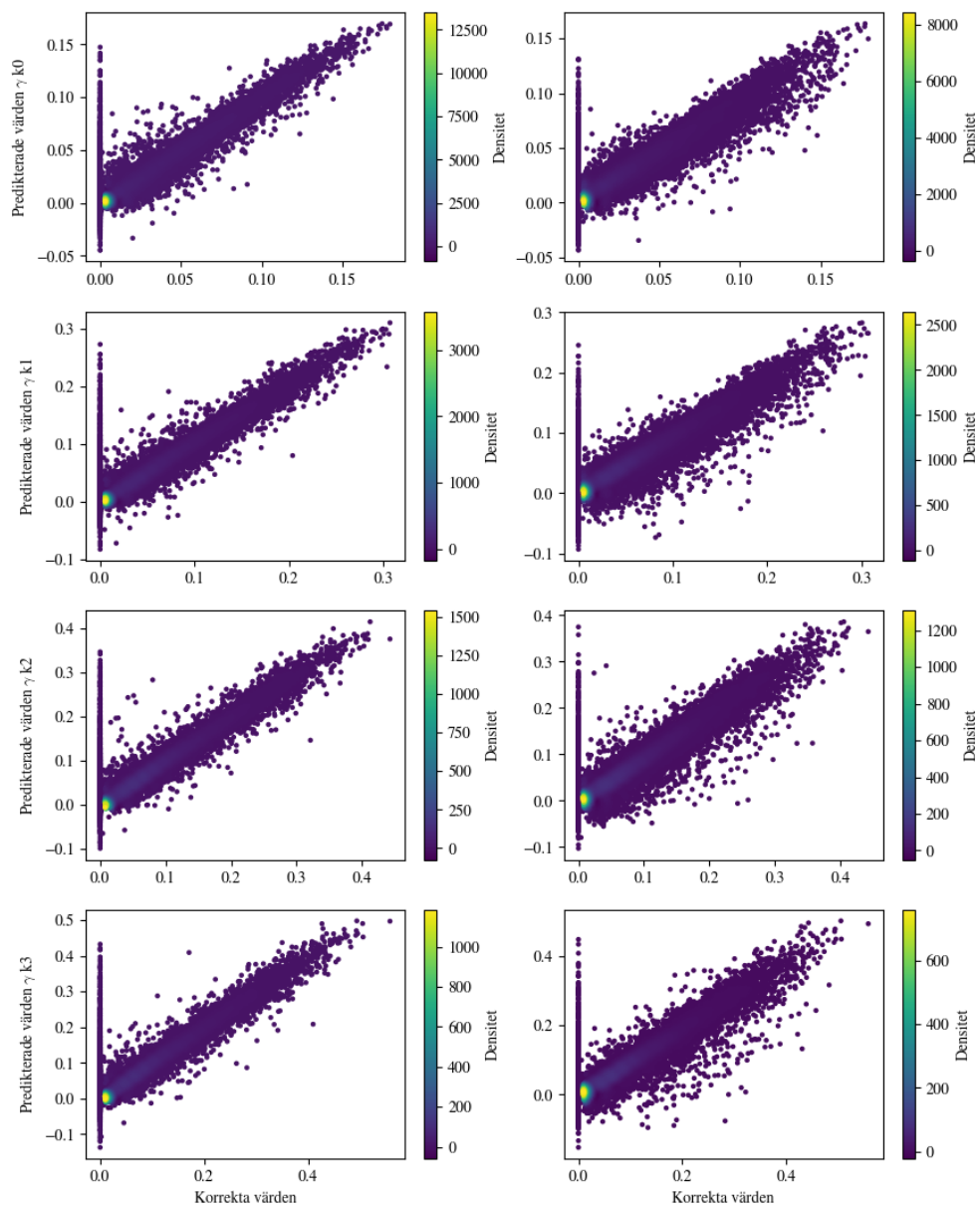


Figur A.1: Träddiagram över den balanserade databasens innehåll. Databasen (`balanced_qlk_v262_jet_dataset_kspectra.h5`) består av fyra grupper: “input”, “spectrum”, “wavenumber” och “index”. Varje grupp innehåller ett antal dataset, till exempel “axis0”, “axis1”, “axis2”, “block0_items” och “block0_values”, vars innehåll och storlek beskrivs av motsvarande rad i tabellen till höger. Notera att `block0_values` har tre dimensioner, då den innehåller en matris för varje krho.

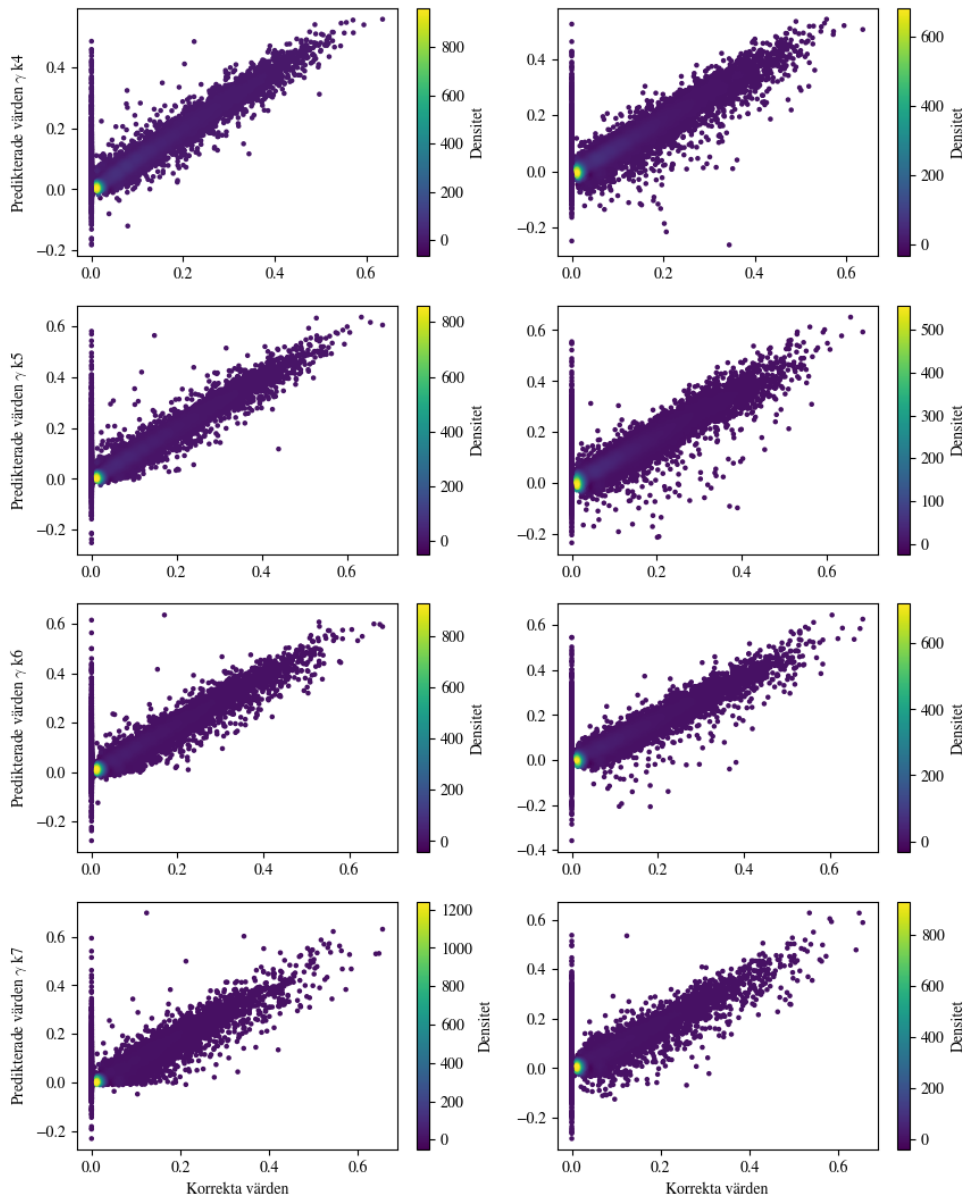
B

Spridningsdiagram för modellprediktioner

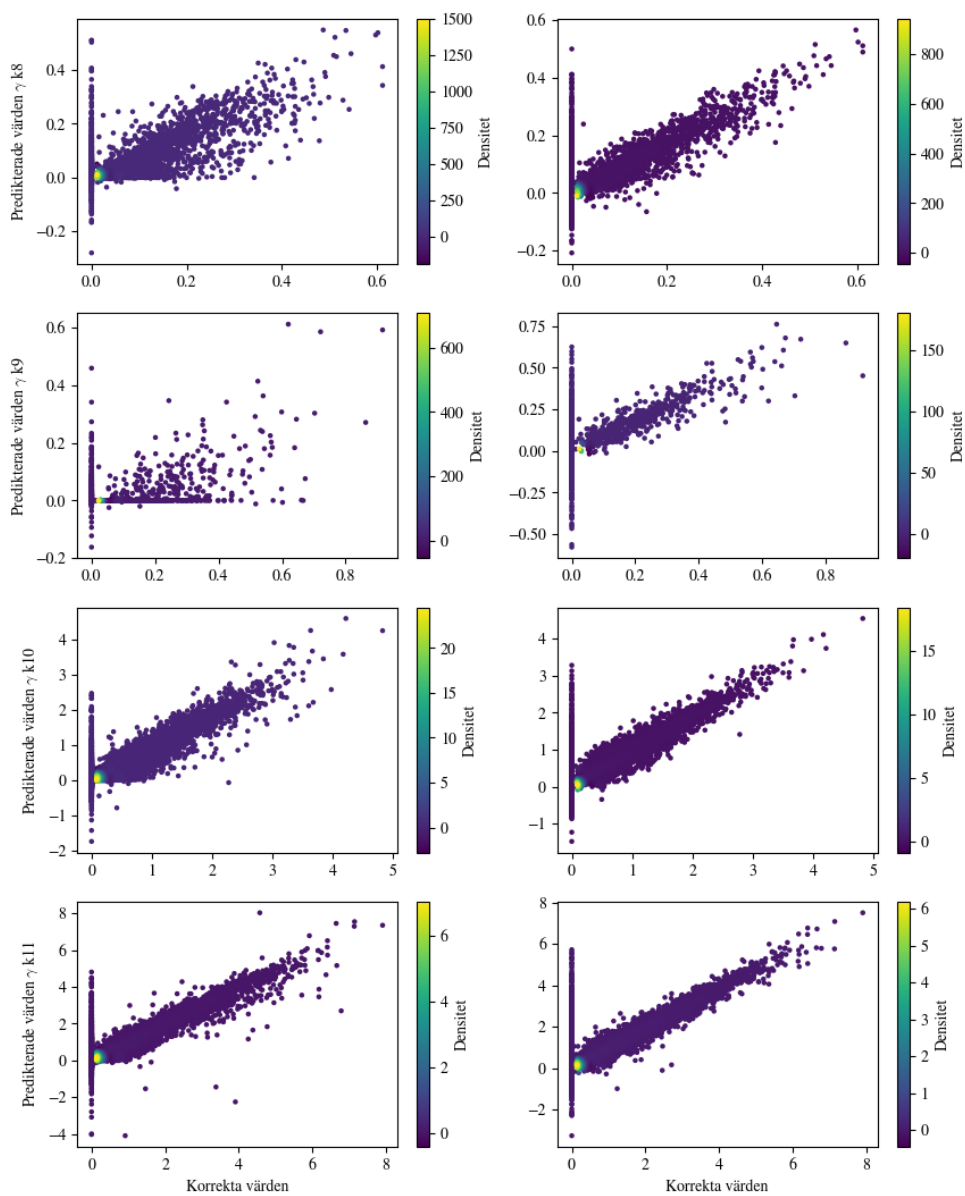
B. Spridningsdiagram för modellprediktioner



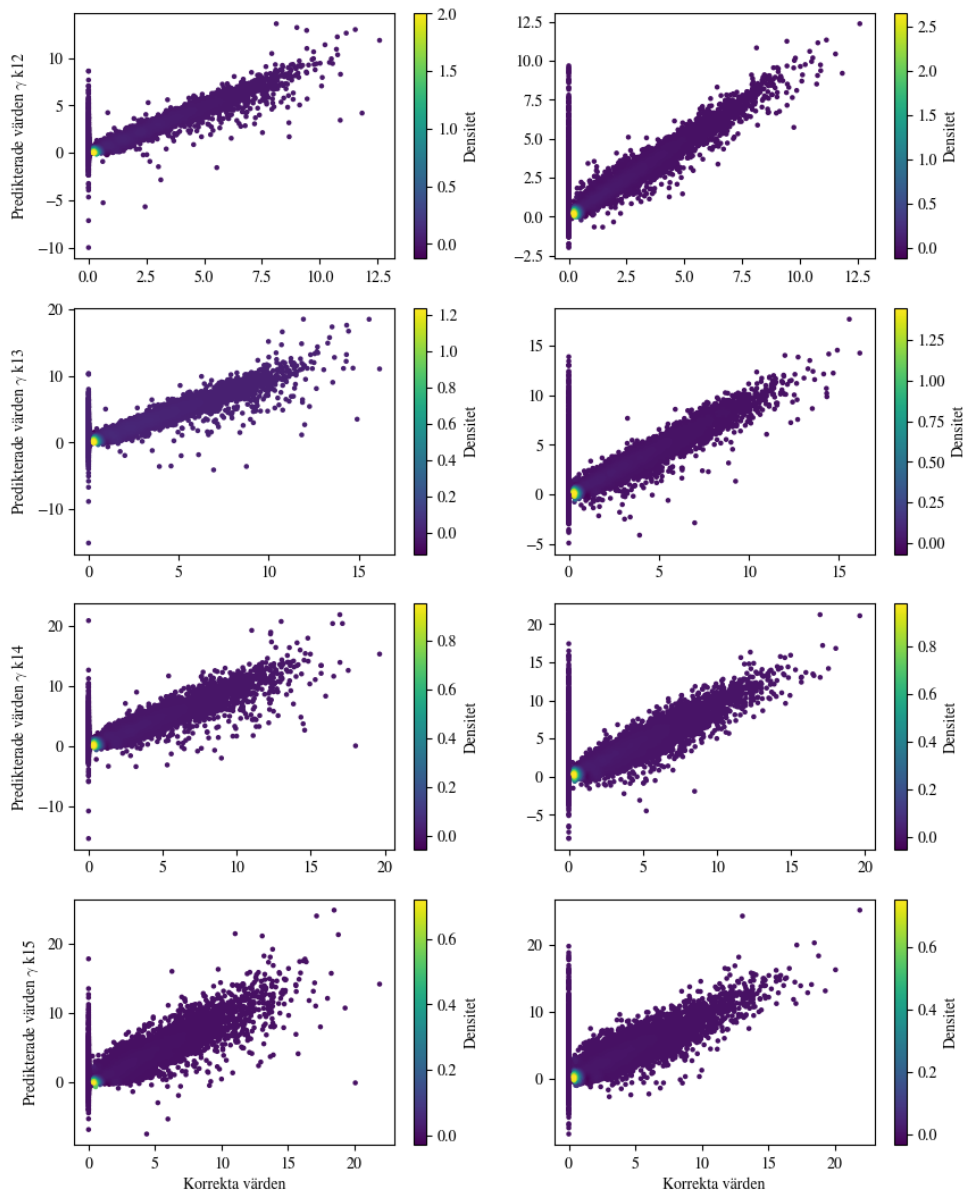
B. Spridningsdiagram för modellprediktioner



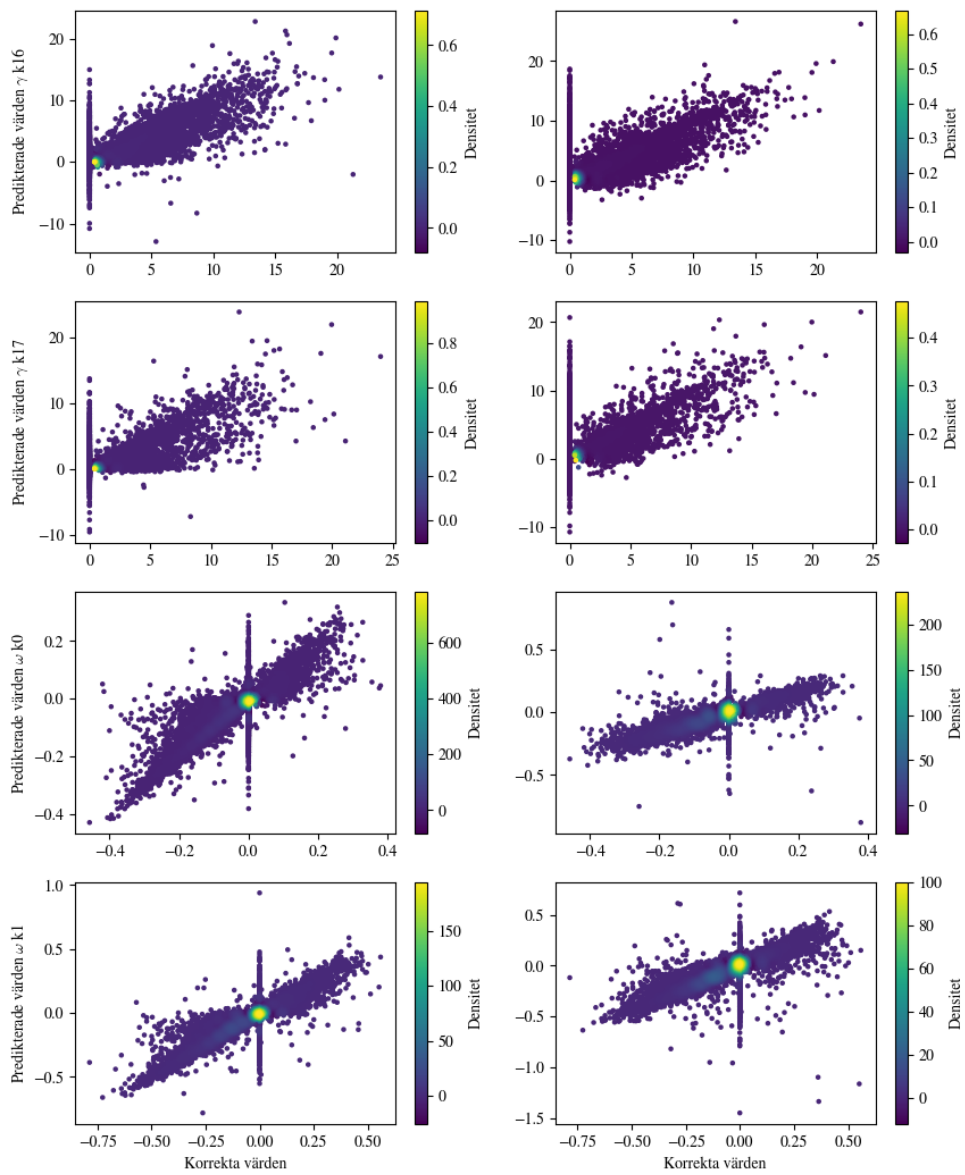
B. Spridningsdiagram för modellprediktioner



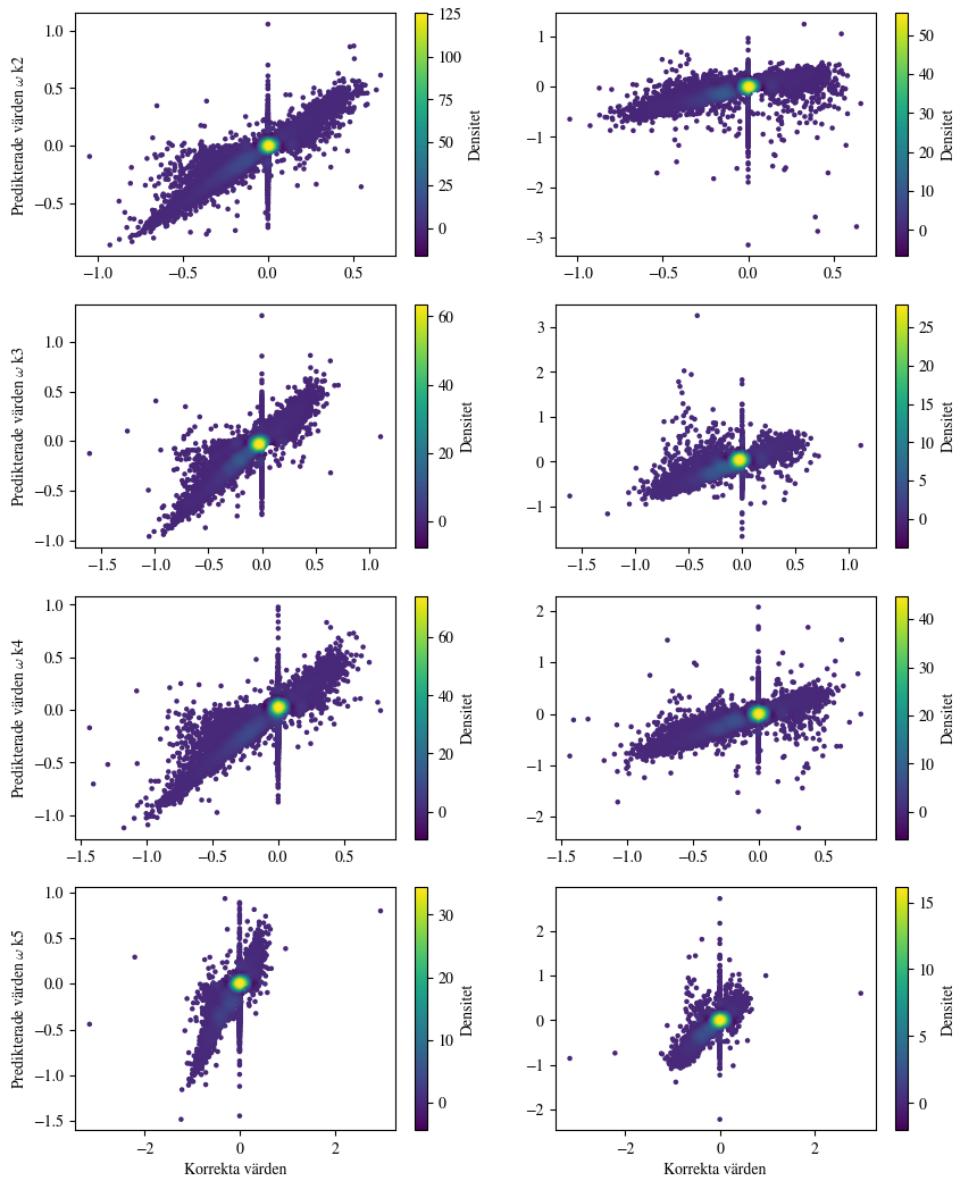
B. Spridningsdiagram för modellprediktioner



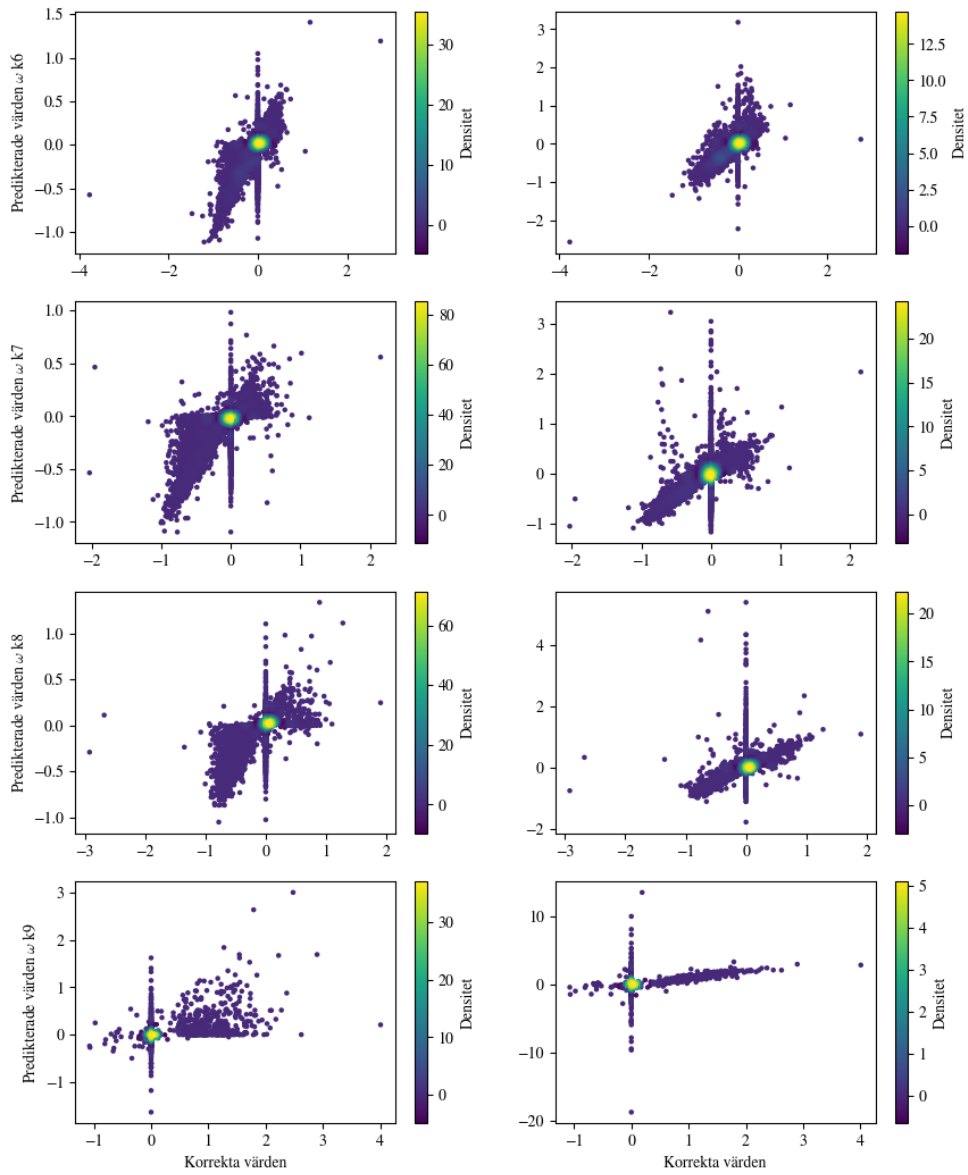
B. Spridningsdiagram för modellprediktioner



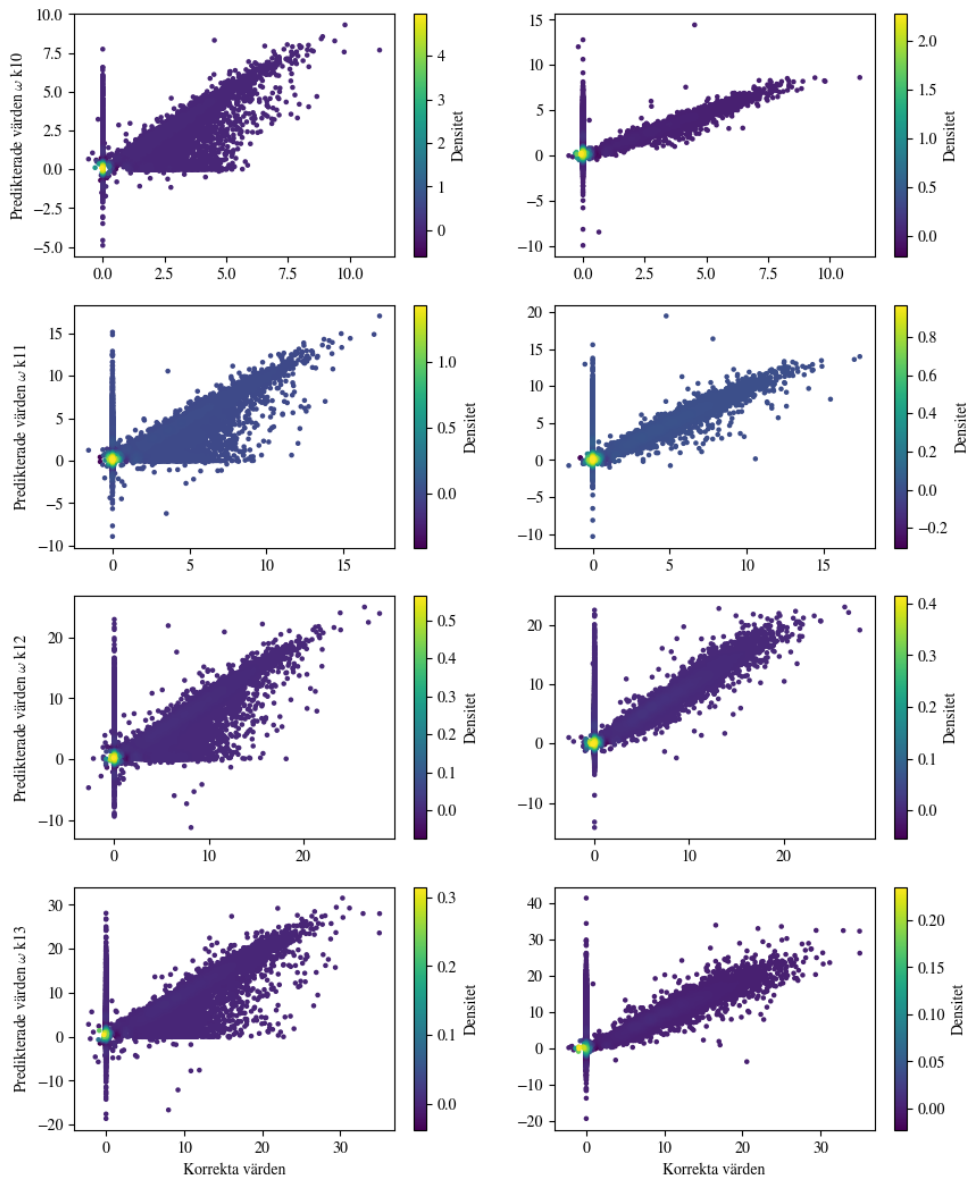
B. Spridningsdiagram för modellprediktioner



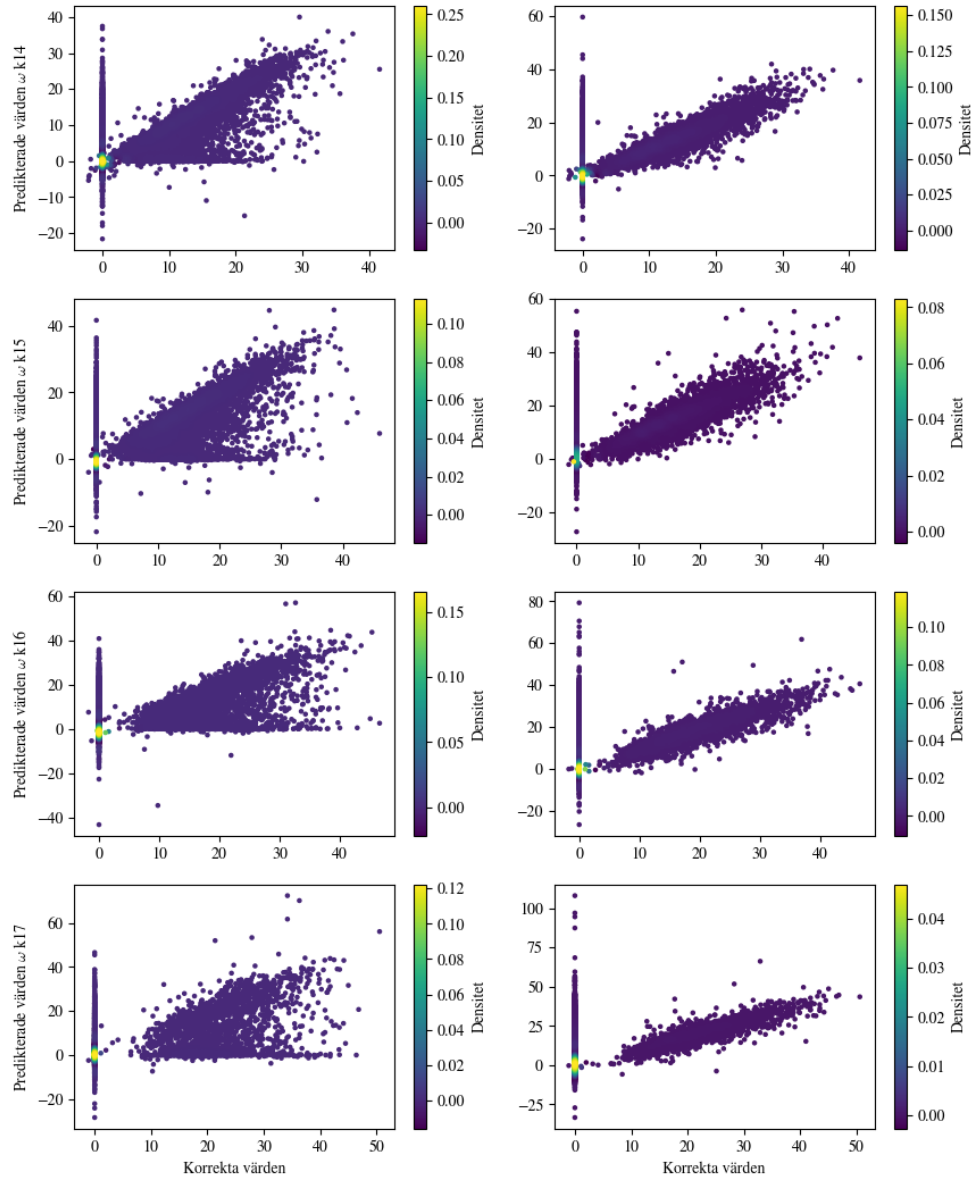
B. Spridningsdiagram för modellprediktioner



B. Spridningsdiagram för modellprediktioner



B. Spridningsdiagram för modellprediktioner



Figur B.1: Spridningsdiagram för alla $k\rho$ -index för både tillväxthastigheten (γ) och realfrekvensen (ω). Figurerna till vänster använder originellt dataset och de till höger använder balanserat dataset

INSTITUTIONEN FÖR RYMD- GEO OCH MILJÖVETENSKAP
CHALMERS TEKNISKA HÖGSKOLA

Göteborg, Sverige

www.chalmers.se



CHALMERS