

Korrelationsutvärdering av lustgasutsläpp på Klagshamns ARV med hjälp av neurala nätverk

INSTITUTIONEN FÖR PROCESSTEKNIK OCH TILLÄMPAD BIOVETENSKAP | LUNDS UNIVERSITET

JONAS TORNKVIST OCH MALTE TÖRNQVIST | EXAMENSARBETE 2026



Korrelationsutvärdering av lustgasutsläpp på Klagshamns ARV med hjälp av neurala nätverk

av

Jonas Tornkvist & Malte Törnqvist

Examensarbete nr: 2026-06

Vattenförsörjnings- och Avloppsteknik
Avdelningen för kemiteknik
Lunds universitet

Maj 2026

Handledare: Åsa Davidsson
Handledare VA SYD: David J.I. Gustavsson
Examinator: Michael Cimbritz

Bild på framsidan: Klagshamns ARV foto: Malte Törnqvist

Postadress
Box 124
221 00 Lund
Hemsida
www.ple.lth.se

Besöksadress
Naturvetarvägen 14

Telefon
046-222 82 85
046-222 00 00

Förord

Detta examensarbete har genomförts inom ämnet Vattenförsörjnings- och avloppsteknik vid Lunds universitet och markerar avslutningen på vår civilingenjörsutbildning i Kemiteknik. Arbetet har varit både lärorikt och utvecklande, och har gett oss möjlighet att fördjupa oss inom ett område som är mycket aktuellt och betydelsefullt ur såväl miljö- som samhällsperspektiv.

Vi vill rikta ett stort tack till avdelningen för kemiteknik vid Lunds universitet för möjligheten att genomföra detta arbete inom ramen för utbildningen. Vi är även mycket tacksamma för det stöd och den vägledning vi fått under arbetets gång.

Ett särskilt stort tack vill vi framföra till David Gustavsson på VA SYD, som varit vår handledare och en mycket värdefull resurs genom hela projektet. Med stort engagemang, tillgänglighet och generositet har han bidragit med kunskap, vägledning och viktiga synpunkter som varit avgörande för arbetets utveckling. Hans stöd har inte bara hjälpt oss framåt i de tekniska och metodmässiga delarna, utan också gett oss en djupare förståelse för de praktiska utmaningar som finns inom avloppsrening.

Vi vill också tacka Sándor Kelemen på VA SYD som varit den processtekniker på Klagshamn som hjälpt oss med alla processtekniska frågor. Sen vill vi tacka Jose Porro för samarbetet med Cobalt Water Group och för att han varit mycket givande och hjälpsam i all diskussion kring vårt arbete.

Vi vill självklart även tacka vår handledare Åsa Davidsson vid Lunds universitet för värdefulla kommentarer och stöd över hela examensarbetet, framförallt rapportskrivandet, samt alla övriga personer som på olika sätt bidragit med hjälp, diskussioner och återkoppling under arbetets gång.

Slutligen vill vi också rikta ett varmt tack till familj och närstående för deras stöd och uppmuntran under hela utbildningen och under arbetet med detta examensarbete.

Lund, Maj 2026

Jonas Tornkvist
Malte Törnqvist

Populärvetenskaplig sammanfattning – Kan en AI minska reningsverkens lustgasutsläpp?

Varje gång du spolrar toaletten börjar en lång resa mot ett reningsverk. Där renas vattnet, men processen har en dold baksida. Den kan nämligen läcka ut lustgas, en växthusgas nästan 300 gånger kraftigare än koldioxid. På Klagshamns reningsverk i Malmö har vi försökt ta reda på varför det händer, och om en AI-modell kan bidra till att sätta stopp för det.

Vi har analyserat stora mängder mätdata från reningsverket och studerat hur olika faktorer hänger ihop med lustgasutsläpp. Mätdata har bestått av ett större antal parametrar i hopp om att lyckas fånga komplexa samband i processen. För att få en bättre bild av vad som faktiskt händer i processen tog vi också hänsyn till att vattnet rör sig genom flera steg i anläggningen som ett paket. Detta innebär att effekter och reaktioner inte direkt uppstår efter ett linjärt mönster. Med bakgrund av data och en mängd fysikalisk samt biokemisk teori har vi skapat en typ av AI-modell som kan lära sig mönster i data. Modellen har gett oss möjlighet att testa olika sätt att driva reningsverket, utan att behöva göra förändringar i verkligheten.

Resultaten pekar på att mängden luft som tillförs processen spelar en stor roll. Mindre luftning i bassängerna indikerar på att man kan få lägre utsläpp av lustgas. Samtidigt får inte luftningen minska för mycket, eftersom det då finns risk att reningen fungerar sämre. Det gäller alltså att hitta en balans där både miljöpåverkan och reningsresultat blir så bra som möjligt. Resultaten visar sig ha varit behjälpliga då den berörda linjen på Klagshamn, med hjälp av arbetet, har börjat drivas med markant lägre luftning.

I framtiden är förhoppningen att metoder av denna art skall kunna bidra till både minskade utsläpp och lägre energianvändning, vilket är viktigt i arbetet mot ett mer hållbart samhälle.

Sammanfattning

Utsläpp av lustgas från avloppsreningsverk (ARV) utgör en betydande klimatpåverkan och är starkt kopplade till utökad biologisk kväveavskiljning. Syftet med detta arbete har varit att analysera sambanden mellan processparametrar och lustgasutsläpp vid Klagshamns avloppsreningsverk samt att undersöka hur datadriven modellering kan användas som stöd för driftoptimering.

Inledningsvis genomfördes en korrelationsanalys där både linjära och monotona samband mellan processvariabler och lustgasutsläpp studerades. En central del av analysen var att införa tidsförskjutningar baserade på hydraulisk uppehållstid, vilket visade sig avgörande för att erhålla processrelevanta samband. Flera kopplingar framträdde först när hänsyn togs till transporttiden genom systemet.

Därefter utvecklades en sekventiell maskininlärningsmodell där aktivslamanläggningens varje zon representerades separat samtidigt som kopplingar mellan zonerna beaktades. Modellen användes som differentiabel komponent i en gradientbaserad optimering, vilket möjliggjorde simulering av hur förändrad luftning påverkar både lustgasutsläpp och utgående ammonium. Utöver den egna analysen bedrevs ett samarbete med Cobalt Water Global som utvecklat en tillämpad modell av liknande typ. Även resultaten från denna pekade i samma riktning som de egna.

Det samlade resultatet är att den nuvarande driften vid Klagshamns ARV sannolikt befinner sig på den höga sidan av det luftningsintervall som minimerar lustgasutsläpp, och att en mer restriktiv luftningsstrategi förefaller möjlig utan att äventyra reningsresultatet. Samtidigt utgör begränsningarna i datamängden, särskilt det i princip konstanta syrebörvärdet under perioden, en grundläggande restriktion. Modellen kan vägleda riktning men inte fastställa optimal driftpunkt.

Arbetet visar att kombinationen av statistiska metoder och maskininläring utgör ett värdefullt verktyg för att förstå komplexa avloppsprocesser, samtidigt som det tydliggör behovet av systematiskt varierad driftdata för att sådana metoder ska kunna ge robust beslutsstöd i praktiken.

Summary

Emissions of nitrous oxide from wastewater treatment plants represent a significant climate impact and are closely linked to enhanced biological nitrogen removal. The aim of this study was to analyze the relationships between process parameters and nitrous oxide emissions at the Klagshamn wastewater treatment plant, and to investigate how data-driven modeling can be used to support operational optimization.

Initially, a correlation analysis was performed in which both linear and monotonic relationships between process variables and nitrous oxide emissions were examined. A central element of the analysis was the introduction of time shifts based on hydraulic retention time, which proved decisive for obtaining process-relevant relationships. Several connections only became apparent once the transport time through the system was taken into account.

Subsequently, a sequential Long Short-Term Memory (LSTM) model was developed in which each zone is represented separately while interactions between zones are also accounted for. The model was used as a differentiable component in a gradient-based optimization, enabling simulation of how altered aeration affects both nitrous oxide emissions and effluent ammonium. A parallel analysis carried out by Cobalt Water Global, who developed a prediction model of similar type, pointed in the same direction as the results obtained in this work.

The overall finding is that current operation at the Klagshamn plant is likely to lie on the high side of the aeration range that minimizes nitrous oxide emissions, and that a more restrictive aeration strategy appears feasible without compromising treatment performance. At the same time, the limitations of the dataset, particularly the essentially constant dissolved oxygen set-point during the studied period, constitute a fundamental restriction. The model can guide the direction of operational changes but cannot determine the optimal operating point.

The study demonstrates that the combination of statistical methods and machine learning offers a valuable tool for understanding complex wastewater treatment processes, while also highlighting the need for systematically varied operational data in order for such methods to provide robust decision support in practice.

Innehållsförteckning

1	Inledning.....	1
1.1	Problemformulering.....	1
1.2	Syfte.....	2
2	Bakgrund	3
2.1	Lustgasbildning i aktivslamprocesser.....	3
2.2	Modellering av lustgasutsläpp	7
2.3	Driftstrategier för lustgasmitigering	8
2.4	Kunskapsluckor	8
3	Metod.....	9
3.1	Aktivslamsystemet vid Klagshamns ARV	9
3.2	Motivering av vald ansats.....	10
3.3	Korrelationsanalys	10
3.4	Maskininlärningsmodell	14
3.5	Optimering.....	17
3.6	Samarbete med Cobalt Water Global	18
3.7	AI-användning	19
4	Resultat och diskussion	21
4.1	Inledande översikt av lustgasutsläpp	21
4.2	Korrelationsanalys	21
4.3	Prediktionsmodell.....	24
4.4	Optimering och simulering.....	28
4.5	Jämförelse med Cobalt Water Global.....	32
4.6	Reglering	34
4.7	Begränsningar.....	35
4.8	Jämförelse med ny data	35
5	Slutsats.....	39
6	Framtida arbete	41
7	Referenser.....	43
8	Appendix	47
A.	Detaljerad LSTM-metod.....	47

B. Detaljerad optimeringsmetod.....	49
C. Korrelationsanalys – Scatterplots.....	50
D. Känslighetsanalyser	69
E. Jämförelse efter huvjustering – zon 2 och zon 5.....	70

1 Inledning

Utsläpp av växthusgaser från avloppsreningsverk har under de senaste åren fått ökad uppmärksamhet inom vatten- och avloppssektorn i takt med att klimatpåverkan blivit en allt viktigare del av hållbarhetsarbetet inom sektorn. En särskilt betydelsefull gas i detta sammanhang är lustgas (N_2O), som har uppvisat en 273 gånger högre klimatpåverkan jämfört med koldioxid räknat per massenhet (IPCC, 2021), och samtidigt utgör en oönskad biprodukt från utökad biologisk kväveavskiljning. Trots att avloppsreningsverk är nödvändiga för att minska utsläpp av näringsämnen till recipienter kan driften av dessa anläggningar alltså också ge upphov till indirekta miljöproblem i form av växthusgasutsläpp. För att kunna minska denna påverkan krävs en bättre förståelse för när, var och varför lustgas bildas i processen.

Lustgas uppkommer främst i samband med nitrifikation och denitrifikation, det vill säga de biologiska processer, förutom assimilation, som används för att omvandla och till slut avlägsna kväve ur avloppsvatten. Tidigare studier visar att lustgasbildningen påverkas av flera samverkande drift- och processparametrar, bland annat löst syre, ammonium- och nitritkoncentrationer, temperatur samt variationer i belastning och internrecirkulation (Kampschreur et al., 2009). Sambanden är ofta komplexa, dynamiska och delvis icke-linjära eftersom flera biologiska och fysikaliska mekanismer verkar samtidigt och dessutom varierar över tid och mellan olika delar av processen. I praktiken innebär detta att lustgasutsläpp kan variera kraftigt både över dygnet och mellan olika zoner i samma reningslinje (Vasilaki et al., 2019). För att kunna identifiera driftstrategier som minskar utsläppen utan att försämra reningsresultatet behövs därför metoder som kan hantera både tidsberoende och komplexa samband mellan flera processvariabler samtidigt.

1.1 Problemformulering

Lustgasutsläpp från biologisk avloppsrening varierar både över tid och mellan olika delar av t ex en aktivslamreningslinje, vilket gör det svårt att identifiera vilka driftförhållanden som faktiskt driver utsläppen. I fullskaliga system beror dessutom senare delar av processen av vad som tidigare skett uppströms. Därför är det viktigt att få en bild över vattenpaketets hastighet vilket beskrivs som den hydrauliska uppehållstiden. Samband mellan processparametrar och lustgasutsläpp riskerar därmed att bli missvisande om samtida mätvärden jämförs utan hänsyn till processens tidsfördröjning.

Traditionella statistiska metoder, såsom korrelationsanalys, kan användas för att identifiera samvariation mellan olika processparametrar och lustgasutsläpp. Samtidigt är deras förmåga begränsad när systemet är dynamiskt, tidsförskjutet och påverkat av flera samverkande variabler. Datadrivna modeller baserade på maskininlärning erbjuder därför ett alternativt angreppssätt där komplexa samband kan beskrivas direkt utifrån historiska processdata.

1.2 Syfte

I detta arbete undersöktes lustgasutsläpp från aktivslamlinje 2 vid Klagshamns avloppsreningsverk (ARV), beläget söder om Malmö. Studien kombinerade tidsförskjuten korrelationsanalys med en zonbaserad maskininlärningsmodell för att undersöka hur processparametrar samvarierar med lustgasutsläpp samt om modellen kan användas för simulering och analys av driftstrategier med lägre mängd utsläpp. Det huvudsakliga vetenskapliga bidraget var utvecklingen och utvärderingen av en zonbaserad modellstruktur som tar hänsyn till aktivslamprocessens seriella uppbyggnad och hydrauliska tidsfördröjning.

Examensarbetets syfte följer:

- Utvärdera hur korrelationerna mellan processparametrar och lustgasutsläpp påverkas av hydraulisk uppehållstid och tidsförskjutning.
- Utröna vilka variabler som uppvisar starkast samband med utsläppen i olika zoner.
- Ta reda på i vilken utsträckning en zonbaserad maskininlärningsmodell kan prediktera lustgasutsläpp.
- Utvärdera om modellen kan användas som indikation för hur förändrad luftning påverkar lustgasutsläppen utan att försämra kväveavskiljningen.

2 Bakgrund

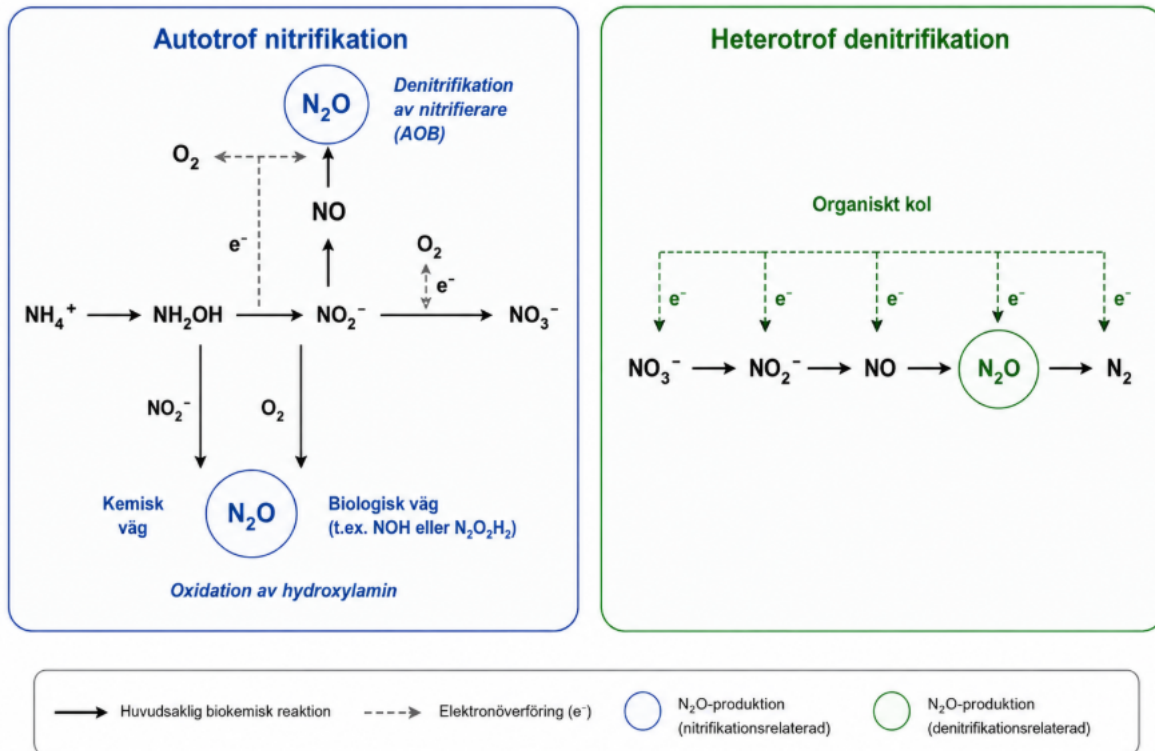
Detta kapitel ger den teoretiska och vetenskapliga grunden för arbetet. Inledningsvis följer en genomgång av de biokemiska och fysikaliska mekanismer som ligger bakom lustgasbildning, samt en översikt över de processfaktorer och driftförhållanden som påverkar utsläppen. Vidare behandlas tidigare arbeten kring modellering av lustgasutsläpp och kapitlet avslutas med en genomgång av utsläppsminskande driftstrategier samt en identifiering av de kunskapsluckor som motiverar föreliggande studie.

2.1 Lustgasbildning i aktivslamprocesser

För att fullt kunna förstå vad som ska modelleras i detta projekt är det viktigt att klargöra vilka mekanismer som ligger bakom lustgasutsläppen. De processparametrar som valts i analysen är hämtade från det som enligt litteraturen uppvisar starkast samband med lustgasbildning, och dessa samband behandlas i detalj nedan.

2.1.1 Biokemiska och kemiska reaktionsvägar

Lustgas bildas från flera reaktionsvägar som illustreras i figur 2 och är en intermediär eller biprodukt i både nitrifikation och denitrifikation. Bildningen är starkt kopplad till driftförhållanden som påverkar mikroorganismernas metabolism. Nitrifikation är en aerob autotrof process där ammonium (NH_4^+) oxideras i två steg. I det första steget oxiderar ammoniakoxiderande bakterier (AOB) ammonium till nitrit (NO_2^-) via hydroxylamin (NH_2OH) som intermediär, en reaktion som katalyseras av enzymerna ammoniakmonooxygenas (AMO) och hydroxylaminoxidoreduktas (HAO). Vid denna oxidation kan lustgas (N_2O) bildas direkt via kemisk eller biologisk oxidation av hydroxylamin (NH_2OH), vilket utgör två av de huvudsakliga reaktionsvägarna för lustgasbildning hos AOB. I det andra steget oxiderar nitritoxiderande bakterier (NOB) nitrit vidare till nitrat (NO_3^-) (Schreiber et al., 2012; Wunderlin et al., 2012). Utöver AOB och NOB kan även ammoniakoxiderande arkéer (AOA) bidra till nitrifikationen i aktivslamanläggningar, som likt AOB oxiderar ammonium till nitrit men tillhör domänen arkéer istället för bakterier. Dessutom finns så kallade comammox-bakterier (complete ammonia oxidizers), det vill säga vissa Nitrospira-arter som ensam utför fullständig nitrifikation från ammonium till nitrat (Daims et al., 2015). Vid stabila aeroba förhållanden sker reaktionen effektivt och med begränsad bildning av intermediärer. När syrehalten däremot blir låg eller fluktuerande kan AOB istället reducera nitrit via så kallad nitrifierande denitrifikation, vilket leder till ökad produktion av lustgas (Wunderlin et al., 2012).



Figur 1. Schematisk översikt av de fyra huvudsakliga reaktionsvägarna för lustgasbildning (N_2O) i aktivslamprocessen, tre biokemiska och en kemisk. Till vänster visas autotrof nitrifikation, där ammonium (NH_4^+) oxideras stegvis till nitrat (NO_3^-) via hydroxylamin (NH_2OH) och nitrit (NO_2^-). N_2O kan här bildas både kemiskt och biologiskt via hydroxylamin, samt biologiskt genom nitrifierande denitrifikation av AOB. Till höger visas heterotrof denitrifikation, där NO_3^- stegvis reduceras till kvävgas (N_2) via NO_2^- , NO och N_2O med organiskt kol som elektronkälla. Heldragna pilar markerar huvudsakliga biokemiska reaktioner och streckade pilar markerar elektronöverföring.

Denitrifikation är en anoxisk process där heterotrofa bakterier reducerar nitrat och nitrit stegvis till kvävgas (N_2) med hjälp av organiskt kol som elektronkälla. Reaktionskedjan omfattar reduktion av NO_3^- till NO_2^- via nitratreduktas (Nar/Nap), vidare till kvävemoxid (NO) via nitritreduktas (Nir), till lustgas (N_2O) via NO -reduktas (Nor) och slutligen till N_2 via N_2O -reduktas (Nos). Under optimala anoxiska förhållanden fullföljs reaktionen till N_2 , vilket innebär att lustgas endast förekommer i mycket små mängder. Om processen däremot störs, exempelvis genom syreinträngning, brist på kolkälla eller höga nitritkoncentrationer, kan det sista reaktionssteget hämmas så att kedjan avstannar vid lustgassteget och betydande lustgasutsläpp uppstår (Schreiber et al., 2012).

Den kritiska kopplingen mellan nitrifikation och denitrifikation uppstår i miljöer där löst syre förekommer i låga eller instabila koncentrationer samtidigt som ammonium och nitrit är tillgängliga. I sådana övergångszoner kan båda processerna pågå parallellt, ofta inom samma

slamflock eller biofilm, där aeroba mikrozooner existerar nära anoxiska kärnor. Dessa så kallade simultana nitrifikations- och denitrifikationsmiljöer (SND) har visat sig vara särskilt benägna att producera lustgas (Kampschreur et al., 2009).

Vilken av reaktionsvägarna som dominerar beror starkt på driftpunkten. I batchexperiment med aktivslam från kommunalt avloppsvatten (Wunderlin et al., 2012) har det visats att hydroxylaminoxidationen var den viktigaste lustgaskällan vid hög ammoniumhalt och låg nitrithalt, medan nitrifierande denitrifikation av AOB tog över när nitrithalten ökade och ammoniumhalten sjönk. I en undersökning av kemiska och biologiska vägar för bildning och konsumering av lustgas och kvävemonoxid (Schreiber et al., 2012) sammanfattades att samtliga organismer som ingår i den katabola delen av kvävecykeln har potential att bilda lustgas via reduktion av nitrit till kvävemonoxid och vidare till lustgas, medan bildningen via hydroxylaminoxidation är begränsad till AOB. Detta innebär att även små förskjutningar i driftpunkt kan ändra både var i processen lustgasen bildas och vilken mekanism som dominerar.

Höga nitritkoncentrationer utgör en av de starkaste drivkrafterna för lustgasbildning. Nitrit fungerar både som substrat för denitrifikation och som elektronacceptor vid nitrifierande denitrifikation hos AOB (Colliver et al., 2000). Vid ackumulering av nitrit, exempelvis till följd av begränsad nitritoxidation eller obalanserad processtyrning, ökar därför risken för kraftigt förhöjda lustgasutsläpp. Samtidigt kan snabba förändringar i syrehalt, ofta orsakade av belastningsvariationer eller intern recirkulation, skapa förhållanden där mikroorganismer växlar mellan aeroba och anoxiska nedbrytningsvägar, vilket ytterligare ökar lustgasproduktionen (Kampschreur et al., 2009).

2.1.2 Fysikaliska och hydrodynamiska faktorer

Utöver de biokemiska reaktionsvägarna spelar även fysikaliska och hydrodynamiska effekter en roll i lustgasens bildning och utsläpp. Särskilt luftningsintensiteten har visat sig vara en kritisk styrparameter, då både för låg och för hög syretillförsel kan leda till ökade lustgasutsläpp.

En direkt effekt av kraftig luftning är så kallad fysisk stripping, där luftbubblor effektivt transporterar lösta gaser från vätskefasen till atmosfären. Detta innebär att lustgas som bildats i exempelvis slamflockar snabbt avlägsnas från vattnet innan den hinner reduceras vidare till kvävgas. Därmed minskar uppehållstiden för lustgas i systemet, vilket begränsar möjligheten för fullständig denitrifikation (Ahn, et al., 2008; Kampschreur et al., 2009; Law et al., 2012).

Hög luftningsintensitet kan dessutom ge upphov till skjuvkrafter som påverkar slamflockarnas struktur (Feng et al., 2019; Jin & Lant, 2004). I intakta flockar uppstår naturligt syregradienter där aeroba zoner samexisterar med anoxiska mikrozooner i flockens inre, och dessa mikrozooner är en förutsättning för att denitrifikationen ska kunna fortlöpa (Li & Bishop, 2004). När syre tränger djupare in i flockarna till följd av turbulens, eller när flockstrukturen förändras, kan omfattningen av anoxiska mikrozooner minska, vilket påverkar förutsättningarna för fullständig denitrifikation och har associerats med ökad lustgasbildning (Kampschreur et al., 2009).

Även nitrifikationsprocessen kan påverkas negativt av kraftig luftning. Trots att nitrifierande bakterier är beroende av syre är de känsliga för snabba förändringar i sin omgivning. Hög turbulens och förändrade mikrobiella förhållanden kan skapa instabila driftförhållanden där AOB:s metabolism påverkas, vilket kan öka sannolikheten för lustgasbildning under ammoniakoxidationen (Kemmu & Amanatidou, 2023; Schreiber et al., 2012).

2.1.3 Drivkrafter och processpåverkan

Utöver de grundläggande reaktionsvägarna finns en rad processparametrar som tillsammans avgör hur stor lustgasproduktionen blir vid en given driftpunkt. Förståelsen för dessa drivkrafter är central för det fortsatta arbetet, eftersom det är dessa parametrar som senare ingår i korrelationsanalysen och i prediktionsmodellen. I sin översikt över utsläpp från avloppsreningsverk identifierade Kampschreur et al. (2009) tre operationella parametrar som genomgående uppvisade starkast samband med förhöjda lustgasutsläpp: låg syrehalt i nitrifikations- och denitrifikationssteget, förhöjda nitritkoncentrationer i båda stegen, samt låg COD/N-kvot i denitrifikationssteget. Snabba förändringar i processförhållanden, exempelvis belastningstoppar eller övergångar mellan aerob och anoxisk drift, utgör en fjärde, ofta lika betydande drivkraft.

Sambandet mellan löst syre och lustgas har en u-formad karaktär som är central för hur driftpunkten kan optimeras. Vid mycket låga nivåer av löst syre (DO) uppstår lokal syrebrist som aktiverar nitrifierande denitrifikation hos AOB, vilket ökar bildningen av lustgas (Kampschreur et al., 2009). Vid höga nivåer av löst syre verkar flera mekanismer samtidigt. Den höga metaboliska aktiviteten hos AOB kan ge ökad lustgasbildning via hydroxylaminoxidation (Wunderlin et al., 2012), samtidigt som syre tränger djupare in i slamflockarna och begränsar de anoxiska mikrozonerna där lustgas annars reduceras vidare till kvävgas (Li & Bishop, 2004). Det höga luftflödet bidrar dessutom till att redan bildad lustgas strippas från vätskefasen innan den hinner reduceras (Kampschreur et al., 2009). Mellan dessa ytterligheter finns en driftpunkt där det totala utsläppet är som lägst, men dess exakta läge är anläggningsspecifikt och beror på flockstruktur, belastning och temperatur. I pilotstudier har syrebörvärdet visat sig vara den enskilt viktigaste styrparametern för lustgasutsläpp (Lotito et al., 2012), och liknande slutsatser drogs vid fullskala där en sänkning av syrebörvärdet från en högre baslinje till en optimerad nivå minskade utsläppsfaktorn med 35 % (Duan et al., 2020).

Belastningstoppar och recirkulationsflöden skapar dessutom snabba processförändringar där bakteriernas metabolism inte hinner anpassa sig, vilket leder till tillfällig ackumulering av nitrit och hydroxylamin. Sådana övergångsförlopp har vid fullskaliga anläggningar visat sig stå för en oproportionerligt stor del av det totala årliga utsläppet. Lotito et al. (2012) observerade exempelvis vid en pilotanläggning att större delen av dygnsutsläpp koncentrerades till en klocklik topp under förmiddagens belastningsmaximum, korrelerad till ammonium- och nitrittoppar i bassängen. Lustgasens typiska utsläppsfaktorer för konventionella aktivslamanläggningar med biologisk kväveavskiljning ligger enligt sammanställningar i intervallet 0,5–2 % av den ingående kvävebelastningen, med betydligt högre värden vid sidoströmsbehandling av rejektvatten från avvattnings av rötat slam (Kampschreur et al., 2009; Vasilaki et al., 2019).

2.1.4 Variation i fullskaliga anläggningar

Mycket av kunskapen om lustgasbildning kommer från labbstudier, där processförhållandena är välkontrollerade och tydliga samband kan etableras. I full skala är bilden mer komplex, eftersom belastningsvariationer, säsongseffekter och anläggningsspecifika utformningar leder till stor variation i uppmätta utsläppsfaktorer.

I en kritisk genomgång av ett decennium av lustgas-mätningar vid fullskaliga avloppsreningsverk visade Vasilaki et al. (2019) att den rapporterade utsläppsfaktorn varierar stort mellan studier, från under 0,5 % till över 5 % av N-belastningen. En del av denna spridning förklaras av processkonfiguration, men en lika viktig faktor är mätkampanjens längd. Studier som mätte under mindre än en månad rapporterade medelvärdet för utsläppsfaktorer om ca 0,7 % av N-belastningen, medan studier som spände över ett helt år hade ett medelvärde på 1,7 %. Skillnaden beror på att utsläpp från fullskaliga avloppsreningsverk uppvisar starka säsongsvariationer och episodiska toppar som korta kampanjer ofta missar. Vid det fullskaliga verket i Amsterdam West rapporterade Seshan et al. (2024) exempelvis att 54 % av den årliga lustgasmassan släpptes ut under mars och april. Dessa observationer talar för att sammanställningar som baseras på korta kampanjer tenderar att underskatta de verkliga utsläppen och försvårar jämförelser mellan anläggningar.

2.2 Modellering av lustgasutsläpp

Modellering av lustgasutsläpp i aktivslamprocesser är ett aktivt forskningsfält där två huvudriktningar dominerar: mekanistiska modeller som bygger på fysikaliska och biokemiska samband, och datadrivna modeller som lär sig samband direkt från mätdata. Mekanistiska modeller utgår vanligen från utvidgningar av den klassiska aktivslammodellen (Henze et al., 2000), och beskriver kinetiken hos de tre huvudsakliga lustgasvägarna explicit (Massara et al., 2017). Deras främsta styrka är fysikalisk tolkbarhet, men de är ofta överparameteriserade, känsliga för ändringar i processförhållanden och väldigt krävande att kalibrera och validera (Mehrani et al., 2022). Datadrivna ansatser, baserade på maskininlärning, undviker detta genom att lära sig samband direkt från drift- och utsläppsdata. I en översikt över området konstaterar Fang et al. (2026) att support vector machines, random forest och artificiella neurala nätverk är de vanligast tillämpade algoritmerna, och att dessa typiskt uppnår R^2 -värden över 0,90. För prognoser över längre tidshorisonter har återkopplade nätverksarkitekturer visat sig särskilt lämpade. Long Short-Term Memory-nätverk (LSTM) är en typ av rekurrent neuron-nätverk som genom interna cell- och grindstrukturer kan lära sig långsiktiga beroenden i sekventiell data (Hochreiter & Schmidhuber, 1997). LSTM har tillämpats med god framgång på lustgasdata vid fullskaliga reningsverk, både för soft-sensor-implementation (Hwangbo et al., 2021) och för flerstegsprognos avsedd att integreras i prediktiv reglering (Seshan et al., 2024). De datadrivna modellernas främsta begränsningar är svårtolkad inre struktur, stora datakrav vid träning samt begränsad förmåga att extrapolera utanför den driftdomän som modellen tränats på.

2.3 Driftstrategier för lustgasmitigering

Förståelse av reaktionsmekanismer och drivkrafter blir praktiskt relevant först när den kan översättas till driftbeslut. Den enskilt viktigaste styrparametern för lustgasutsläpp i aerob aktivslamdrift är luftningen, eftersom den både påverkar de biokemiska bildningsvägarna och den fysiska stripping-effekten. Sänkning av syrebörvärdet har i flera studier visats minska lustgasutsläppen utan att äventyra reningsresultatet. Vid en fullskale-SBR i Adelaide implementerade Duan et al. (2020) permanent en sådan strategi och uppnådde 35 % lägre utsläppsfaktor (från 0,89 % till 0,58 % av N-belastningen) samtidigt som luftningens energiåtgång minskade med 20 %. Sambandet är dock icke-monotont, och vid alltför låga syrenivåer aktiveras nitrifierande denitrifikation hos AOB så att utsläppen ökar igen. Den optimala driftpunkten är därför anläggningsspecifik och måste identifieras genom kombination av mätningar, modellering och kontrollerade driftexperiment.

Den syrereglering som beskrivs i avsnitt 2.1 är ett exempel på kaskadreglering, där flera regulatorer kopplas i serie så att en yttre regulators utsignal utgör börvärde för en inre regulator. I avloppsreningsverk är ammoniumbaserad luftningsstyrning i kaskad med syre- och luftflödesregulatorer en vedertagen lösning (Åmand, 2013). Varje regulator i kaskaden är typiskt utformad som en PI- eller PID-regulator. I tillämpningar där processdynamiken är mer komplex har även modellbaserad prediktiv reglering (MPC) studerats. MPC använder en processmodell för att förutsäga framtida tillstånd och optimera styråtgärder över en glidande tidshorisont, och har visats kunna minska både energianvändning och lustgasutsläpp i avloppsreningsverk (Debel et al., 2025). Inom samma utvecklingslinje har LSTM-baserade prognosmodeller föreslagits som indata till framtida MPC-ramverk för lustgasstyrning (Seshan et al., 2024).

2.4 Kunskapsluckor

Mot denna bakgrund framträder flera kvarstående kunskapsluckor som föreliggande arbete adresserar. För det första visar litteraturen att lustgasutsläpp från fullskaliga anläggningar uppvisar stark anläggningsspecifik variation som korta mätkampanjer inte fångar (Vasilaki et al., 2019), vilket motiverar studier baserade på långa, sammanhängande dataserier från enskilda verk. För det andra har de hittills tillämpade datadrivna ansatserna huvudsakligen behandlat anläggningen som en svart låda och predikerat aggregerade utsläpp (Hwangbo et al., 2021; Seshan et al., 2024). Den seriella, zonindelade strukturen i en pluggflödesreaktor som Klagshamns linje 2 utgör, medför dock att olika zoner befinner sig i olika driftpunkter samtidigt och därmed bidrar olika mycket till de totala utsläppen. En modell som fångar denna zonvisa skillnaden har potential att ge mer mekanistisk insikt än en aggregerad modell, samtidigt som den undviker den fullständiga parameterkomplexiteten i en mekanistisk ASM-utvidgning. För det tredje har tidigare optimeringsstudier främst fokuserat på isolerade enskilda utsläppsminskande strategier, eller bedrivits som rena simuleringsstudier utan koppling till verklig driftdata. Kombinationen av tidsförskjutet korrelationsanalys, zonvis prediktionsmodell och optimering mot fullskaliga driftdata vid ett svenskt verk är därmed, såvitt vi har kunnat finna i litteraturen, en outforskad ansats.

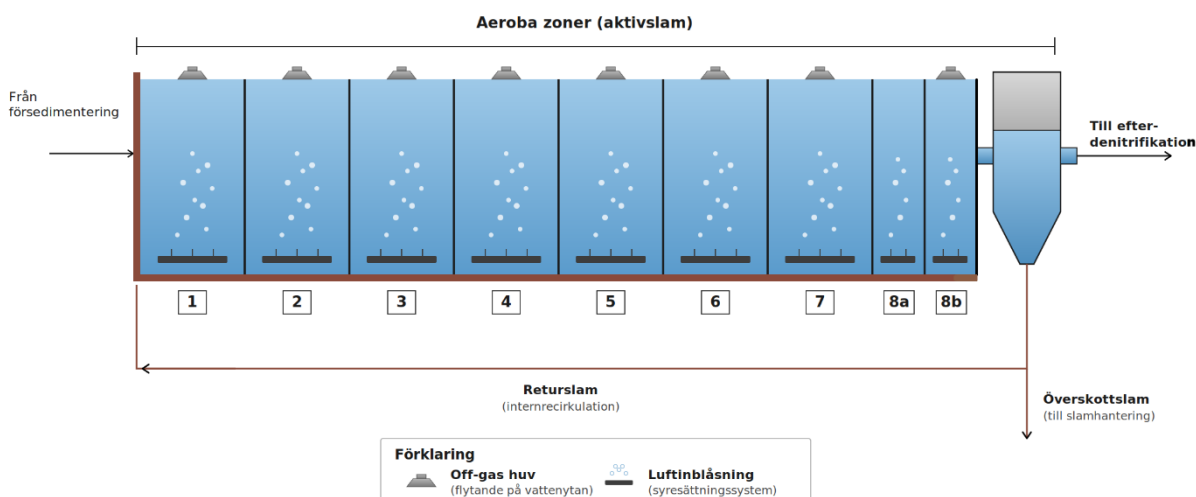
3 Metod

3.1 Aktivslamsystemet vid Klagshamns ARV

Aktivslamanläggningen bestod av två identiska parallella linjer där vardera fungerade som en pluggflödesreaktor uppdelad i åtta zoner. Dessa utgjorde den biologiska aktivslamhanteringen vid Klagshamns ARV och linje två blev föremål för den huvudsakliga undersökningen, se figur 1. Anledningen till att aktivslamanläggningen valdes att studeras var att majoriteten av lustgasproduktionen sker i form av bi- och mellanprodukter från den utökade biologiska kväveavskiljningen. Varje zon hade ett eget system för luftflödesstyrning i syfte att kontrollera oxidation av organiskt material och nitrifikationsgrad. Varje zon hade också varsin huv för att fånga upp frånluft till gasmätningar för syre (O_2), koldioxid (CO_2), lustgas och metan (CH_4) vilket visualiseras i figur 1. Mätningarna gjordes automatiskt av ett mätningssystem i ett skåp kopplat till huvarna från varje zon. Gasflödet från varje zon kopplades in till mätskåpet i tur och ordning, alltså kunde inga zoner mätas samtidigt. Vid undersökningar har det visat sig att lustgasbildningen är mycket dynamisk både på dygnssikt och längs med bassängerna. En hot-spot har kunnat påvisas i zonerna 2–4 där höga ammoniumhalter och intensiv nitrifikation råder.

Klagshamns ARV - AS linje 2

Schematisk skiss - 8 zoner (zon 8 uppdelad i 8a och 8b)
Alla zoner är aeroba



Figur 2. Schematisk skiss över Klagshamns ARV, Aktivslamlinje 2.

För att styra syretillförseln till de aeroba zonerna användes en trippelkaskadreglering (Åmand, 2013) bestående av tre seriekopplade PI-regulatorer. Överst i kaskaden satt en ammoniumregulator, som arbetade mot ett fastställt börvärde för ammoniumkväve i utgående vatten. Uti-

från avvikelserna mellan uppmätt och önskad ammoniumhalt beräknade regulatorn ett gemensamt syrebörvärde, vilket sedan användes av de fyra underliggande syreregulatorerna. Dessa var placerade i de zoner där syregivare fanns, och styrde syrehalten genom att i sin tur ange ett börvärde på luftflödet till zonens luftflödesregulator. Eftersom syremätning endast skedde i vissa zoner tillämpades det framräknade luftflödesbörvärdet även i närmast uppströms belägen zon utan mätning, det vill säga zon 1, 3, 5 och 7. Både syre- och luftflödesbörvärdena var begränsade till förutbestämda intervall, som kunde variera mellan zonerna. Typiska nivåer låg i storleksordningen 1,3–4,0 mg O₂/l för syrebörvärdet och 50–1000 Nm³/h för luftflödet.

Under den period som arbetet baseras på var den automatiska luftflödesstyrningen i zon 3 ur funktion, och zonen kördes istället manuellt med ett fast luftflöde. Detta innebär att zon 3 under hela mätperioden avvek från den kaskadreglering som beskrivs ovan, vilket är värt att ha i åtanke vid tolkningen av resultaten.

3.2 Motivering av vald ansats

Mot bakgrund av kapitel 2 valdes i detta arbete en datadriven ansats baserad på LSTM-nätverk. Valet motiverades av att reningsprocessen är dynamisk och starkt tidsberoende, där aktuella tillstånd påverkas av tidigare förhållanden i systemet, samtidigt som tillgången till långa, högupplösta driftdatabaser vid Klagshamns ARV gjorde en datadriven ansats praktiskt genomförbar. LSTM-strukturen är specifikt utvecklad för att hantera sådana tidsberoenden (Hochreiter & Schmidhuber, 1997). Tidigare LSTM-tillämpningar på lustgasprediktion har huvudsakligen behandlat det fullskaliga verket som en enhetlig svart låda och predikerat aggregerade utsläpp från processens utlopp (Hwangbo et al., 2021; Seshan et al., 2024). I föreliggande arbete byggs istället den fysiska processens seriella struktur in genom en zonvis kaskad, där varje zons utgångstillstånd utgör ingång till nästa zon, utan att specifika reaktionskinetiker behöver definieras explicit.

3.3 Korrelationsanalys

Tidigare korrelationsstudier av lustgasutsläpp vid linje 2 på Klagshamns ARV har huvudsakligen utgått från Pearsons korrelationskoefficient (VA SYD, 2021), som beskriver styrkan i linjära samband mellan variabler. Pearson är väl etablerad och lämpar sig när relationen mellan två storheter är ungefär linjär, men har en känd begränsning i biologiska system där sambanden ofta är icke-linjära, uppvisar tröskeleffekter eller mätnadsbeteende. Förändringar i exempelvis syrehalt, ammoniumhalt eller redoxförhållanden ger sällan helt proportionerliga förändringar i lustgasbildningen, och korrelationsmättet riskerar då att underskatta verkliga samband (Montgomery & Runger, 2014).

Av denna anledning kompletterades Pearson med Spearmans rangkorrelationskoefficient. Spearman bygger på rangordning av observationer och mäter monotona snarare än strikt linjära samband. Monotona samband skiljer sig från linjära samband på det sätt att de kan vara icke-linjära så länge de rör sig i en och samma riktning, exempelvis att y växer med x utan att

förändringen är proportionerlig. Spearmans koefficient beräknas genom att varje observationsvärde ersätts med sitt rangtal inom respektive variabel, varefter korrelationen mellan rangtalen beräknas. I sin enklaste form kan måttet uttryckas enligt ekvation 1:

$$\rho_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (1)$$

där n är antalet observationer och d_i är skillnaden mellan rangtalen för de två variablerna för observation i . Rangtransformeringen reducerar påverkan av extremvärden och gör måttet bättre lämpat för system med icke-linjär respons. De båda måtten användes parallellt för varje variabelpar, vilket gjorde det möjligt att identifiera både rent linjära samband och samband som är systematiska men icke-linjära (Montgomery & Runger, 2014).

Reningsprocessen vid Klagshamns ARV består av åtta zoner med tydliga inbördes beroenden, där förhållandena i en zon påverkas av vad som skett tidigare i processen. Korrelationsanalysen utfördes därför zonvis, så att lustgas i respektive zon analyserades tillsammans med lokala processvariabler och relevanta tidsförskjutna storheter. Resultaten visualiserades med spridningsdiagram, vilket möjliggjorde en kvalitativ bedömning av sambandens karaktär utöver de numeriska korrelationsvärdena. Analysen användes främst som ett explorativt verktyg för att identifiera relevanta variabler inför det fortsatta arbetet med prediktionsmodellering, och ska därför inte tolkas som ett bevis för kausala samband.

3.3.1 Tidsberoende

I ett pluggflödessystem är det inte rimligt att jämföra samtidiga mätvärden från olika punkter i bassängen direkt med varandra. Ett ammoniumvärde som registreras i inloppet vid en viss tidpunkt påverkar förhållandena i senare zoner först efter en viss hydraulisk uppehållstid, och en korrelationsanalys som inte tar hänsyn till denna transporttid riskerar att jämföra mätvärden som tillhör olika vattenpaket. För att hantera detta infördes tidsförskjutna variabler där varje extern processvariabel synkroniserades till respektive zon utifrån en uppskattad effektiv uppehållstid genom bassängen.

Den effektiva uppehållstiden kan inte beräknas direkt från bassängens geometriska volym, eftersom slamansamling i botten, ofullständig omblandning och kortslutningsvägar gör att den volym som faktiskt deltar i transporten avviker från den nominella. För att estimeras den effektiva volymen utnyttjades istället de tre ammoniummätningar som finns längs processen: i inloppet samt i zon 4 och zon 8. Eftersom ammonium är samma fysikaliska storhet i alla tre mätpunkter och transporteras genom bassängen längs pluggflödet, kan tidsförskjutningen mellan signalerna tolkas som ett bra mått på den effektiva transporttiden. Det är däremot inte rimligt att förvänta sig någon perfekt korrelation då ammonium reagerar under processen och

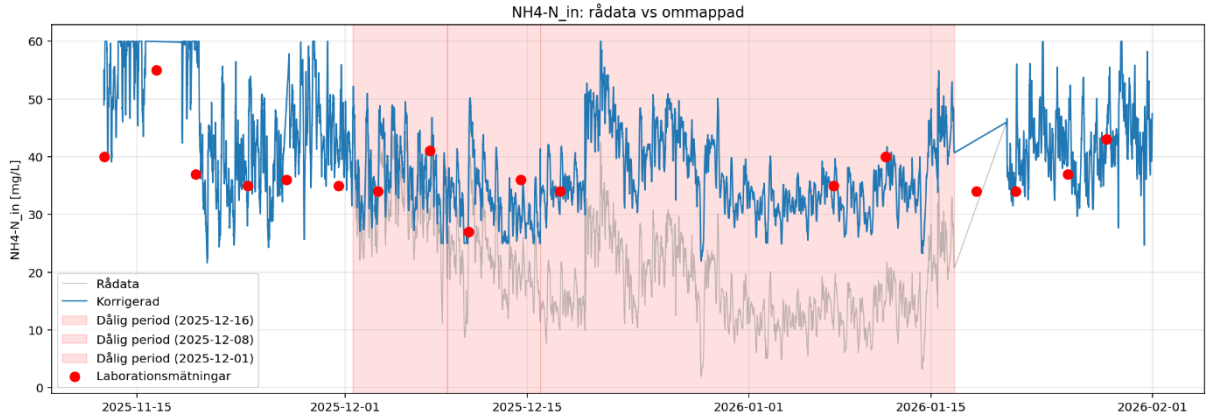
koncentrationen i senare zoner kommer att påverkas av flera parametrar. Men eftersom ammoniumkoncentration i inloppet självklart påverkar ammoniumkoncentration i senare zoner så förväntas en korrekt tidsförskjutning att öka korrelationen.

Korrelationen mellan NH_4 in i linjen och NH_4 i zon 4 respektive zon 8 beräknades vid varierande tidsförskjutningar. Den tidsförskjutning där korrelationen blev maximal användes som estimat av den effektiva uppehållstiden till respektive zon. Genom att två oberoende mätpunkter användes kunde estimaten dessutom kontrolleras mot varandra: om flödet genom bassängen är förhållandevis jämnt förväntas uppehållstiden till zon 4 vara ungefär halva uppehållstiden till zon 8. Den effektiva volym som följer av analysen användes därefter konsekvent vid omräkning av aktuellt flöde till uppehållstid genom hela analysen.

3.3.2 Dataupparbetning

Inför analysen filtrerades data för att exkludera driftfall med orimliga eller extrema data samt perioder med kända datakvalitetsproblem. I vissa tidsintervall korrigerades även signalen för inkommande ammonium, eftersom mätningen under dessa perioder bedömdes vara felkalibrerad. Viktigt att tillägga är att huv 1 och 2 visade sig sitta snett i bassängen och därmed kunde det inte uteslutas att gasläckage kan ha förekommit i dessa zoner. Detta noterades vid besök på anläggningen och kan ha haft påverkan på datakvaliteten. Huvarna justerades men den huvudsakliga datamängden var tvungen att hämtas från perioden med sneda huvar. Den datamängd som fanns från efter justeringen analyserades och jämfördes mot den tidigare, och resultatet behandlas i avsnitt 4.8.

I figur 3 illustreras hur data över den ingående ammoniumkoncentrationen passades om för att bättre representera den faktiska halten vid de tillfällen mätaren uppvisat stora differenser från laboriemätningar. Laboriemätningarna kommer från stick- och dygnsprover från Klagshamnsverket, som analyseras av VA SYDs laborier, som återkommande behöver göra mätningar för att se till att mätarna är välkalibrerade. Denna justering av data gjordes för att slippa stora dataförluster och ansågs kunna ge en bra estimering av de faktiska värdena under perioden då systematiska fel av denna typ oftast genererar fel av konstant magnitud.



Figur 3. Visar hur systematiska datafel justerades för att bättre passa in efter laboratorievärdena. Blå linje visar den korrigerade datan. Röda prickar visar laboratiemätningarna. Grå linje följer ammoniummätarens data för inflödet.

3.3.3 Frånluftsmätning och homogenisering av data

Gasmätningarna för syre, koldioxid, lustgas och metan omvandlades från volymkoncentrationer till massflöden med hjälp av den ideala gaslagen, så att analysen baserades på utsläppsrelevanta storheter snarare än enbart koncentrationer. Luftflödet i respektive zon, $Q_{luft,z}$, mättes kontinuerligt och anges i Nm^3/h , dvs. normerat till standardförhållanden ($T_s = 273,15 \text{ K}$, $P_s = 101\,325 \text{ Pa}$). Det totala molflödet av frånluft i zon z vid tiden t beräknades därför enligt ekvation 2:

$$\dot{n}_{tot,z}(t) = \frac{P_s}{R T_s} \cdot Q_{luft,z}(t) \quad (2)$$

där $\dot{n}_{tot,z}$ är totalt molflöde (mol/h), P_s är standardtryck (101 325 Pa), R är den allmänna gaskonstanten ($8,314 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$) och T_s är standardtemperatur (273,15 K). Det antas här att frånluftens molflöde är lika med inluftens, vilket är en god approximation eftersom syreförbrukningen och koldioxidproduktionen är av samma storleksordning och därmed i stort sett tar ut varandra volymetriskt.

Frånluftskoncentrationerna omvandlades till molfraktioner $x_{i,z}$ genom $x_{i,z} = C_{i,z} / 100$ för syre och koldioxid (givna i vol-%) respektive $x_{i,z} = C_{i,z} / 10^6$ för lustgas och metan (givna i ppm). Massflödet av gas i i zon z beräknades sedan enligt ekvation 3:

$$\dot{m}_{i,z}(t) = x_{i,z}(t) \cdot \dot{n}_{tot,z}(t) \cdot M_i \quad (3)$$

där $\dot{m}_{i,z}$ är massflödet av gas i och M_i är gasens molmassa. De molmassor som användes var $M_{syre} = 31,998 \text{ g/mol}$, $M_{koldioxid} = 44,010 \text{ g/mol}$, $M_{lustgas} = 44,013 \text{ g/mol}$ och $M_{metan} = 16,043$

g/mol. Koldioxid uttrycks i kg/h medan lustgas och metan uttrycks i g/h på grund av deras betydligt lägre koncentrationer.

För syret beräknades istället det upptagna syreflödet som differensen mellan ingående luft och uppmätt frånluftshalt, eftersom det är denna storhet som speglar processens biologiska aktivitet:

$$\dot{m}_{\text{syre,upptag},z}(t) = \frac{C_{\text{syre,luft}} - C_{\text{syre},z}(t)}{100} \cdot \dot{n}_{\text{tot},z}(t) \cdot M_{\text{syre}} \quad (4)$$

där $C_{\text{syre,luft}} = 20,95$ vol-% är syrehalten i omgivningsluft. Genom denna omvandling kunde gasflödena uttryckas i kg/h (upptaget syre, koldioxid) respektive g/h (lustgas, metan), vilket gav en mer representativ grund för att analysera hur olika processparametrar samvarierade med lustgasutsläppen i respektive zon.

Gasmätningarna görs via en metod som mäter en zon i taget i ca sex minuter. Det är alltså ca 48 minuter mellan nya gasmätningar i varje zon. I datan kopieras det senast tillgängliga nya värdet så att man får ca 48 minuter av samma värde. I verkligheten kan det vara helt andra värden under de 48 minuterna av kopierade värden, men i och med att de andra parametrarna har värden varje minut och för att vi skulle kunna koppla ihop zonerna med varandra ville vi ändå använda dessa kopierade, uppskattade värden. För att göra denna 48-minutersperiod mer användbar började vi med att göra om hälften av värdena. Den första hälften behölls, 1-24 minuter efter ett nytt mätvärde, medan den andra hälften, 25-48 minuter efter mätvärdet, gjordes om till värdet för det nästkommande nya mätvärdet. De värden som gjordes om blir alltså som mest 24 minuter ifrån ett nytt mätvärde istället för 48 minuter. Fortsättningsvis skapade vi vikter mellan 0-1 för hur ”gamla” värdena var, baserat på minuter ifrån ett nytt mätvärde. Viktningen utformades som en linjär funktion där vikten var 1 vid en ny mätning och avtar till 0 vid 24 minuter från närmaste mätvärde, enligt ekvation 5.

$$W = 1 - \frac{t}{24} \quad (5)$$

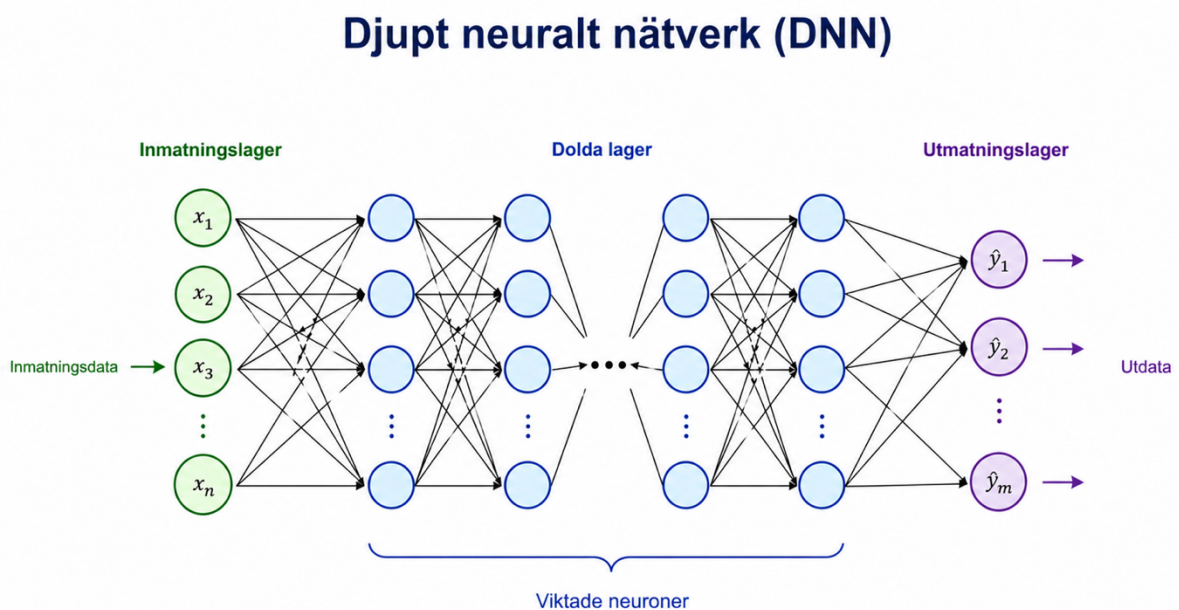
Vikten (W) beroende av tiden för hur gammalt ett värde är (t). Således har varje datapunkt en viktning för hur mycket vi kan lita på värdet, vilket gör varje minut av gasdata mycket mer användbart.

3.4 Maskininlärningsmodell

Maskininlärningsmodeller av olika typer har i litteraturen testats för prediktion av processrelaterade utsläpp med varierande resultat (Fang et al., 2026). I takt med den ökande utvecklingen inom artificiell intelligens har mer avancerade modeller blivit tillgängliga, vilket möjliggör analys av komplexa och högdimensionella datamängder. I detta arbete utnyttjas denna möjlighet genom att utveckla en datadriven prediktionsmodell för lustgasutsläpp baserad på ett stort antal processparametrar.

3.4.1 Modellarkitektur

Grundtanken med modellen är att beskriva sambandet mellan indata, såsom luftflöden, kvävefraktioner och övriga processtillstånd, och utdata i form av lustgasutsläpp. Modellen kan konceptuellt betraktas som en komplex matematisk funktion med flera in- och utvariabler, där relationen mellan dessa inte explicit specificeras utan i stället lärs in från data. Detta sker genom ett neuralt nätverk där informationen passerar genom flera lager av artificiella neuroner. Varje neuron tilldelas vikter som successivt justeras under träningen, så att modellens prediktioner över tid närmar sig de uppmätta värdena (Kinsley & Kukiela, 2020). Förloppet illustreras i figur 4.



Figur 4. Uppbyggnad av ett DNN (Djupt Neuralt Nätverk). Figuren visar beståndsdelarna inmatnings-, dolt- och utmatnings- lager samt hur dessa är sammanlänkade.

I detta arbete används en sekventiell modell baserad på Long Short-Term Memory-nätverk (LSTM), vilket är en typ av rekurrenta neurala nätverk särskilt lämpade för tidsseriesdata. Valet av modell motiveras av att reningsprocessen är dynamisk och starkt tidsberoende, där aktuella tillstånd påverkas av tidigare förhållanden i systemet. LSTM-modellen möjliggör därmed att dessa tidsberoenden kan fångas på ett mer effektivt sätt än med statistiska modeller. (Hochreiter & Schmidhuber, 1997)

Modellen byggdes som en sekventiell zonmodell där vardera av de åtta zonerna representeras av ett eget neuralt nätverk. Varje zonmodell predikterar flera processtillstånd samtidigt, däribland lustgas, syre och koldioxid i frånluften, och i de zoner där ytterligare mätningar finns tillgängliga även löst syre och ammonium i vattenfasen. Denna multioutput-struktur valdes för att bättre fånga kopplingen mellan olika delar av processen, snarare än att enbart prediktera lustgas isolerat.

3.4.2 Kaskadstruktur

Zonmodellerna kopplas samman till en kaskad där predikterade tillstånd från en zon används som indata till efterföljande zon. Detta efterliknar processens fysiska struktur, där förhållanden uppströms påverkar vad som sker nedströms. Utöver dessa kaskadvariabler används som indata till varje zon: zonens luftflöde, tidsförskjutna externa processvariabler (NH₄ in, pH, redoxpotential, suspenderad substans, temperatur) samt cykliska tidsvariabler för timme och veckodag. Tidsvariablerna möjliggör att modellen kan beskriva återkommande dygns- och veckomönster i processen. Inputvalen baserades dels på korrelationsanalysen, dels på de processparametrar som enligt avsnitt 2.1 är kända drivkrafter för lustgasbildning. Själva kaskadmodellen är något som skiljer sig från tidigare LSTM-modeller för lustgasutsläpp inom avloppsrening, såsom Hwangbo et al. (2021) och Seshan et al. (2024).

3.4.3 Träning och utvärdering

Den data vi fick tillgång till består av cirka 120 000 datapunkter som hämtats in varje minut under perioden 2025-11-10 till 2026-01-31. Tillgängliga mätvariabler per zon framgår av tabell 1.

Tabell 1. Respektive zon och dess tillgängliga mätdata. Kolumn 1 korresponderar till luftmätare vid inlopp till respektive bassäng underifrån. Kolumn 2–4 motsvarar frånluftsmätare och kolumner 7–14 korresponderar till vätskefasmätare, utplacerade i ett färre antal zoner. ORP står för oxidation-reduktionspotential, dvs redoxpotentialen.

Zon	Nm ³ /h	%	%	ppm	ppm	mg/L	mg/L	mg/L	mg/L	mg/L	-	mV	°C	mg/L
-	Luft	O ₂	CO ₂	N ₂ O	CH ₄	N ₂ O	NH ₄ -N	NO ₃ -N	NO ₂ -N	O ₂	pH	ORP	T	SS
1	X	X	X	X	X	X	X							
2	X	X	X	X	X					X				
3	X	X	X	X	X									
4	X	X	X	X	X		X			X				
5	X	X	X	X	X									
6	X	X	X	X	X					X				
7	X	X	X	X	X									
8	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Databehandlingen följde principen som beskrivs i avsnitt 3.3.2. Tidsserierna delades upp i dygnssekvenser där varje dygn representerar en komplett indata- och utdataföljd med 1440 tidssteg. Endast dygn med tillräcklig datakvalitet inkluderades. Datasetet delades randomiserat in i dygn för tränings- och testdata, där 80 % av dygnen användes för träning och resterande 20 % för utvärdering. För varje zon tränades en separat LSTM-modell i Python, där PyTorch användes med Adam-optimizer och viktad medelkvadratfel (MSE) som förlustfunktion. För att minska risken för överanpassning användes early stopping baserat på valideringsdata.

Förlustfunktionen viktades på två sätt. Dels användes de viktningsfaktorer som beskrivs i avsnitt 3.3.3 för att låta modellen lita mer på färskare gasmätningar än på interpolerade värden. Dels viktades ammoniumprediktionen med en faktor 3 jämfört med övriga outputs, eftersom NH_4 i zon 8 representerar reningsresultatet och därför är särskilt viktig att fånga korrekt. Övriga hyperparametrar och detaljer om modellarkitekturen sammanställs i appendix.

Modellernas prestanda utvärderades med determinationskoefficient (R^2), medelabsolutfel (MAE) och root mean square error (RMSE), via två kompletterande metoder. I teacher forcing matas varje zonmodell med uppmätta värden från föregående zon, vilket isolerar prestandan för en enskild zon. I kaskadinferens används istället den föregående zonens predikterade värden som indata, vilket innebär att fel kan fortplanta sig genom kedjan på samma sätt som vid praktisk användning där modellen körs i serie utan tillgång till mätvärden mellan zonerna. Den senare metoden är därför mer representativ för en realistisk driftsituation.

3.5 Optimering

Efter att prediktionsmodellen tränats användes den som grund för att analysera möjliga processförändringar. Optimeringsproblemet formulerades så att luftflödet i varje zon tilläts variera inom databaserade gränser, definierade utifrån observerade percentiler för att undvika orealistisk extrapolering.

Målet med optimeringen var att minimera den totala dygnssumman av lustgas över samtliga zoner. Samtidigt infördes ett bivillkor för ammonium i zon 8, så att den optimerade lösningen inte fick försämrade utgående ammoniumhalt jämfört med baslinjen, där baslinjen är modellens prediktion och inte den faktiska datan. Detta implementerades som en straffterm i målfunktionen, där överskridande av ammoniumgränsen gav en kraftigt ökad bestraffning. Även orimligt stora ändringar på kort tid straffades för att få bort att optimeraren föreslår hopp mellan högsta och lägsta tillåtna värde fram och tillbaka. Optimeringen genomfördes med gradientbaserad uppdatering där de tränade neurala nätverken användes som differentiabla delmodeller i en sammanhängande kaskad.

3.5.1 Implementering och begränsningar

Optimeringen implementerades i Python med PyTorch, där de tränade LSTM-modellerna för zon 5–8 utgjorde differentiable komponenter i en sammanhängande kaskad. Här valdes zon 1–4 bort på grund av sämre prediktion från LSTM-modellen, vilket förklaras mer i nästa kapitel, framförallt avsnitt 4.3. Luftflödet i zon 5–8 utgjorde beslutsvariabler med ett värde per minut och zon, medan uppmätta värden från zon 4 användes som indata till kaskadens första zon. Som startpunkt användes det historiskt uppmätta luftflödet, vilket innebär att optimeringen sökte efter förbättringar i ett område nära den faktiska driften. Beslutsvariablerna parameterades så att luftflödet alltid hölls inom datadrivna gränser, definierade per zon som 3:e respektive 99:e percentilen av observerat luftflöde i hela datasetet. Dessa gränser valdes för att säkerställa att optimeringen håller sig inom ett område där tillräckligt med träningsdata finns, samtidigt som extrema värden som inte är representativa för normal drift utesluts. Optimeringen genomfördes med Adam-optimizer där gradienterna beräknades automatiskt genom hela kaskaden av zonmodeller. En mer detaljerad beskrivning av optimeringen återfinns i appendix B.

3.5.2 Känslighetsanalyser

Utöver optimeringen genomförde vi även känslighetsanalyser, där luftflödet i en zon i taget varierades över ett intervall medan övriga zoner hölls konstanta. Detta möjliggjorde analys av hur luftflödet i respektive zon påverkar de totala lustgasutsläppen och om sambanden är monotona eller uppvisar ett lokalt optimum. Slutligen gjorde vi enklare simuleringsfall där luftflödet i samtliga zoner förändrades med ett fast värde, för att ge en mer direkt tolkning av hur generella förändringar i luftning påverkar lustgasutsläpp och ammoniumhalter.

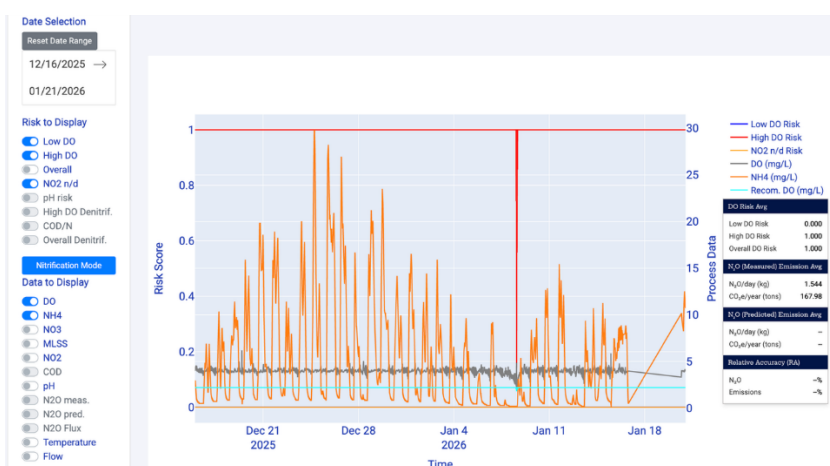
Sammanfattningsvis möjliggör den valda metodiken att komplexa, tidsberoende och icke-linjära samband i reningsprocessen kan analyseras och utnyttjas för att identifiera potentiella processförbättringar. Samtidigt bör resultaten tolkas med viss försiktighet, eftersom modellen är datadriven och därmed beroende av kvaliteten och representativiteten hos den använda datamängden.

3.6 Samarbete med Cobalt Water Global

Under arbetets gång fick vi även kontakt med Jose Porro, grundare av Cobalt Water Global. Han har sedan 2020 drivit upp sin startup med motivet att inkorporera AI och maskininlärning för att minska lustgasutsläppen från ARV. Tillsammans med vår handledare David J.I. Gustavsson och Porro kom vi fram till att vi även skulle få testköra vårt dataset genom Cobalt Waters modell för lustgasutsläpp. Modellen kallas ”N₂O Risk DSS” och bygger likt vår egen modell på datadriven uppträning av ett neuralt nätverk med optimerade styrparametrar som huvudsaklig output. Porros modell testar 15 olika maskininlärningsmodeller och rankar dem efter vilken som ger bäst resultat. Den använder forskning från olika reningsverk som grund till att räkna ut riskvärden för syrekonzentrationen från datan. Riskvärdena används sen, tillsammans med datan, för att räkna ut ett rekommenderat börvärde för syrekonzentrationen.

Metodiken är beskriven i Porro et al. (2017), där en fyrstegsprocess används för att identifiera ”Low DO risk” och ”High DO risk”.

CSV-filer om 8, en per zon, skickades in och användes för att träna upp deras modell. Via inlogg till hemsidan kunde vi sedan komma åt en GUI (Guided User Interface) där val av parametrar, tidsperiod för plot, zon och typ av riskkurva kunde illustreras. Från DSS-modellen kunde även en kurva för ny rekommenderad syrekonzentration ritas upp. Denna modell bygger med andra ord på samma tes som vår egenutvecklade och är tänkt att agera som en intressant bedömning av den egenutvecklade modellens trovärdighet. Tanken var att resultat från de respektive modellerna skulle jämföras för identisk inputdata och därmed förhoppningsvis kunna styrka slutsatserna ytterligare. Gränssnittet finns illustrerat i figur 5.



Figur 5. Visar Cobalt Waters online-gränssnitt. Till vänster i menyn väljs vilket datum-intervall som ska plottas. Här väljs även vilken risk- och data- typ som ska visas. Cyanfärgad linje illustrerar ML-modellens rekommenderade syrekonzentration i mg/l.

3.7 AI-användning

Hur man använder generativ AI är en högst relevant och viktig fråga i dagens värld, både inom utbildning och arbete. Vi har under detta examensarbete haft attityden att vi ska använda generativ AI och dess styrkor så bra som möjligt för att kunna vara mer effektiva och få mer gjort, men också för att få bättre kvalitet på det vi gör. Vi har använt flera olika generativa AI-verktyg under arbetets gång då vi anser att det viktigt att jämföra dem med varandra för att eliminera misstag och öka den kvalitativa potentialen. De AI-verktyg som främst har använts är Claude, ChatGPT, Microsoft Copilot och Google Gemini.

AI-verktyg har framförallt använts under arbetet i samband med all programmering vi har gjort. Utan AI hade det vi har kunnat testa under denna period förmodligen tagit oss flera år. Hjälpen vi har fått att skriva kod när vi programmerar har gjort att vi sparat extremt mycket tid. Vi har kunnat skapa och testa många olika modeller för att till slut landa i det vi skriver

om i denna uppsats. Vi har även använt AI för att få hjälp att formulera våra tankar och det vi vill få sagt i text, som att skriva om delar av uppsatsen för att få ett bättre språk.

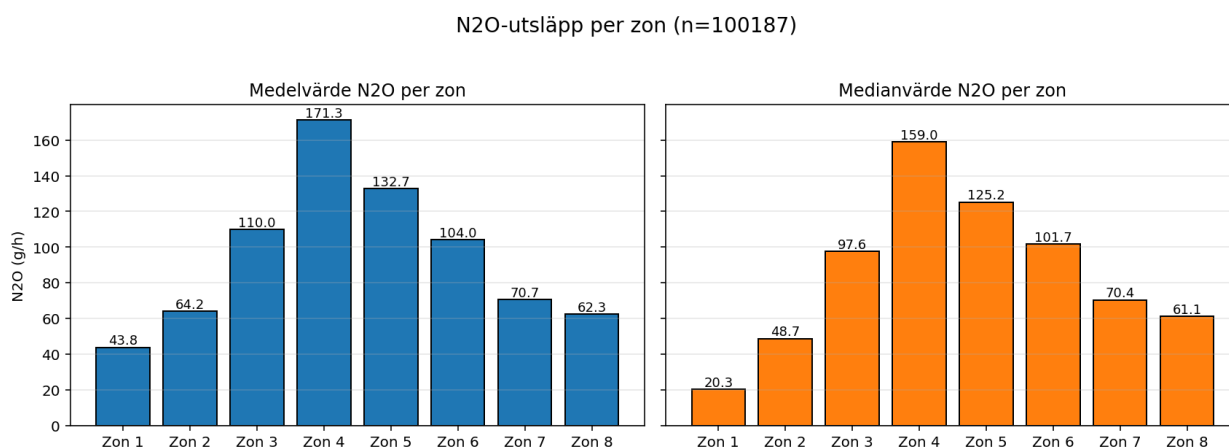
Vi vill självklart också nämna att vi aldrig litar blint på AI-verktygen. Både när det kommer till programmering och omskrivande av text så har vi läst igenom, granskat och justerat allt vi presenterar som eget. Vi ser generativ AI som ett utmärkt verktyg att få fram det vi hade kunnat göra själva, fast oerhört mer effektivt.

4 Resultat och diskussion

I detta kapitel presenteras resultaten från analys av de totala lustgasutsläppen, korrelationsanalysen, maskininlärningsmodellen och den efterföljande optimeringen av luftningsstrategin. Resultaten redovisas stegvis för att först utvärdera modellens prediktionsförmåga och slutligen undersöka möjligheter till minskade utsläpp genom optimerad drift.

4.1 Inledande översikt av lustgasutsläpp

Resultaten är baserade på den data som beskrivs i avsnitt 3.2 och av ca 120 000 tillgängliga minuter av data rensades ca 20 000 av de bort. Figur 6 visar data för hur fördelningen av lustgasutsläpp ser ut över de olika zonerna under den period resultaten är baserade på. Zon 1 har tydligt lägst mängd utsläpp medan zon 4 har klart mest utsläpp. Den klart lägre nivån lustgasutsläpp i zon 1 och 2 är en indikation på att nitrifikationen inte tar fart i de första zonerna och att det istället sker mer nedbrytning av organiskt material. Vid sammanställning av lustgasutsläppen över hela mätperioden motsvarade dessa en utsläppsfaktor på 2,7 % av N-belastningen, vilket ligger inom det spann på 0,5–5 % som nämns i avsnitt 2.1.4.



Figur 6. Medel- och medianvärdet för utsläpp av N₂O (g/h) i varje zon, där n är antalet tillgängliga datapunkter för varje zon.

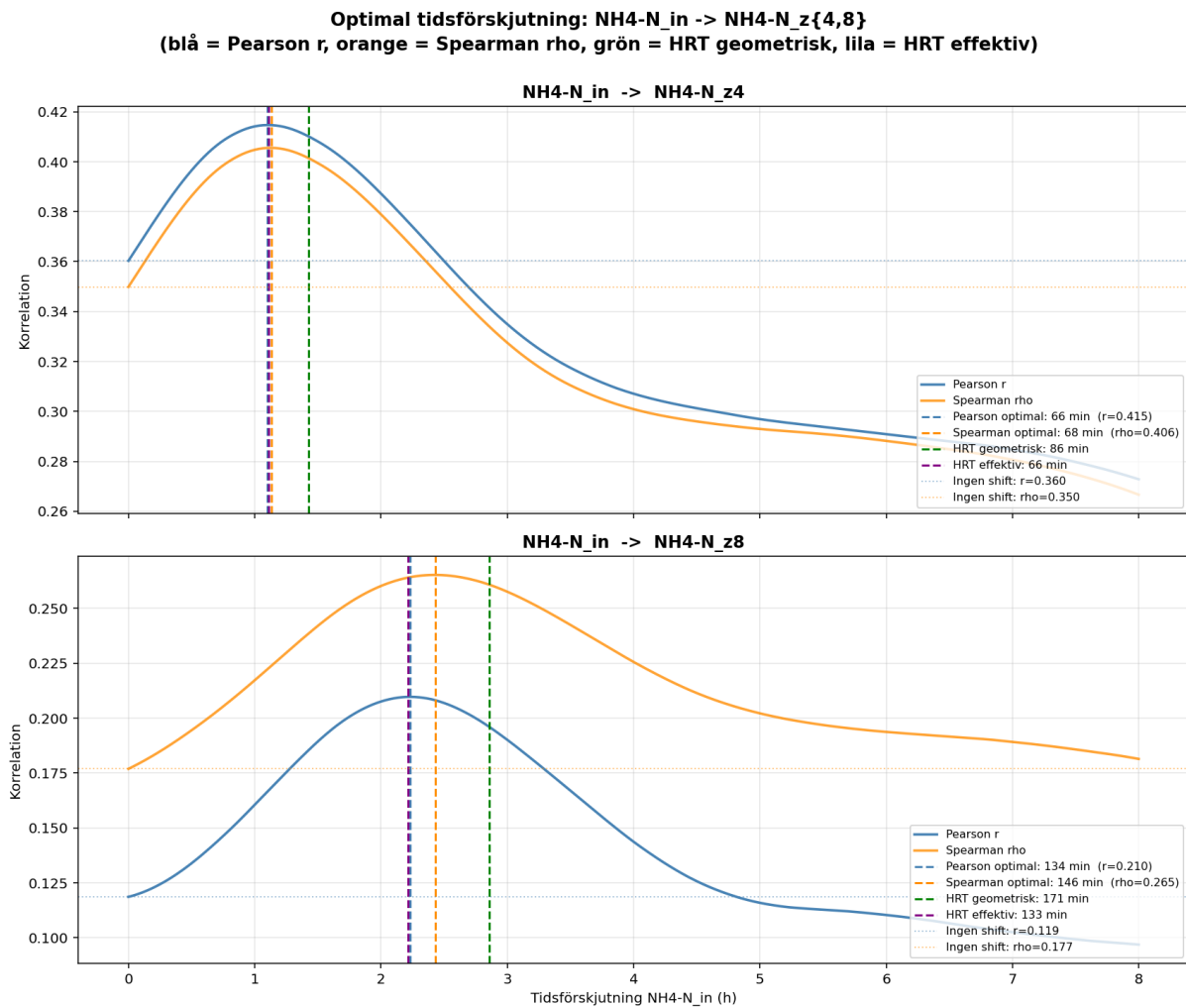
4.2 Korrelationsanalys

Korrelationsanalyserna är baserade på Pearson- och Spearman-korrelation som beskrivs i avsnitt 3.3 och användes först för att hitta en bra estimering av vattnets uppehållstid i varje zon. Upphållstiden säger hur mycket varje zons värden måste förskjutas i tid för att analyseras mot värden från en annan zon. Vid fortsatt korrelationsanalys av parametrar från olika zoner kunde den estimerade uppehållstiden användas för att hitta mer korrekta samband.

4.2.1 Tidsförskjutning

Den korrelationsbaserade estimeringen av effektiv uppehållstid gav ett tydligt och konsistent resultat. Figur 7 visar Pearson- och Spearman-korrelationen mellan NH₄ in och NH₄ i zon 4

respektive zon 8 som funktion av tidsförskjutning. För zon 4 maximerades Pearson-korrelationen vid en förskjutning om 66 minuter ($r = 0,42$), och för zon 8 vid 134 minuter ($r = 0,21$). Att den optimala tidsförskjutningen till zon 8 är ungefär dubbelt så lång som till zon 4 ger en oberoende validering den estimerade tidsförskjutningen. Med dessa uppehållstider beräknas en effektiv volym på cirka 1 710 m³, att jämföra med bassängens geometriska volym på 2 200 m³. Den effektiva volymen användes därefter konsekvent vid omräkning av flöde till uppehållstid genom hela analysen.



Figur 7. Uppskattning av effektiv hydraulisk uppehållstid baserat på korrelation mellan NH₄-N i inloppet och NH₄-N i zon 4 (övre panelen) respektive zon 8 (nedre panelen) som funktion av tidsförskjutning. Den blå kurvan visar Pearsons korrelationskoefficient och den orangea kurvan visar Spearmans rangkorrelationskoefficient. Den gröna streckade linjen markerar den geometriska hydrauliska uppehållstiden baserad på bassängvolym (86 min till zon 4 och 171 min till zon 8), medan den lila streckade linjen markerar den effektiva uppehållstiden vid Pearson-korrelationens maxvärde (66 min respektive 133 min).

Korrelationsanalysen genomfördes både med och utan tidsförskjutning, vilket gjorde det möjligt att utvärdera om hänsyn till hydraulisk transport faktiskt påverkar de statistiska sambanden. När rådata analyserades utan tidsförskjutning framstod kopplingarna mellan inkommande processvariabler och lustgasutsläpp i de senare zonerna som svaga och delvis motsägelsefulla. Efter införande av tidsförskjutning förändrades dessa korrelationer, generellt mot starkare och mer tolkbara samband mellan inkommande variabler och tillstånd nedströms i processen. Som referens kan exempelvis korrelationen mellan NH_4 in och NH_4 i zon 4 nämnas: utan tidsförskjutning var Pearson-korrelationen 0,36, medan den efter införande av tidsförskjutning ökade till 0,42 (Spearman: 0,35 respektive 0,41).

Vissa korrelationer blev tvärtom svagare efter tidsförskjutning. Detta tolkas som att en del av de ursprungliga sambanden var skenbara, drivna av att flera variabler samvarierar i tid utan att ha en direkt processmässig koppling, och att de stabiliserades till mer rättvisande nivåer när jämförelsen istället gjordes inom samma vattenpaket. Sammantaget stödjer detta att tidsförskjutning är ett nödvändigt steg för att korrelationsanalysen ska bli processmässigt meningsfull i ett system av denna typ. Detta ligger i linje med den teoretiska beskrivningen av aktivslamlinjen som ett pluggflödesliknande system, där processförhållanden nedströms påverkas av belastning och omvandlingar uppströms först efter en hydraulisk fördröjning. Att sambanden stärks efter tidsförskjutning stödjer därför antagandet att simultana mätvärden från olika delar av bassängen inte nödvändigtvis representerar samma vattenvolym. Liknande behov av att beakta processens dynamik och tidsberoende har lyfts fram i tidigare studier av lustgasutsläpp från fullskaliga avloppsreningsverk, där utsläppen beskrivs som starkt varierande över både tid och rum (Vasilaki et al., 2019; Seshan et al., 2024).

4.2.2 Pearson- och Spearmankorrelation

Den parallella användningen av Pearson- och Spearmankorrelation visade sig värdefull. För flera variabelpar gav de båda måtten likvärdiga värden, vilket talar för att sambandet är någorlunda linjärt. För andra par var Spearman-koefficienten tydligt högre än Pearson, vilket indikerar att sambanden är systematiska men icke-linjära, exempelvis med tröskeleffekter eller mättnad. Sådana mönster syns också i de zonvisa scatter- och täthetsdiagrammen i appendix C, där flera variabelpar uppvisar krökta eller s-formade strukturer snarare än räta linjer. Förekomsten av icke-linjära samband är konsistent med den u-formade relation mellan löst syre och lustgasbildning som beskrivs i avsnitt 2.1.3 (Lotito et al., 2012; Duan et al., 2020), liksom med de tröskel- och mättnadseffekter som rapporterats i biologiska kväveavskiljningsreningsystem (Kampschreur et al., 2009). Observationen att en betydande del av sambanden är icke-linjära utgjorde en del av motiveringen till att gå vidare med en icke-linjär modelleringsansats (LSTM) i nästa steg, eftersom en rent linjär modell skulle missa en del av den information som finns i datan.

Korrelationsanalysen användes också för att identifiera vilka processvariabler som uppvisade starkast samband med lustgasbildning i respektive zon, och därmed utgöra ett underlag för val av indata till prediktionsmodellen. I zon 1 framträder överlag svaga samband mot lustgas, där ingen enskild processvariabel når Pearson- eller Spearman-värden över ca 0,3, vilket talar för

att nitrifikationen ännu inte tagit fart i denna zon. Från och med zon 2 dominerar istället kopplingen till föregående zons frånluftsmätningar av lustgas, koldioxid och syre, där exempelvis lustgas i föregående zon uppvisar Pearson-korrelationer på 0,75 eller högre i zon 3–8 och så högt som 0,97 i zon 8. Detta utgör en stark motivering för en zonvis och kaskadkopplad modellstruktur i nästa steg. Ett tydligt mönster är att zoner som delar luftflödesstyrning med nedströms liggande zon (zon 1 styrs av zon 2, zon 5 av zon 6 och zon 7 av zon 8) uppvisar särskilt höga korrelationer mot frånluften i föregående zon, med Pearson-värden upp mot 0,84–0,97. Zon 3 och 4 styrs inte tillsammans på samma sätt eftersom luftflödesstyrningen i zon 3 inte fungerar som avsett, och just dessa zoner uppvisar tydligt lägre korrelation mot föregående zons frånluft än de övriga paren. Detta förstärker bilden av att zoner som delar styrning också uppvisar starkast samvariation i frånluften. Metanmätningen utmärker sig genom att genomgående uppvisa betydligt svagare och delvis negativa samband med lustgas jämfört med övriga frånluftsvariabler, och valdes därför bort som indata till modellen. Lokalt luftflöde uppvisar successivt starkare samband med lustgas ju längre ned i linjen man kommer, från svaga värden ($r \approx 0,13$ – $0,28$) i zon 1–3 till tydliga samband ($r \approx 0,60$ – $0,81$) i zon 5–8. Tidsförskjutet inkommande ammonium bidrar främst med information i zon 5–7 ($r \approx 0,42$ – $0,43$). Mer detaljerade scatterdiagram per zon redovisas i appendix C.

Det är värt att betona att korrelationsanalysen i sig endast beskriver statistiska samband och inte fångar tidsdynamik eller interaktioner mellan flera variabler samtidigt. Den utgör därför ett första, explorativt steg, och resultaten används främst som vägledning för variabelval och för att motivera den modelleringsansats som beskrivs i avsnitt 4.3.

4.3 Prediktionsmodell

Modelleringen inleddes med en enklare ansats där samtliga tillgängliga processvariabler användes för att prediktera lustgasutsläppen i en gemensam modell. Det blev dock snabbt tydligt att en sådan ansats leder till brus och begränsad tolkbarhet, och att modelleringen behövde förenklas och struktureras tydligare för att ge meningsfulla resultat. Korrelationsanalysen, tillsammans med teorin inom området, användes därför för att välja ut de variabler som uppvisade starkast samband med lustgasbildningen, vilket minskade risken att irrelevanta signaler dominerade modellen. Variablerna som var med i modellen förklaras i 3.2 och kan ses i appendix A, där de bortvalda variablerna berodde på dålig korrelation med lustgas och otillräckligt stöd från teorin. Exempelvis valdes metan bort på grund av dålig korrelation, vilket kan ses i appendix C.

I nästa steg implementerades en LSTM-struktur för att bättre fånga de tidsberoenden som präglar processen. Modellen byggdes som en kedja av åtta zonvisa delmodeller, där varje zon tar föregående zons predikterade utvärden som del av sin indata. På så sätt efterliknar modellens struktur den fysiska processens seriella karaktär.

4.3.1 Modellprestanda

Modellens prestanda utvärderades med två metoder, dels genom *teacher forcing*, där varje zonmodell matades med uppmätta värden från föregående zon, dels genom *kaskadinferens*,

som innebär att varje zon använder predikterade värden från föregående zon istället för uppmätta värden. Resultaten tyder på ett tydligt mönster där zonerna 4–8 predikteras betydligt bättre än zonerna 1–3. Att huvarna i zon 1–2 var snedställda under perioden (se avsnitt 4.8) är sannolikt en bidragande orsak till den sämre prediktionsprestandan i dessa zoner, och stödjer beslutet att fokusera den fortsatta analysen på zon 5–8. Ytterligare en faktor är att luftflödesstyrningen i zon 3 var ur funktion under mätperioden (se avsnitt 2.1), vilket innebär att zonen kördes manuellt och därmed inte följde samma styrlogik som övriga zoner. Detta bidrar sannolikt också till att prediktionen är sämre i zon 3 och påverkar via kaskadkopplingen även zon 4. Skillnaden mellan de tidiga och senare zonerna kan även beskrivas utifrån processdynamik. En möjlig förklaring är att nedbrytning av organiskt material dominerar i de första zonerna, och att nitrifikationen ännu inte har etablerats fullt ut i samma utsträckning som längre ned i linjen. Detta stöds av figur 6, där lustgasutsläppen i zon 1 och 2 ligger klart lägre än i zon 3–5, vilket kan tyda på att nitrifikationen ännu inte dominerar processen i de inledande zonerna. I zon 5–8 är processen mer stabil och utsläppsmönstret styrs i högre grad av återkommande dygnsdynamik och luftningsnivå, vilket sannolikt är en av anledningarna till att modellen presterar bättre där.

4.3.2 Resultaterande lustgasprediktion

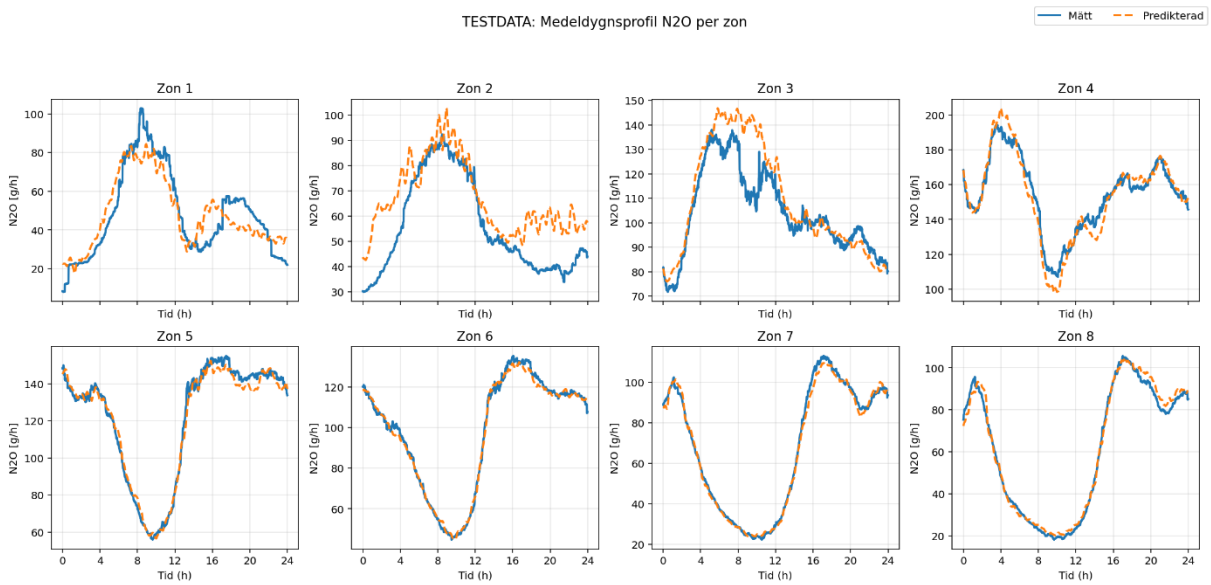
Tabell 2 och 3 visar R^2 -, MAE- och RMSE-värdena för de olika modellerna där man kan se den stora skillnaden mellan de tidiga och de senare zonerna. Medeldygnsprofilen för hur väl prediktionen följer mätvärdena finns illustrerat för de olika modellerna i figur 8 och 9. De R^2 -värden som modellen uppnår i zon 5–8 ligger i nivå med vad tidigare datadrivna ansatser rapporterat för lustgasprediktion vid fullskaliga verk (Hwangbo et al., 2021; Seshan et al., 2024; Mehrani et al., 2022), som genomgående redovisar $R^2 > 0,90$ enligt översikten av Fang et al., 2026. Skillnaden här är att prestandan uppnås zonvis i kaskad snarare än som en aggregerad utsläppsprognos, vilket adresserar den svartlåde-begränsning som lyftes i avsnitt 2.2.

Tabell 2. Modell för zon 1–8. Värden för prestanda vid prediktion av N_2O i alla zoner, både för teacher forcing och kaskadinferens, determinationskoefficient (R^2), medelabsolutfel (MAE) och root mean square error (RMSE).

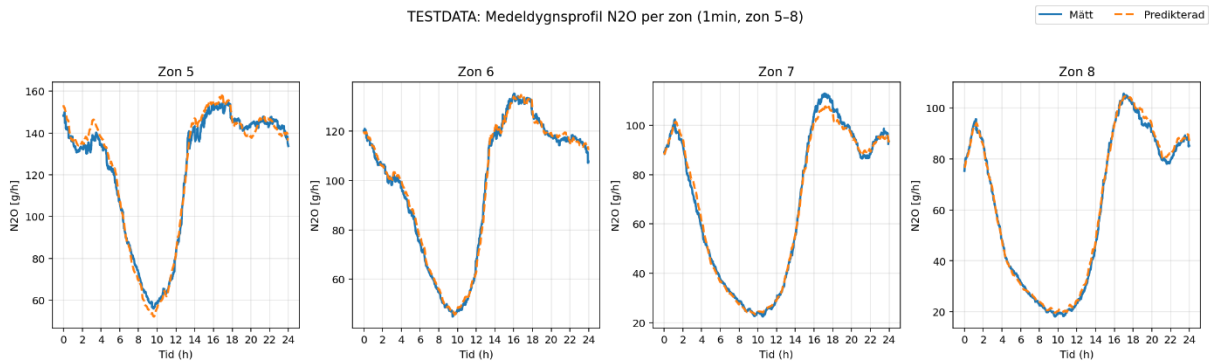
Zon	Teacher forcing			Kaskad-inferens		
	R^2	MAE	RMSE	R^2	MAE	RMSE
1	0,60	23,9	41,5	0,60	23,0	39,7
2	0,24	19,8	31,0	0,24	22,2	33,0
3	0,71	17,2	26,2	0,42	29,1	41,3
4	0,90	15,1	21,1	0,73	26,7	36,7
5	0,94	10,7	15,0	0,84	18,9	26,2
6	0,95	7,6	10,6	0,86	12,8	18,5
7	0,95	6,5	10,0	0,91	9,3	13,7
8	0,90	7,6	12,6	0,90	8,7	13,4

Tabell 3. Modell för zon 5–8. Värden för prestanda vid prediktion av N_2O i alla zoner, både för teacher forcing och kaskad-inferens, determinationskoefficient (R^2), medelabsolutfel (MAE) och root mean square error (RMSE).

Zon	Teacher forcing			Kaskad-inferens		
	R^2	MAE	RMSE	R^2	MAE	RMSE
5	0,95	10,2	13,4	0,96	10,1	13,6
6	0,96	6,4	8,9	0,94	8,6	11,7
7	0,96	5,4	8,8	0,94	7,6	11,5
8	0,97	4,7	6,6	0,94	6,9	10,0



Figur 8. Medeldygnsprofil av lustgasutsläpp (N_2O , g/h) per zon för testdata, baserad på prediktion med teacher forcing. Den blå linjen visar uppmätta värden och den orangea streckade linjen visar modellens prediktion. Profilen visar genomsnittlig dygnsvariation för zon 1–8.



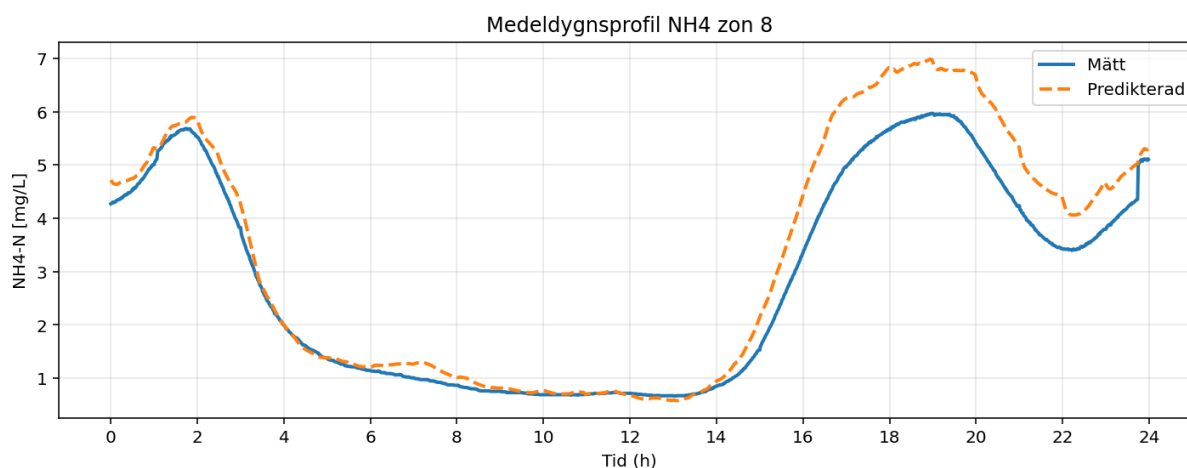
Figur 9. Medeldygnspröfil av lustgasutsläpp (N₂O, g/h) för zon 5–8 baserad på testdata och prediktion från modellen tränad enbart på de fyra sista zonerna (1-minutsupplösning). Den blå linjen visar uppmätta värden och den orangea streckade linjen visar modellens prediktion.

4.3.3 Prediktion av utgående ammoniumhalt

Prediktionen av ammonium i zon 8 är den variabel som visar sig vara svårast att återge korrekt, samtidigt som den är särskilt viktig eftersom den representerar utgående reningsresultat. Kaskadansatsen ger här bättre resultat än en isolerad prediktion, och även om R²-värdet är lägre än för lustgas följer modellen de övergripande dygnsmönstren väl. R²-värdet blev 0,87 för båda modellerna vid kaskad-inferens och är presenterat med värden för MAE och RMSE i tabell 4. Figur 10 visar medeldygnspröfilen av mätt ammonium i zon 8 mot prediktionen av ammonium från modellen med de fyra sista zonerna. Även om R²-värdet är lägre än önskat kan man se på medeldygnspröfilen att prediktionen är dynamisk och generellt sett följer trenderna bra. Detta innebär att modellen i första hand bör ses som ett verktyg för att beskriva trender och relativa förändringar snarare än exakta absolutnivåer, vilket är fullt tillräckligt för den optimering som studien syftar till.

Tabell 4. Värden för modellernas prestanda vid prediktion av NH₄ i zon 8, determinationskoefficient (R²), medelabsolutfel (MAE) och root mean square error (RMSE).

Modell	R ²	MAE	RMSE
Zon 1–8	0,87	0,79	1,18
Zon 5–8	0,87	0,81	1,17



Figur 10. Medeldygnspröfil av ammoniumkoncentration ($\text{NH}_4\text{-N}$, mg/L) i zon 8. Den blå linjen visar uppmätta värden och den orangea streckade linjen visar prediktion från modellen som är tränad på zon 5–8. Modellen följer den övergripande dygnsdynamiken väl, även om absolutnivåerna avviker tidvis.

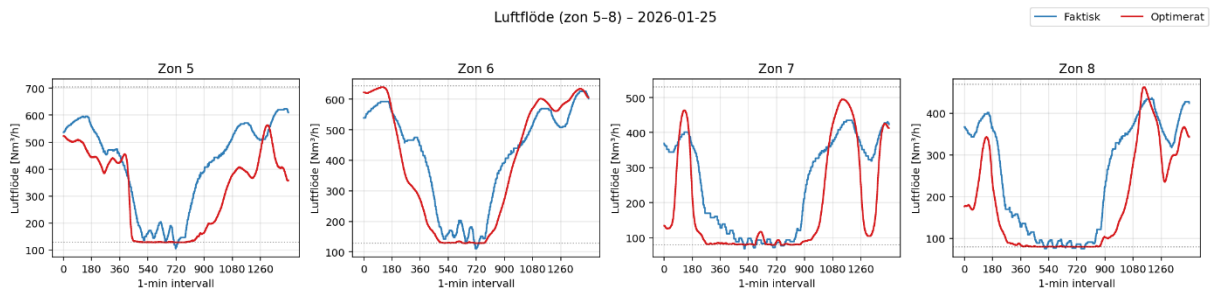
Sammantaget är modellen inte fullständig i absolut mening, men tillräckligt robust för att användas som underlag för simulering och optimering. Den genomgående indikationen är att resultaten från olika delar av analysen pekar i samma riktning. Lustgasutsläppen tycks kunna reduceras genom minskad luftning, och modellen indikerar att en sådan minskning åtminstone i vissa driftfall kan ske utan att utgående ammoniumhalt försämras.

4.4 Optimering och simulering

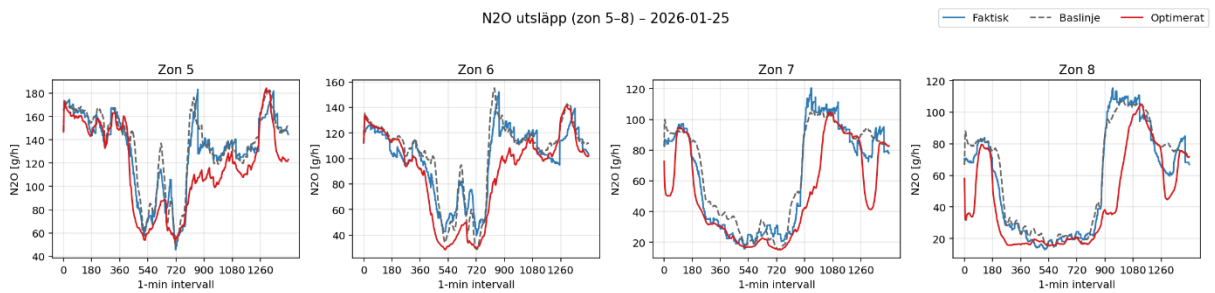
Optimeringen genomfördes på dygnsnivå där luftflödet i varje zon tilläts variera inom data-drivna gränser. Den totala dygnssumman av lustgas användes som målvariabel, medan utgående ammonium i zon 8 utgjorde ett bivillkor som inte fick försämrats jämfört med baslinjen. Optimeringen ska främst ses som ett verktyg för att undersöka vilka driftområden som modellen pekar ut som potentiellt fördelaktiga, snarare än en bra prediktion av utfall. De numeriska reduktionerna nedan ska därför tolkas som indikationer på riktning, inte som verifierade utfall.

4.4.1 Kvantitativt resultat

Resultaten från optimeringen visar att det, enligt maskininlärda modellen, är möjligt att reducera de samlade lustgasutsläppen för zon 5–8 påtagligt, mellan 17–38 %, samtidigt som bivillkoret för ammonium uppfylls. Figur 11–14 visar resultaten för optimering på dagen 2026-01-25. Flera tillgängliga testdagar optimerades, precis som för dygnet 2026-01-25, och resultaten för alla dygn är sammanställda i figur 15 och tabell 5. Dessa dygn var slumpmässigt utvalda.

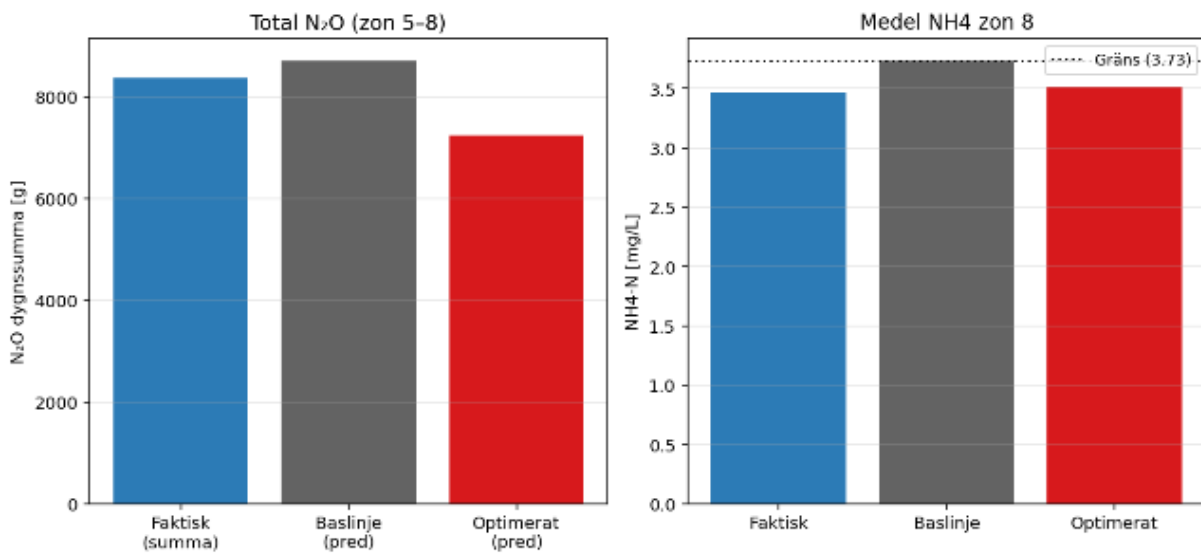


Figur 11. Luftflöde (Nm^3/h) per zon för zon 5–8 under dygnet 2026-01-25. Den blå linjen visar det faktiska luftflödet och den röda linjen visar det optimerade luftflödet enligt modellen. Optimeringen tillåts variera inom data-drivna gränser med målsättningen att minimera dygnssumman av N_2O utan att försämra utgående $\text{NH}_4\text{-N}$ i zon 8.

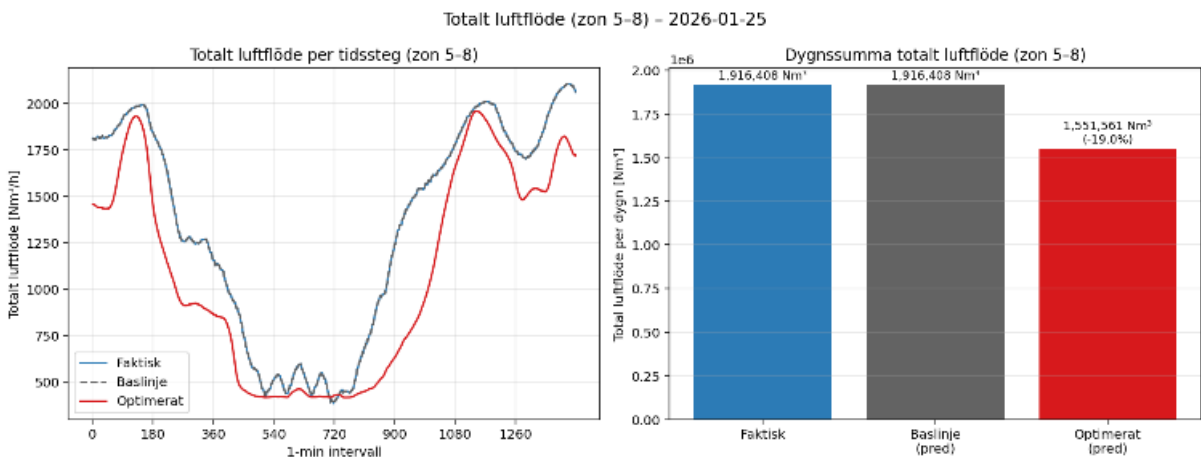


Figur 12. Lustgasutsläpp (N_2O , g/h) per zon för zon 5–8 under dygnet 2026-01-25. Den blå linjen visar det faktiska utsläppet, den grå linjen visar modellens baslinjeprediktion och den röda linjen visar predikerat utsläpp efter optimering av luftflödet.

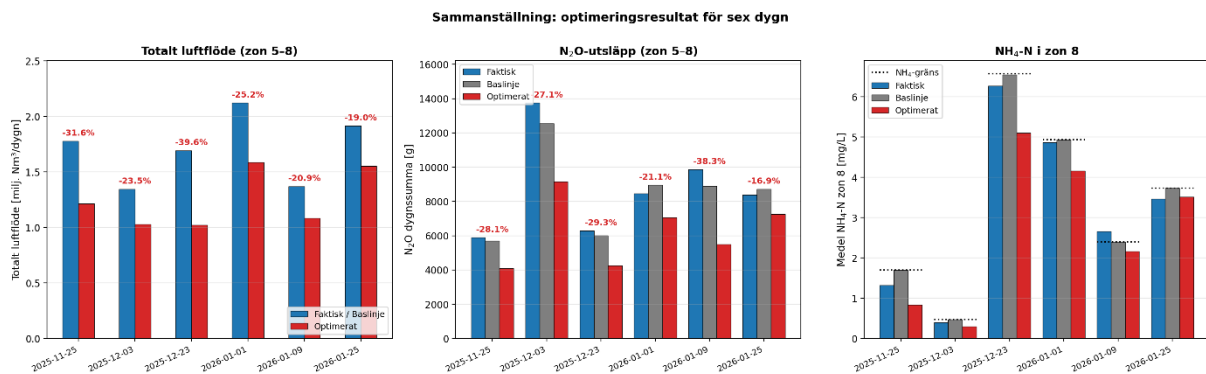
Optimeringssammanfattning (zon 5-8) - 2026-01-25



Figur 13. Sammanfattning av optimeringen för dygnet 2026-01-25. Vänster panel visar total dygns-summa av N₂O (g) för zon 5–8 vid faktiskt drift, baslinje och optimerad lösning. Höger panel visar medel-NH₄-N (mg/L) i zon 8, där bivillkoret innebär att den optimerade lösningen inte får överstiga baslinjens nivå.



Figur 14. Luftflöde under dygnet 2026-01-25 för zon 5–8. Vänster panel visar luftflödesprofilen (Nm³/h) per minut, där den blå linjen är det faktiska luftflödet och den röda linjen det optimerade. Baslinjen döljer sig bakom det faktiska luftflödet då de är identiska. Höger panel visar dygnssumma av totalt luftflöde (Nm³) för zon 5–8 vid faktiskt drift, baslinje och optimerad lösning.



Figur 15. Sammanställning av optimeringsresultat för sex testdygn. Vänster panel visar totalt luftflöde (Nm^3/dygn) för zon 5–8, mittpanelen visar dygnssumma av N_2O -utsläpp (g) och höger panel visar medelvärde av $\text{NH}_4\text{-N}$ (mg/L) i zon 8. Procentuella förändringar mellan baslinje och optimerad lösning är angivna ovanför staplarna. Den röda streckade linjen i höger panel markerar bivillkorets gräns för $\text{NH}_4\text{-N}$.

Tabell 5. Sammanställning av optimeringsresultat för sex testdygn.

Datum	Luft faktisk [Nm^3/dygn]	Luft optimerat [Nm^3/dygn]	Δ luft [%]	N_2O baslinje [g]	N_2O optimerat [g]	Δ N_2O [%]	NH_4 optimerat [mg/L]	NH_4 -gräns [mg/L]
2025-11-25	1 777 218	1 215 650	-31,6%	5 700	4 100	-28,1%	0,83	1,70
2025-12-03	1 345 545	1 028 727	-23,5%	12 560	9 150	-27,1%	0,30	0,47
2025-12-23	1 692 337	1 022 205	-39,6%	6 010	4 250	-29,3%	5,10	6,57
2026-01-01	2 121 644	1 586 832	-25,2%	8 950	7 060	-21,1%	4,16	4,93
2026-01-09	1 369 312	1 083 380	-20,9%	8 900	5 490	-38,3%	2,16	2,40
2026-01-25	1 916 408	1 551 561	-19,0%	8 720	7 250	-16,9%	3,51	3,73
Medel	1 703 744	1 248 059	-26,6%	8 473	6 217	-26,8%	2,68	3,30

Parallellt med optimeringen genomfördes känslighetsanalyser där luftflödet i en zon i taget varierades medan övriga hölls konstanta. Dessa visade att sambandet mellan luftflöde och lustgasutsläpp i princip genomgående är monotont avtagande inom det studerade intervallet, det vill säga att ett lägre luftflöde konsekvent ger lägre predikterad lustgasutsläpp. Inget tydligt lokalt optimum kunde identifieras i någon zon.

4.4.2 Återkoppling till litteraturen

Detta resultat avviker från den teoretiska bilden i avsnitt 2.1, där lustgasbildningen förväntas öka både vid för låg och för hög luftning. Vid syrebrist drivs nitrifikationen mot nitrifierande denitrifikation hos AOB, medan kraftig luftning bidrar till deflockulering, fysisk stripping och störd anoxisk denitrifikation. Teoretiskt bör det därför finnas en mellannivå där lustgasbildningen är som lägst. Att modellen inte återger detta beteende beror sannolikt inte på att fenomenet saknas i verkligheten, utan på en begränsning i underlaget. Som diskuteras i avsnitt 4.7 har maxbörvärdet för löst syre varit i stort sett konstant under hela perioden för den tillgängliga datan, vilket innebär att modellen aldrig fått observera systemets beteende vid tydligt lägre luftningsnivåer. Det monotont avtagande sambandet ska därför tolkas som giltigt inom det driftområde som faktiskt finns representerat i träningsdatan, och inte som en utsaga om att ännu lägre luftning alltid skulle ge ännu lägre utsläpp. För att kunna identifiera den nedre vändpunkten där syrebrist börjar driva upp lustgasbildningen krävs sannolikt en mätkampanj med systematiskt varierat maxbörvärde, vilket diskuteras vidare i kapitlet om framtida arbete. Bilder från känslighetsanalyserna finns i appendix D.

Modellens resultat indikerar att luftning och lustgas följer varandra väldigt nära och monotont. Detta kan vara en indikation på att luftningen i sig leder till att en större del lustgas kan emittera från vattnet via luftbubblorna, mer än att det är syret i vattnet som gör att större lustgasproduktion sker.

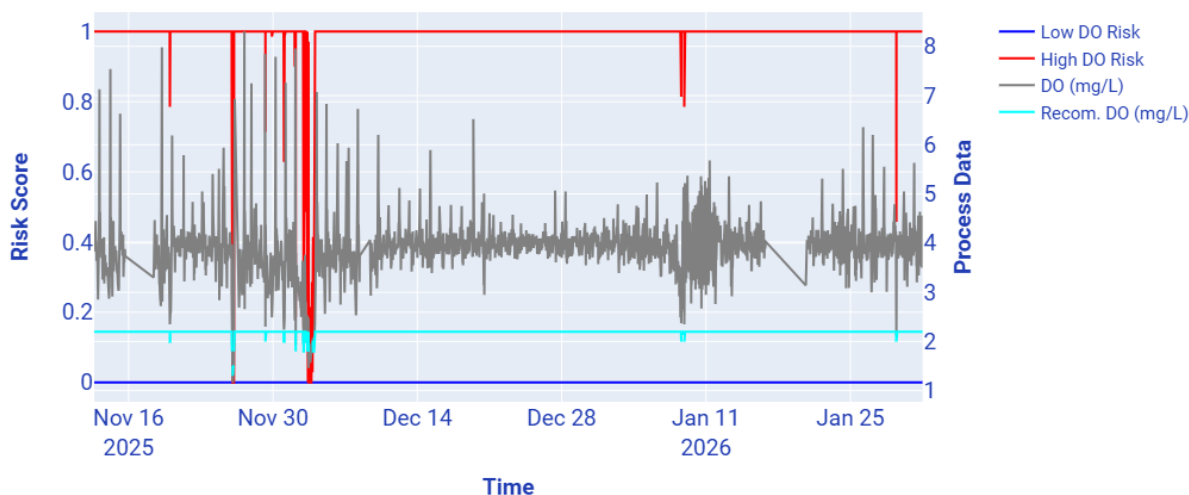
Enklare simuleringsfall där luftflödet i samtliga zoner ändrades med ett fast värde bekräftar huvudriktningen i optimeringen. Ett lägre luftflöde ger generellt minskade lustgasutsläpp, men också stigande ammoniumhalter i zon 8 när luftningen minskas för kraftigt. Detta tydliggör den avvägning som finns mellan minskade växthusgasutsläpp och bibehållen kväveavskiljning, och visar att en mer restriktiv luftningsstrategi bör implementeras gradvis och följas upp noggrant i drift.

Att optimeringen främst reducerar luftflödet är i linje med att luftning i litteraturen beskrivs som en central styrparameter för lustgasutsläpp. Luftning påverkar dels syretillgången för nitrifikation och denitrifikation, dels den fysiska överföringen av redan bildad lustgas från vätskefas till gasfas. Tidigare studier har visat att ändrade syrebörvärden kan ge betydande effekt på lustgasutsläpp, men också att sambandet är anläggnings specifikt och beroende av belastning, flockstruktur och kvävefraktioner (Kampschreur et al., 2009; Law et al., 2012; Duan et al., 2020). Resultatet ska därför inte tolkas som att lägre luftning generellt alltid är bättre, utan som att den studerade driften vid Klagshamn sannolikt ligger på en nivå där en viss reduktion av luftningen är möjlig.

4.5 Jämförelse med Cobalt Water Global

Det parallella arbetet med Cobalt Water Global och deras ”N₂O Risk DSS”-modell gav en kompletterande referenspunkt. Exempelvis visar figur 16 vad Cobalt Waters modell får för resultat när den körs på data från zon 6. Alla zoner gav väldigt liknande resultat och zon 6 visas här för att kunna jämföra modellerna med en av de senare zonerna som innefattar mätningar

för löst syre (DO). Även Cobalt Waters modell indikerar att den löst syrehalten vid Klagshamns reningsverk under den studerade perioden har legat systematiskt över det optimala området. Den rekommenderade börvärdet för löst syre enligt deras analys ligger på cirka 2,2 mg/L, att jämföra med det faktiska maxbörvärdet på 4 mg/L som tillämpats under perioden. Båda modellerna pekar därmed åt samma håll, en lägre luftningsnivå skulle sannolikt minska lustgasutsläppen utan att äventyra reningsresultatet i någon större utsträckning.

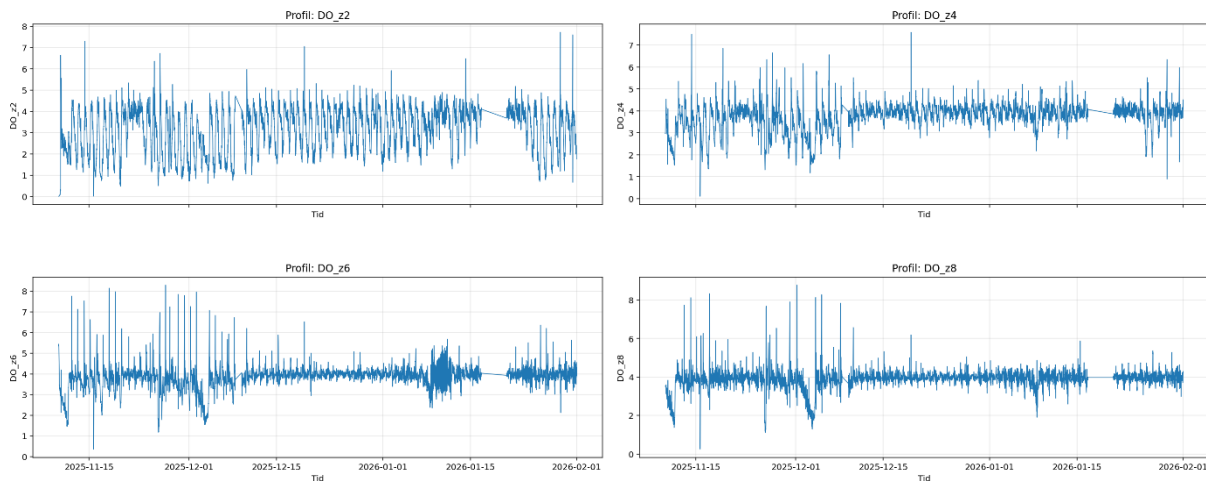


Figur 16. Resultat från Cobalt Waters modell för data tillhörande zon 6. Risken att det är för hög eller för låg koncentration löst syre bedöms med ett riskvärde (vänster y-axel), där 1 är högst risk och 0 är ingen risk. Den röda linjen visar riskvärde för att det är hög koncentration löst syre och den blå linjen visar riskvärde för att det är låg koncentration löst syre. Riskvärdena lägger grund för det uträknade resultatet av rekommenderat börvärde för löst syre, vilket ses av den cyanfärgade linjen i mg/L (höger y-axel). Den gråa linjen visar de uppmätta värdena för löst syre i zon 6.

Att två parallella datadrivna angreppssätt ger samma kvalitativa slutsats stärker trovärdigheten i resultaten, även om modellernas exakta numeriska rekommendationer skiljer sig åt. Under arbetets gång har maxbörvärdet för löst syre vid Klagshamns reningsverk dessutom sänkts från 4 till 3 mg/L, vilket ytterligare bekräftar den praktiska relevansen av resultaten.

4.6 Reglering

Den syrereglering som tillämpas på Klagshamns ARV är, som beskrivits i avsnitt 2.1, uppbyggd som en kaskadstruktur där samtliga regulatorer är PI-reglerade. Figur 17 visar profiler för löst syre (DO) i de fyra zoner där mätning sker. Av figuren framgår att DO-värdena uppvisar tydlig oscillation kring börvärdet och under perioder överstiger maxbörvärdet på 4 mg/L, trots att regulatorerna är inställda att hålla nivån vid eller under detta värde.



Figur 17. Tidsserie av löst syre (DO, mg/L) i de fyra zoner där DO mäts (zon 2, 4, 6 och 8). Profilen visar tydlig oscillation kring börvärdet, där värden återkommande överstiger maxbörvärdet på 4 mg/L. Detta indikerar att den nuvarande PI-regleringen inte är tillräckligt prediktiv för att hantera processens belastningsvariationer.

Eftersom PI-strukturen saknar en deriverande, prediktiv komponent reagerar regulatorn först när en avvikelse redan har uppstått, vilket i kombination med systemets dödtdider kan ge upphov till översläng och oscillation. En övergång till PID-reglering, eller en komplettering med MPC, skulle teoretiskt kunna minska denna oscillation och hålla koncentrationen löst syre närmare börvärdet. Debel et al. (2025) visar i en datadriven simuleringsstudie på ett fullskaligt verk att MPC kan reducera lustgasutsläppen och luftningen avsevärt jämfört med konventionell reglering av löst syre. En reglerteknisk utvärdering ligger utanför ramen för föreliggande arbete men utgör en relevant aspekt för fortsatta studier, eftersom en mer stabil syrereglering kan bidra till både lägre lustgasutsläpp och ett mer förutsägbart processunderlag för datadriven modellering.

4.7 Begränsningar

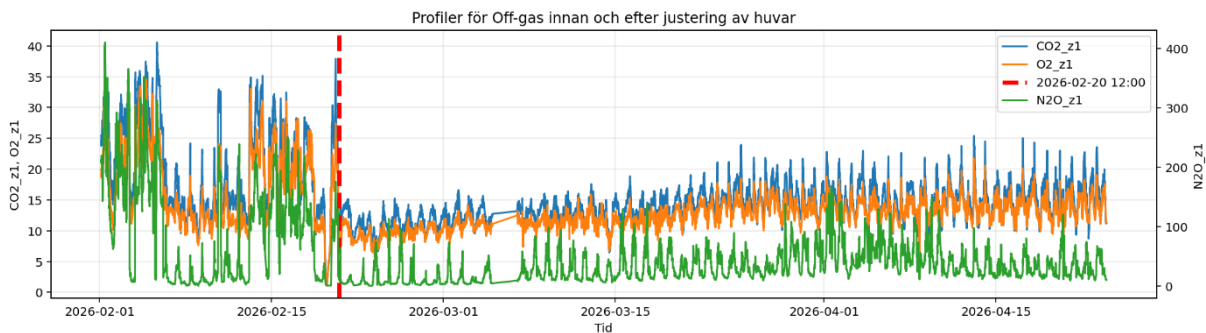
Trots att studien visar tydliga och samstämmiga resultat finns flera begränsningar som bör beaktas vid tolkningen. Den mest betydande är att maxbörvärdet för löst syre varit i princip konstant under hela den tillgängliga datamängden. Detta innebär att modellen inte haft möjlighet att lära sig systemets beteende vid annorlunda driftförhållanden, vilket medför att simulering av lägre luftningsnivåer delvis utgör en extrapolering utanför träningsdatans domän. Det är ingen direkt extrapolering på så sätt att vi håller oss inom de ramar som det finns data för när vi optimerar, men eftersom det inte finns någon data på hur systemet beter sig med lägre maxbörvärde för löst syre så blir det ändå en tydlig form av extrapolering att simulera med lägre luftflöden. De numeriska värdena för reduktion av lustgas, samt bibehållen NH_4 -koncentration i zon 8, bör därför tolkas med viss försiktighet, även om riktningen i resultaten bedöms som robust.

Vidare är modellens prestanda som konstaterats betydligt sämre i zon 1–4 än i zon 5–8. Detta begränsar möjligheten att dra kvantitativa slutsatser för hela linjen och motiverar att optimeringen i första hand genomfördes för zon 5–8. För att förbättra modelleringen av de tidigare zonerna krävs sannolikt både utökad mätutrustning, särskilt kvävefraktioner i vätskefas, och en mer varierad driftdata när det kommer till maxbörvärde för löst syre.

En ytterligare begränsning är att modellen bygger på den givna datamängden med alla dess egenheter, inklusive perioder med kalibreringsavvikelse i ammoniumsignalen och den sekventiella karaktären hos gasanalysatorn. Även om dessa hanterats genom omstrukturering och viktning kvarstår osäkerhet kopplad till datakvaliteten. Korrigeringen av ammoniumsignalen innebär att en del av modellens indata vilar på en estimering snarare än direkt mätning. Resultaten bör därför tolkas med detta i åtanke, särskilt eftersom ammonium in är en central variabel i både korrelationsanalysen och LSTM-modellen. En systematiskt utformad mätkampanj med varierande maxbörvärden och noggrann övervakning av mätutrustningen skulle sannolikt öka modellens tillförlitlighet påtagligt, vilket diskuteras vidare i kapitlet om framtida arbete.

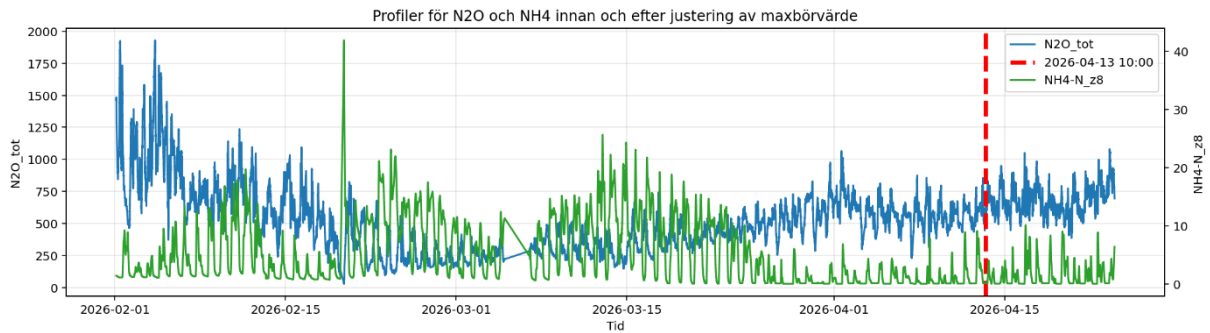
4.8 Jämförelse med ny data

Vid slutet fick vi tillgång till data som samlats in under tiden av detta arbete. Justeringar som vi varit med och påverkat Klagshamns ARV att göra har gjort att denna nya data är intressant att jämföra med den data som huvuddelen av detta arbete är baserat på. Den första justeringen var att huvarna i zon 1 och 2 rättades till från att vara för hårt spända så att de blev sneda. Denna justering av huvarna kan vid enkel analys tyckas gjort att mätningarna blev mer stabila och betedde sig mer som i andra zoner, vilket kan ses för zon 1 i figur 18. Plottarna för zon 2 visar nästan identiska resultat som zon 1 och i jämförelse med andra zoner är det tydligt att zon 1 och 2 uppvisar mer stabilitet efter justeringen, dessa plottar finns i appendix E.



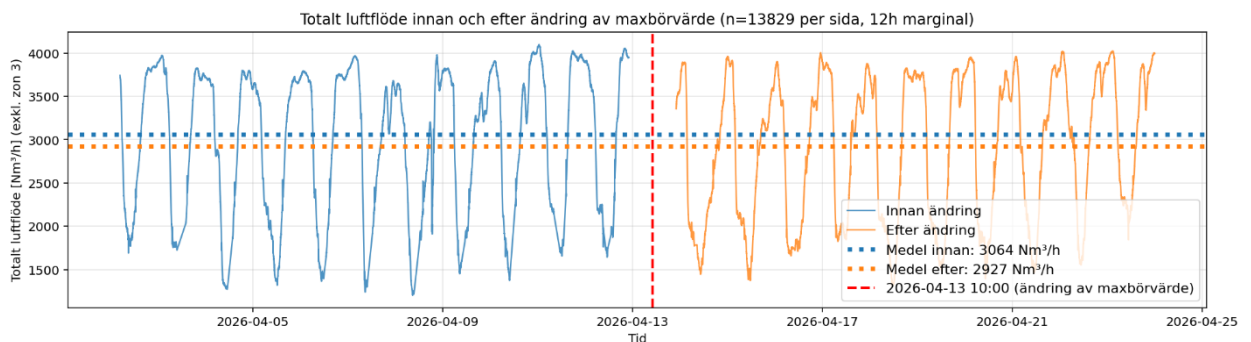
Figur 18. Tidsserie av off-gasmätningar i zon 1 över perioden februari–april 2026. Vänster y-axel är kg/h och höger y-axel är g/h. Den blå linjen visar CO_2 (kg/h), den orangea linjen visar upptaget O_2 (kg/h) och den gröna linjen visar N_2O (g/h). Den röda streckade linjen markerar tidpunkten 2026-02-20 då huvarna i zon 1 och 2 justerades. Mätsignalerna uppvisar tydligt mer stabilt beteende efter justeringen.

Den andra justeringen under arbetets gång var att maxbörvärdet för löst syre i bassängerna ändrades från 4 mg/L till 3 mg/L. Då vi bara har tillgång till en dryg vecka av data från denna period går det inte att dra några säkra slutsatser, men en första analys indikerar att lustgasutsläppen och ammoniumkoncentrationen ligger på ungefär samma nivåer som tidigare, se figur 19. Att utgående NH_4 inte ökat är konsistent med modellens prediktion att ammoniumhalten kan upprätthållas vid en lägre luftningsnivå, vilket är en positiv indikation. Att lustgasnivåerna inte har minskat likt modellernas prediktion kan bero på att den faktiska DO-nivån i bassängen, på grund av oscillation kring börvärdet, i praktiken redan befann sig nära eller under 3 mg/L under delar av dygnet. Det kan även bero på att uppföljningsperioden är för kort för att fånga effekten eller kanske framförallt på de osäkerheter som följer av modellens extrapolering. Det är också värt att nämna att andra saker kan ha större påverkan än vad vi tror, tex regleringen som nämns i 4.6, och att ändringar av maxbörvärde inte kommer att visa några stora förändringar.



Figur 19. Tidsserie av N_2O -utsläpp (g/h totalt, blå linje, vänster axel) och NH_4-N i zon 8 (mg/L, grön linje, höger axel) över perioden februari–april 2026. Vänster y-axel är g/h och höger y-axel är mg/L. Den röda streckade linjen markerar tidpunkten 2026-04-13 då maxbörvärdet för löst syre sänktes från 4 mg/L till 3 mg/L. Ingen tydlig påverkan på vare sig N_2O -utsläpp eller utgående NH_4 -koncentration kan observeras under den korta uppföljningsperioden.

Även om ingen tydlig lustgas-minskning kan påvisas under denna korta period, innebär en sänkning av maxbörvärdet i sig att man borde uppnå en lägre luftningsintensitet och därmed minskad energianvändning. Figur 20 visar jämförelsen av total luftningsintensitet innan och efter ändringen av maxbörvärdet. Här syns inte någon stor skillnad, men medelvärdet minskar med 4,5 % från 3064 till 2927 Nm^3/h . Att bibehålla både rening och utsläppsnivåer vid lägre luftning är därför ett positivt resultat ur driftsynpunkt, även om det krävs en längre uppföljningsperiod för att utvärdera den fulla effekten på lustgasutsläppen.



Figur 20. Totalt luftflöde (Nm^3/h) summerat över zon 1, 2, 4, 5, 6, 7 och 8 under perioden 2026-04-02 till 2026-04-25. Zon 3 är exkluderad eftersom den under denna period styrdes manuellt. Den blå linjen visar luftflödet under perioden innan ändringen av maxbörvärdet, och den orangea linjen visar perioden efter. De prickade horisontella linjerna markerar medelvärdet för respektive period (3064 Nm^3/h innan, 2927 Nm^3/h efter), och den röda streckade linjen markerar tidpunkten 2026-04-13 10:00 då maxbörvärdet för löst syre sänktes från 4 mg/L till 3 mg/L. Lika många datapunkter har använts från båda sidor med en marginal på 12 timmar från ändringstidpunkten. Efter justeringen syns en mindre minskning av det genomsnittliga luftflödet, motsvarande cirka 4,5 %.

5 Slutsats

Syftet med detta arbete har varit att analysera sambanden mellan processparametrar och lustgasutsläpp vid Klagshamns avloppsreningsverk samt att undersöka möjligheten att använda datadriven modellering för att stödja driftoptimering. Resultaten indikerar att luftning, frånluftsp parametrar och hur zonerna påverkar varandra genom systemet är viktigt för att förstå lustgasutsläppen vid Klagshamns ARV. Detta stämmer överens med bilden från litteraturen av att lustgasbildning är en komplex och dynamisk process som påverkas av flera samverkande faktorer.

Den genomförda korrelationsanalysen indikerar att hänsyn till hydraulisk uppehållstid och tidsförskjutning av data är viktigt för att erhålla mer processmässigt relevanta samband. Genom att justera för transporttiden genom systemet framträdde tydligare och mer rimliga korrelationer mellan exempelvis ammoniumhalter och lustgasutsläpp. Detta indikerar att liknande analyser, där det finns flera zoner, kan ge missvisande resultat om man inte tar hänsyn till hydraulisk uppehållstid.

Den utvecklade maskininlärningsmodellen visade att det är möjligt att beskriva och prediktera lustgasutsläpp med hänsyn till processens dynamik. Genom att modellera varje zon separat och samtidigt beakta kopplingarna mellan zonerna kunde en mer realistisk representation av processen uppnås. Samtidigt framkom att modellens tillförlitlighet i hög grad är beroende av datans kvalitet och variation, vilket innebär att resultaten bör tolkas med viss försiktighet, särskilt vid simuleringar utanför det observerade driftområdet. Optimerings- och simuleringsresultaten indikerar att en minskning av luftningen i många fall kan leda till reducerade lustgasutsläpp, utan att omedelbart försämra reningsresultatet. Samtidigt påvisas en tydlig avvägning mellan låga utsläpp och bibehållen kväveavskiljning, där för kraftigt reducerad luftning kommer leda till ökade ammoniumhalter i utgående vatten. Detta understryker vikten av att betrakta processen ur ett helhetsperspektiv, där både miljöpåverkan och reningsprestanda tas i beaktning.

Sammantaget är resultaten konsistenta med den teoretiska bilden av lustgasbildning som en process med ett optimalt driftområde mellan syrebrist och överluftning. Både korrelationsanalysen, modellens identifiering av lokala optima i flera zoner och den empiriska observationen att rening och utsläpp inte försämrats vid sänkt maxbörvärde pekar i samma riktning. Den nuvarande driften befinner sig sannolikt på den högre sidan av detta optimum, och det finns utrymme att försiktigt minska luftningen utan negativa konsekvenser.

6 Framtida arbete

Som konstaterats i detta arbete indikerar resultaten att en lägre luftningsnivå kan vara fördelaktig och bör undersökas vidare. En begränsning i studien är dock avsaknaden av ett dataset där maxbörvärdet för löst syre (DO) systematiskt har varierats under en längre tidsperiod. En mer heterogen datamängd hade sannolikt minskat behovet av extrapolering i den datadrivna modellen och därmed ökat tillförlitligheten i de genererade resultaten. Det är därför rimligt att anta att modellens precision kan förbättras avsevärt med ett dataset som uppvisar större variation. Det hade också varit intressant att undersöka om optimeringen kan anpassas mer avancerat för att försöka minimera extrapolering.

Ett konkret förslag för framtida studier är att systematiskt variera maxbörvärdet för löst syre i steg om 0,5 mg/L, både uppåt och nedåt, under en period om cirka ett år. En sådan tidsram möjliggör att även långsamt varierande faktorer, såsom temperatur, inkluderas i analysen. Under denna period är det av stor vikt att mätutrustningen övervakas noggrant genom frekventa kontroller och kalibreringar. Erfarenheter från detta arbete visar att bristande underhåll av mätutrustning kan leda till betydande dataförluster, vilket i sin tur kräver användning av osäkrare korrigeringsmetoder för att bibehålla tillräcklig datadiversitet. Vid platsbesök noterades att mätutrustningens huvar i vissa fall var placerade för högt över vattenytan, vilket också är något som kan leda till felaktiga koncentrationsbestämningar. Detta bör undersökas vidare, och en analys av variationer i vattennivåerna i bassängerna rekommenderas för att säkerställa att mätutrustningen kontinuerligt är korrekt positionerad. En implementering av automatiserade system för detektion av avvikande mätvärden skulle kunna bidra till snabbare identifiering av felaktiga sensorer och därmed avlasta en stor del av arbetskraften för att hålla igång ett så långvarigt projekt.

Vidare skulle installation av sensorer för specifika kvävefraktioner, såsom ammonium och nitrit, möjliggöra en mer detaljerad modellering av nitrifikationsförloppet längs bassängernas pluggflöde. Samtidigt bör nyttan av ytterligare mätutrustning vägas mot de ökade kostnader och det merarbete som installation och underhåll medför.

En aspekt som inte behandlats i denna studie är det ekonomiska incitamentet kopplat till luftningsstrategin. Även om luftkompressionen vid Klagshamns reningsverk drivs med grön energi, innebär minskad luftning inte enbart potentiellt reducerade klimatutsläpp utan även lägre driftkostnader, vilket gör frågan relevant ur ett ekonomiskt perspektiv.

Som diskuterats i avsnitt 4.6 uppvisar den nuvarande syrereglering vid Klagshamns ARV oscillerande beteende, vilket sannolikt bidrar till att nivåerna för löst syre periodvis avviker från börvärdet på ett sätt som kan påverka lustgasbildningen. En reglerteknisk utvärdering av kaskadstrukturen rekommenderas därför som en del av framtida arbete. En sådan studie bör innefatta analys av processens dödtider, brusnivåer i signalerna för löst syre samt en jämförelse mellan olika regulator typer, exempelvis PI, PID och prediktiva regulatorer som MPC.

Vidare vore det av intresse att undersöka om en mer stabil styrreglering kan minska variabiliteten i lustgasutsläppen och därmed också utgöra en bättre grund för datadriven modellering, eftersom mer stabila driftförhållanden skulle göra det enklare att urskilja effekter av enskilda processparametrar.

7 Referenser

- Ahn, J. H., Kim, S. P., Park, H. K., Rahm, B., Pagilla, K., & Chandran, K. (2010). N₂O emissions from activated sludge processes, 2008–2009: Results of a national monitoring survey in the United States. *Environmental Science & Technology*, 44(12), 4505–4511.
- Colliver, B. B., & Stephenson, T. (2000). Production of nitrogen oxide and dinitrogen oxide by autotrophic nitrifiers. *Biotechnology Advances*, 18(3), 219–232.
- Daims, H., Lebedeva, E. V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N., Palatinszky, M., Vierheilig, J., Bulaev, A., Kirkegaard, R. H., von Bergen, M., Rattei, T., Bendinger, B., Nielsen, P. H., & Wagner, M. (2015). Complete nitrification by *Nitrospira* bacteria. *Nature*, 528, 504–509.
- Debel, L., Stentoft, P. A., Stokholm-Bjerregaard, M. A., Ortiz-Arroyo, D., & Durdevic, P. (2025). Model predictive control for nitrous oxide mitigation in a full-scale wastewater treatment plant: A data-driven simulation study. *Journal of Water Process Engineering*, 76, 108128.
- Duan, H., Ye, L., Erler, D., Ni, B.-J., & Yuan, Z. (2020). Quantifying nitrous oxide production pathways in wastewater treatment systems using isotope technology – A critical review. *Water Research*, 185, 116232.
- Fang, X., Li, J., Gao, C., Li, C., & Wang, Q. (2026). Applicability of machine learning in predicting nitrous oxide emissions from wastewater treatment: A review. *Journal of Environmental Management*, 373, 123787.
- Feng, Q., Tai, X., Sun, Y., & Li, M. (2019). Influence of turbulent mixing on the composition of extracellular polymeric substances (EPS) and aggregate size of aerated activated sludge. *Chemical Engineering Journal*, 378, 122123.
- Henze, M., Gujer, W., Mino, T., & van Loosdrecht, M. C. M. (2000). *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3*. IAWQ Scientific and Technical Report No. 9. London: IWA Publishing.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- Hwangbo, S., Al, R., Chen, X., & Zhuang, K. (2021). Integrated model for understanding N₂O emissions from wastewater treatment plants: A deep learning approach. *Environmental Science & Technology*, 55(3), 2143–2151.

- IPCC. (2021). *Climate Change 2021: The Physical Science Basis*. Cambridge University Press.
- Jin, B., & Lant, P. (2004). Flow regime, hydrodynamics, floc size distribution and sludge properties in activated sludge bubble column, air-lift and aerated stirred reactors. *Chemical Engineering Science*, 59, 2379–2388.
- Kampschreur, M. J., Temmink, H., Kleerebezem, R., Jetten, M. S. M., & van Loosdrecht, M. C. M. (2009). Nitrous oxide emission during wastewater treatment. *Water Research*, 43(17), 4093–4103.
- Kemmou, L., & Amanatidou, E. (2023). Factors affecting nitrous oxide emissions from activated sludge wastewater treatment plants—A review. *Resources*, 12(9), 114.
- Kinsley, H., & Kukiela, D. (2020). *Neural Networks from Scratch in Python*. Egen utgivning. Tillgänglig på www.nnfs.io
- Law, Y., Ye, L., Pan, Y., & Yuan, Z. (2012). Nitrous oxide emissions from wastewater treatment processes. *Philosophical Transactions of the Royal Society B*, 367, 1265–1277.
- Li, B., & Bishop, P. L. (2004). Micro-profiles of activated sludge floc determined using microelectrodes. *Water Research*, 38(5), 1248–1258.
- Lotito, A. M., Wunderlin, P., Joss, A., Kipf, M., & Siegrist, H. (2012). Nitrous oxide emissions from the nitrification step of a pilot activated sludge plant. *Water Research*, 46(13), 4191–4200.
- Massara, T. M., Malamis, S., Guisasola, A., Baeza, J. A., Noutsopoulos, C., & Katsou, E. (2017). A review on nitrous oxide (N₂O) emissions during biological nutrient removal from municipal wastewater and sludge reject water. *Science of the Total Environment*, 596–597, 106–123.
- Mehrani, M.-J., Bagheri, M., Bagheri, M., Naderi, A., Sobotka, D., Kowal, P., Ciesielski, S., & Makinia, J. (2022). Application of hybrid mechanistic/machine-learning model for predicting nitrous oxide (N₂O) production in a nitrifying sequencing batch reactor. *Process Safety and Environmental Protection*, 162, 1015–1025.
- Montgomery, D. C., & Runger, G. C. (2014). *Applied Statistics and Probability for Engineers* (6 uppl.). Wiley.
- Porro, J., Bellandi, G., Rodriguez-Roda, I., Deeke, A., Weijers, S., Vanrolleghem, P., Comas, J., & Nopens, I. (2017). Developing an artificial intelligence-based WRRF nitrous oxide mitigation road map: The Eindhoven N₂O mitigation case study. Water Environment Federation Technical Exhibition and Conference 2017, WEFTEC 2017, 1568–1580.

Schreiber, F., Wunderlin, P., Udert, K. M., & Wells, G. F. (2012). Nitric oxide and nitrous oxide turnover in natural and engineered microbial communities: Biological pathways, chemical reactions, and novel technologies. *Frontiers in Microbiology*, 3, 372.

Seshan, H., Milferstedt, K., Hamelin, J., Desmond, P., & Batstone, D. (2024). Forecasting nitrous oxide emissions from a full-scale wastewater treatment plant using a long short-term memory (LSTM) model. *Water Research*, 252, 121197.

VA SYD. (2023). *Miljörapport Klagshamns avloppsreningsverk 2023 (Miljörapport)*.

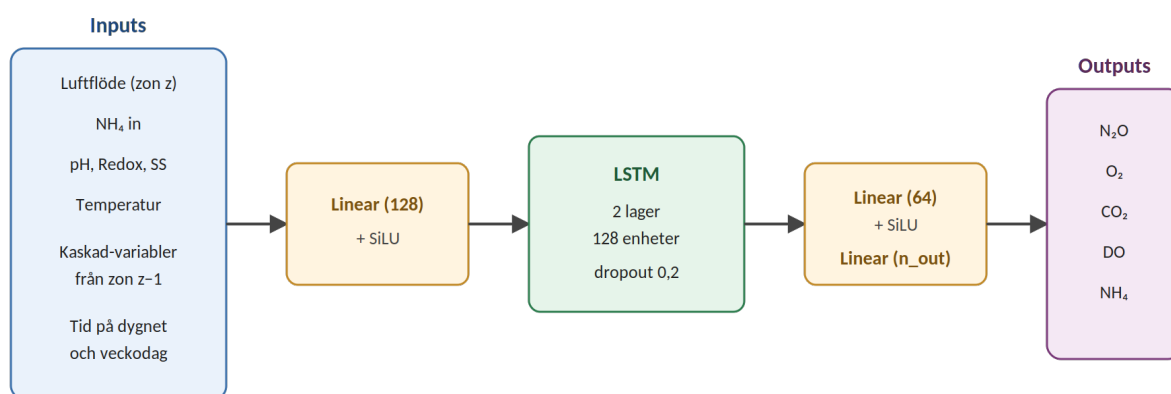
Wunderlin, P., Mohn, J., Joss, A., Emmenegger, L., & Siegrist, H. (2012). Mechanisms of N₂O production in biological wastewater treatment under nitrifying and denitrifying conditions. *Water Research*, 46(4), 1027–1037.

Åmand, L. (2013). *Control of aeration in activated sludge processes (Doktorsavhandling)*. Lunds universitet.

8 Appendix

A. Detaljerad LSTM-metod

Varje zon representeras av en separat LSTM-modell med multipla outputs. Inputs består av zonens luftflöde, externa processvariabler (NH₄ in, pH, redoxpotential, suspenderad substans och temperatur), kaskad-variabler från föregående zon samt tidsfeatures som beskriver tid på dygnet och veckodag. Outputs består av N₂O, O₂ och CO₂ för alla zoner samt DO och NH₄ i de zoner där dessa mätningar finns.



En modell per zon. Outputs från zon z matas in som kaskad-variabler till zon z+1.

Figur A. Översikt av LSTM-modellens arkitektur per zon.

De linjära lagren följs av en SiLU-aktiveringsfunktion (Sigmoid Linear Unit, även kallad swish), definierad som $f(x) = x \cdot \sigma(x)$. Aktiveringsfunktionen introducerar en icke-linjäritet i nätverket, vilket gör att modellen kan fånga mer komplexa samband än vad rent linjära lager skulle tillåta.

Modellen tränades med viktad medelkvadratfel (MSE) som förlustfunktion. MSE beräknas som medelvärdet av de kvadrerade skillnaderna mellan predikterade och uppmätta värden, och är ett standardval för regressionsproblem eftersom den straffar stora avvikelser hårdare än små. Två viktningar tillämpades parallellt: en per-output-vikt där NH₄ prioriterades med faktor 3 jämfört med övriga storheter, samt en per-tidssteg-vikt som avtar linjärt från en färsk gasmätning ned till värdet mitt emellan två gasanalysator-cykler (se avsnitt 3.3.3). Båda viktningarna normaliserades så att medelvikten = 1, vilket gör loss-storleken jämförbar mellan körningar.

Parameter	Värde
Modelltyp	LSTM (rekurrent neuralt nätverk)
Antal modeller	En per zon
Tidsupplösning	1 minut, 1440 tidssteg per dygn
Input-projektion	Linear (128) + SiLU
LSTM-lager	2 lager, 128 enheter, dropout 0,2
Output-head	Linear (64) + SiLU + Linear (n_outputs)
Optimerare	Adam
Inlärningshastighet	$1 \cdot 10^{-3}$
Batch-storlek	4 dygnssekvenser
Maximalt antal epoker	200
Early stopping (patience)	12 epoker
Förlustfunktion	Viktad MSE
Output-vikt NH ₄	3,0 (övriga outputs: 1,0)
Tidsviktning	Linjärt avtagande från senaste gasmätning
Train/test-split	80 % / 20 % (slumpmässigt per dygn)
Primer-kontext	120 minuter från föregående dygn

B. Detaljerad optimeringsmetod

Beslutsvariablerna parametriserades via en logit-transformation enligt

$$Q_{\text{luft},z}(t) = Q_{\text{min},z} + (Q_{\text{max},z} - Q_{\text{min},z}) \cdot \sigma(x_z(t))$$

Målfunktionen som minimerades formulerades som

$$L = \sum_{z \in \{5, \dots, 8\}} \sum_t \frac{\widehat{N_2O}_z(t)}{60} + \lambda_{\text{NH}_4} \cdot \max(0, \widehat{NH}_{4,z8} - C_{\text{NH}_4})^2 + \lambda_s \cdot \frac{1}{N} \sum_{z,t} (Q_{\text{luft},z}(t+1) - Q_{\text{luft},z}(t))^2$$

där den första termen är total dygnssumma av predikterad N_2O i zon 5–8 (uttryckt i gram), den andra termen är ett kvadratisk straff för överskridande av ammoniumgränsen C_{NH_4} i zon 8, och den tredje termen straffar stora förändringar i luftflöde mellan tidssteg för att undvika orealistiska oscillationer i lösningen.

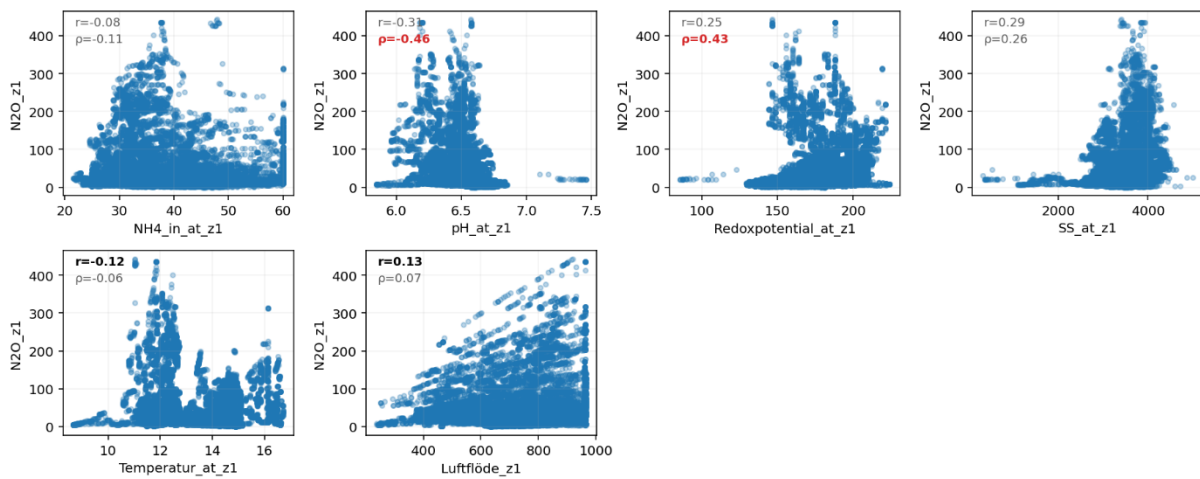
Parameter	Värde
Beslutsvariabler	4 zoner \times 1440 minuter = 5760
Optimerare	Adam
Inlärningshastighet	0,1
Iterationer per dygn	50
Gradientklippning (maxnorm)	1,0
λ_{NH_4} (NH4-straff)	1,0
λ_s (glatthetsstraff)	10
$Q_{\text{min},z}$ och $Q_{\text{max},z}$	3:e resp. 99:e percentilen per zon
Startpunkt	Historiskt luftflöde
Slutgiltig lösning	Iteration med lägst totalkostnad

C. Korrelationsanalys – Scatterplots

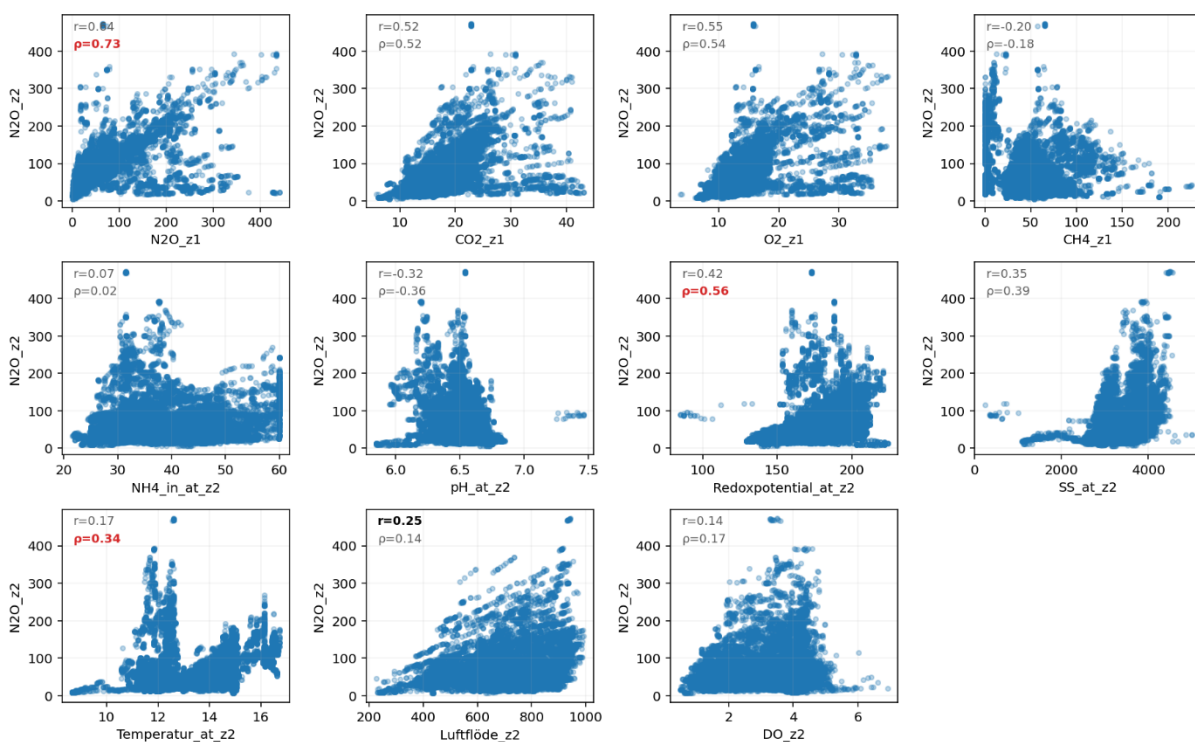
Först visas scatterplots för korrelationsanalys av olika variabler mot N₂O i varje zon. Här kan man exempelvis se hur CH₄ har betydligt sämre korrelation än övriga frånluftsvariabler. Efter detta visas scatterplots för korrelation mot fler outputs i LSTM-modellen: upptaget O₂, CO₂ och NH₄-N.

a) Korrelation mot N₂O

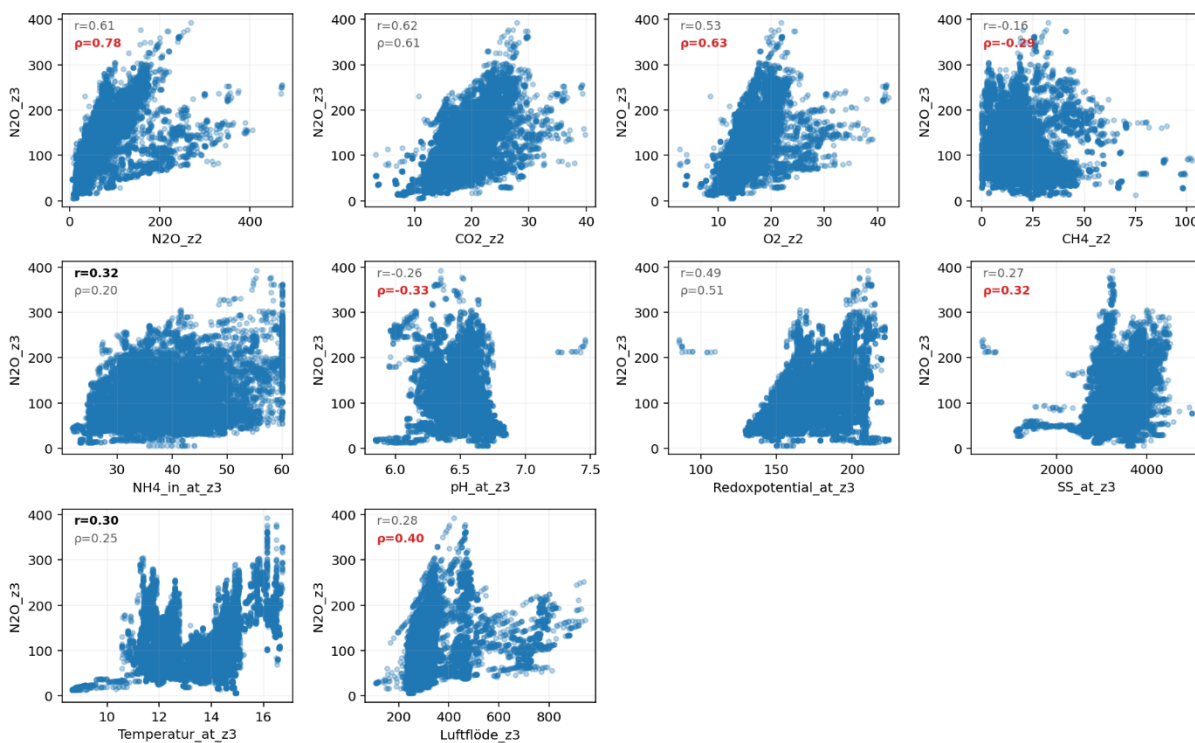
Zon 1 - Scatter mot N2O_z1 (n=20000)



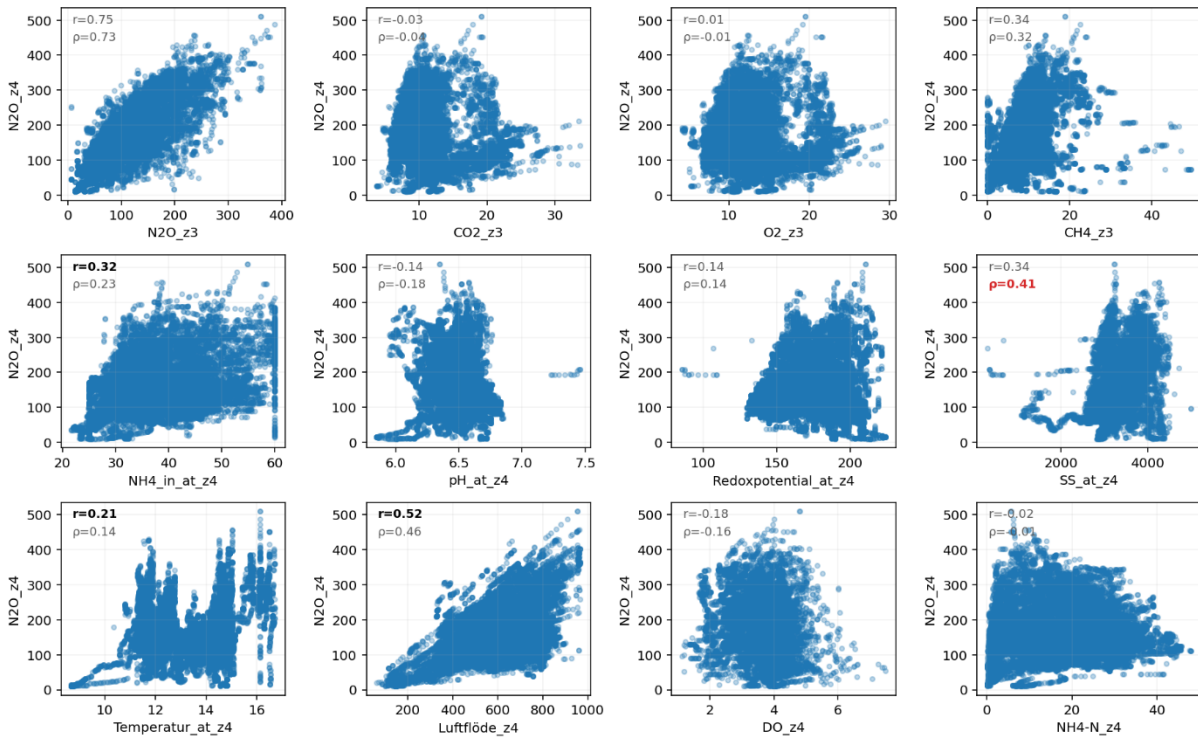
Zon 2 - Scatter mot N2O_z2 (n=20000)



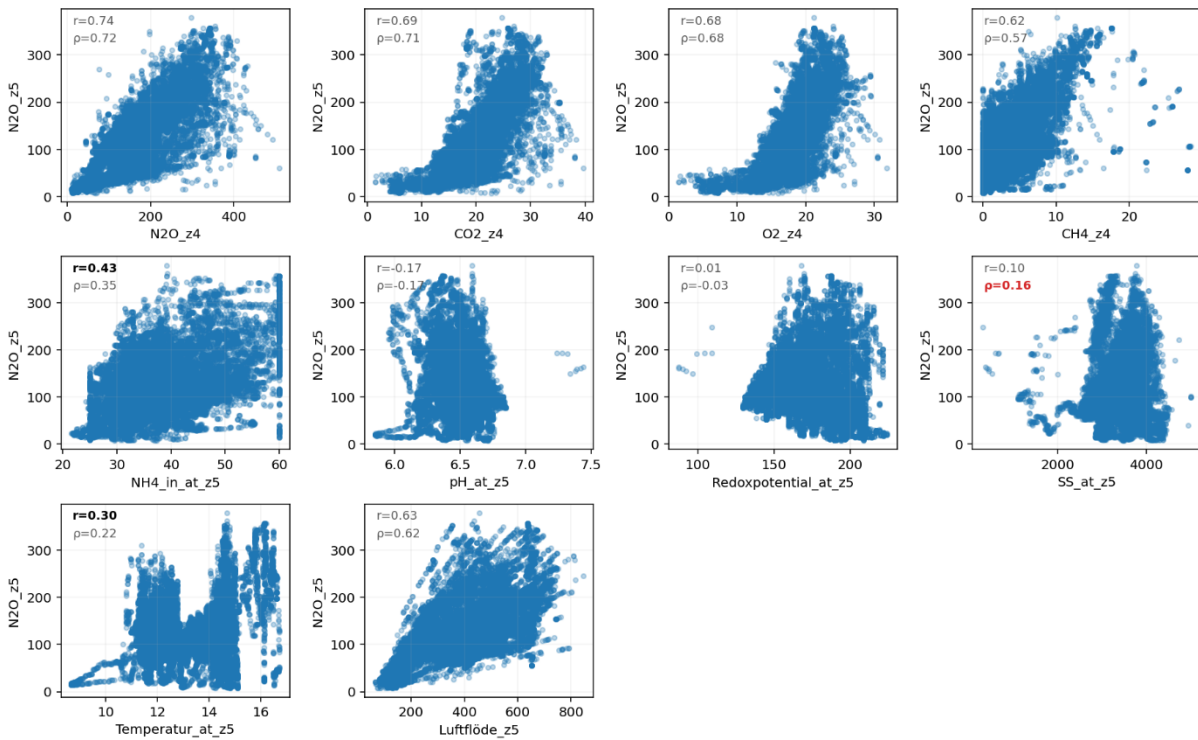
Zon 3 - Scatter mot N2O_z3 (n=20000)



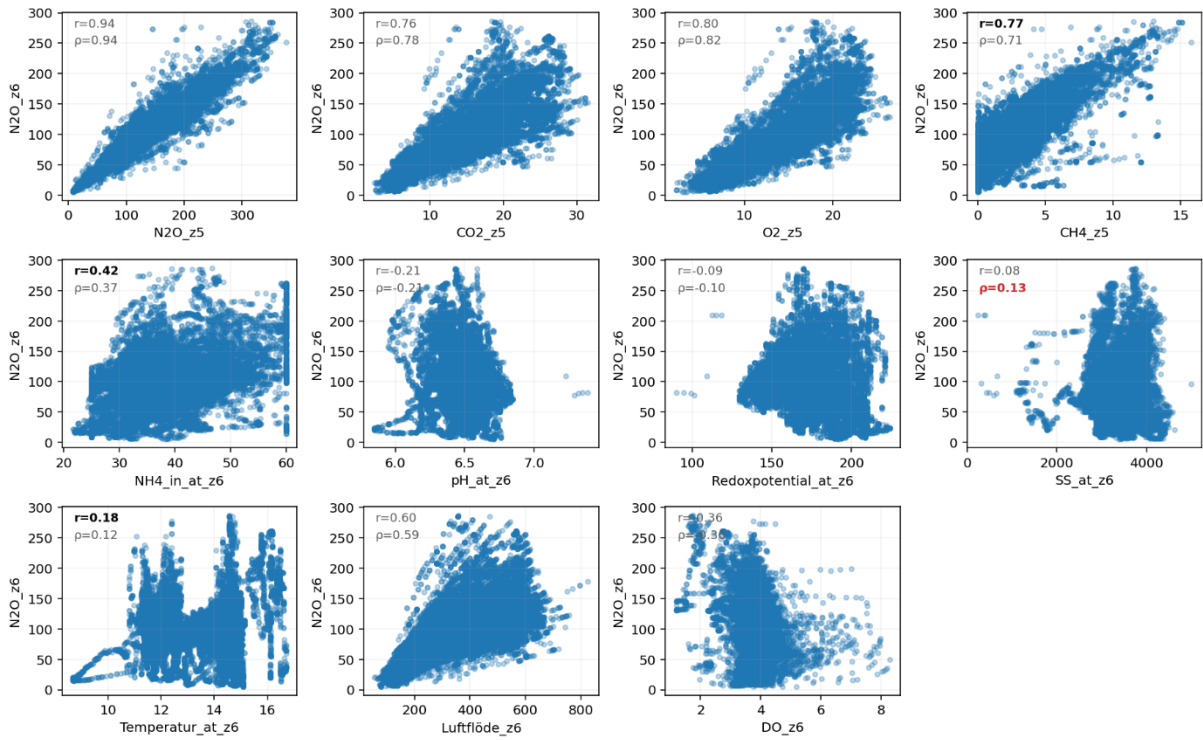
Zon 4 - Scatter mot N2O_z4 (n=20000)



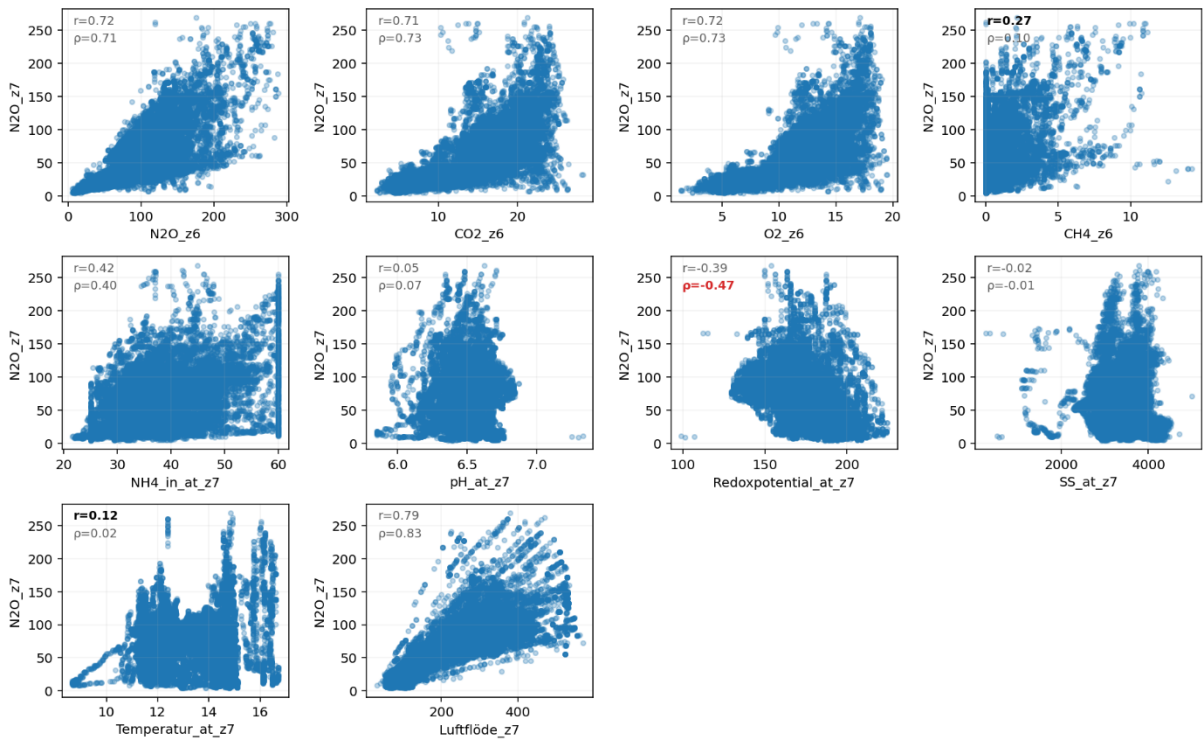
Zon 5 - Scatter mot N2O_z5 (n=20000)



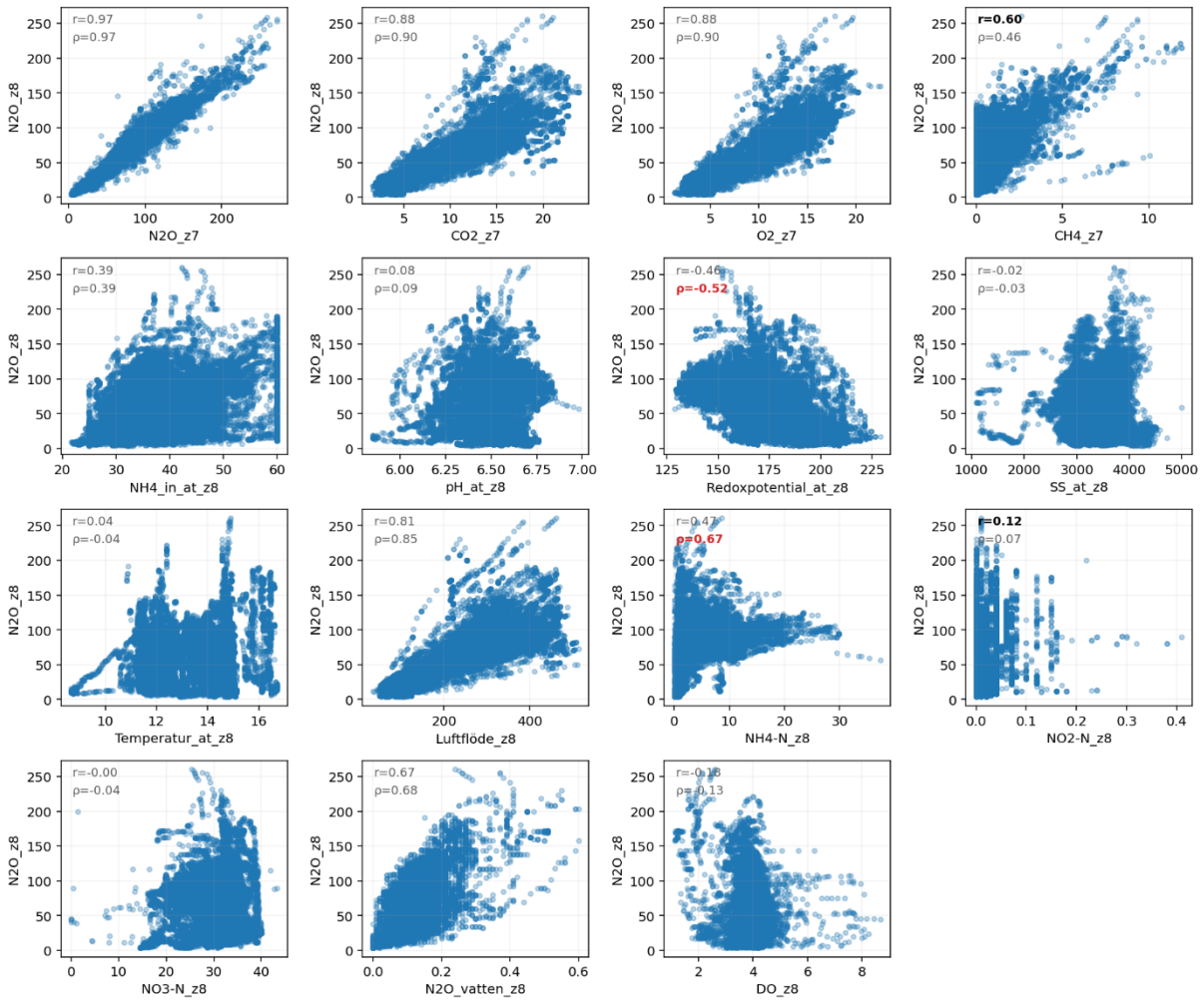
Zon 6 - Scatter mot N2O_z6 (n=20000)



Zon 7 - Scatter mot N2O_z7 (n=20000)

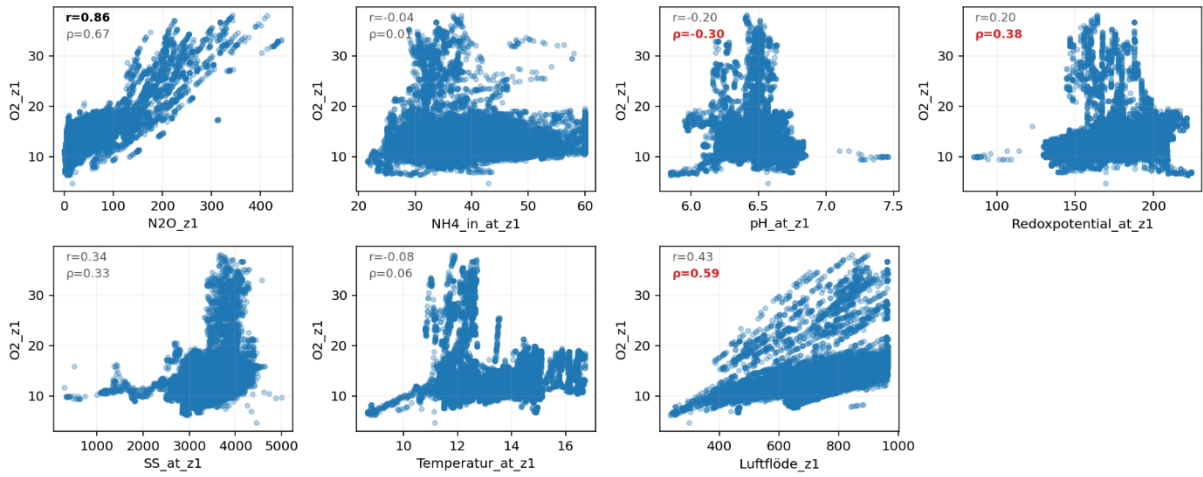


Zon 8 - Scatter mot N2O_z8 (n=20000)

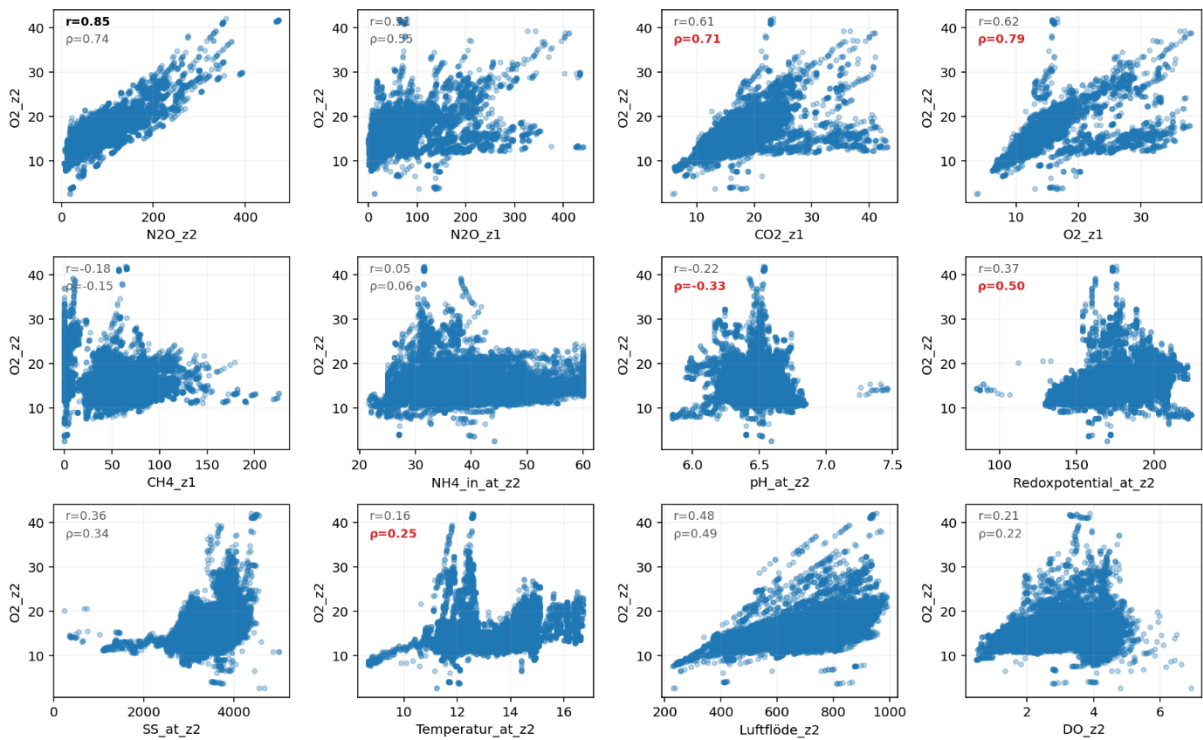


b) Korrelation mot O₂

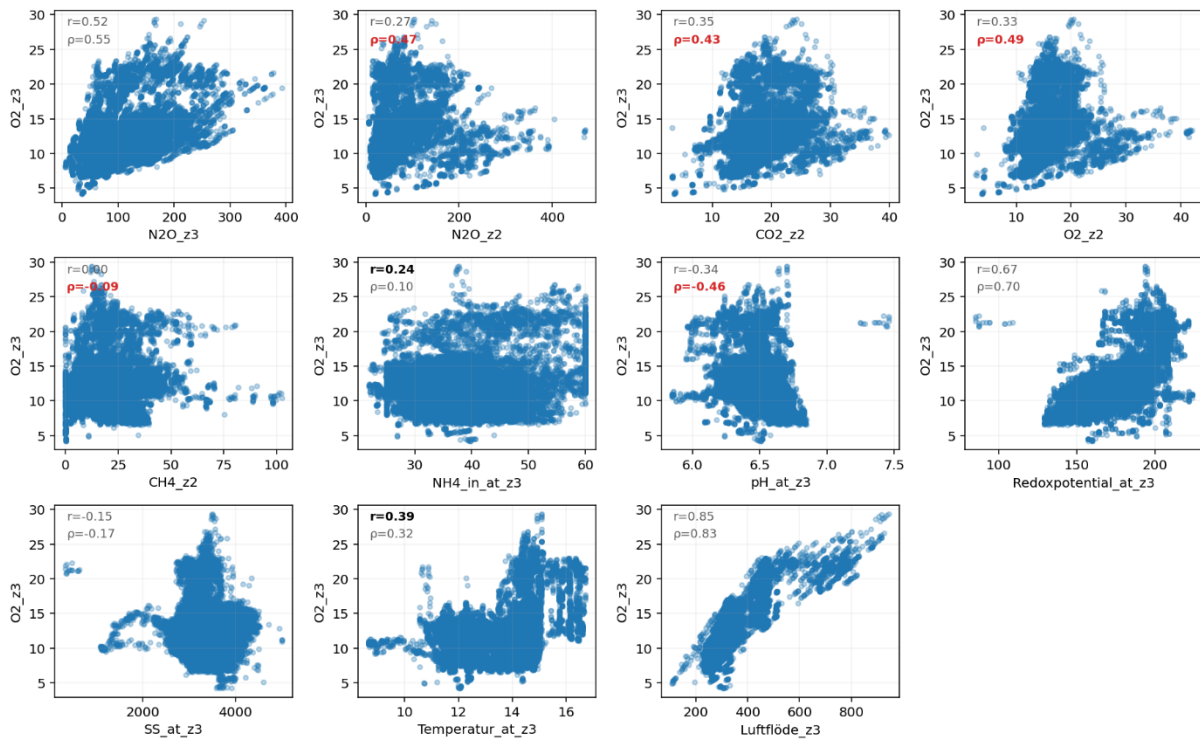
Zon 1 - Scatter mot O₂_z1 (n=20000)



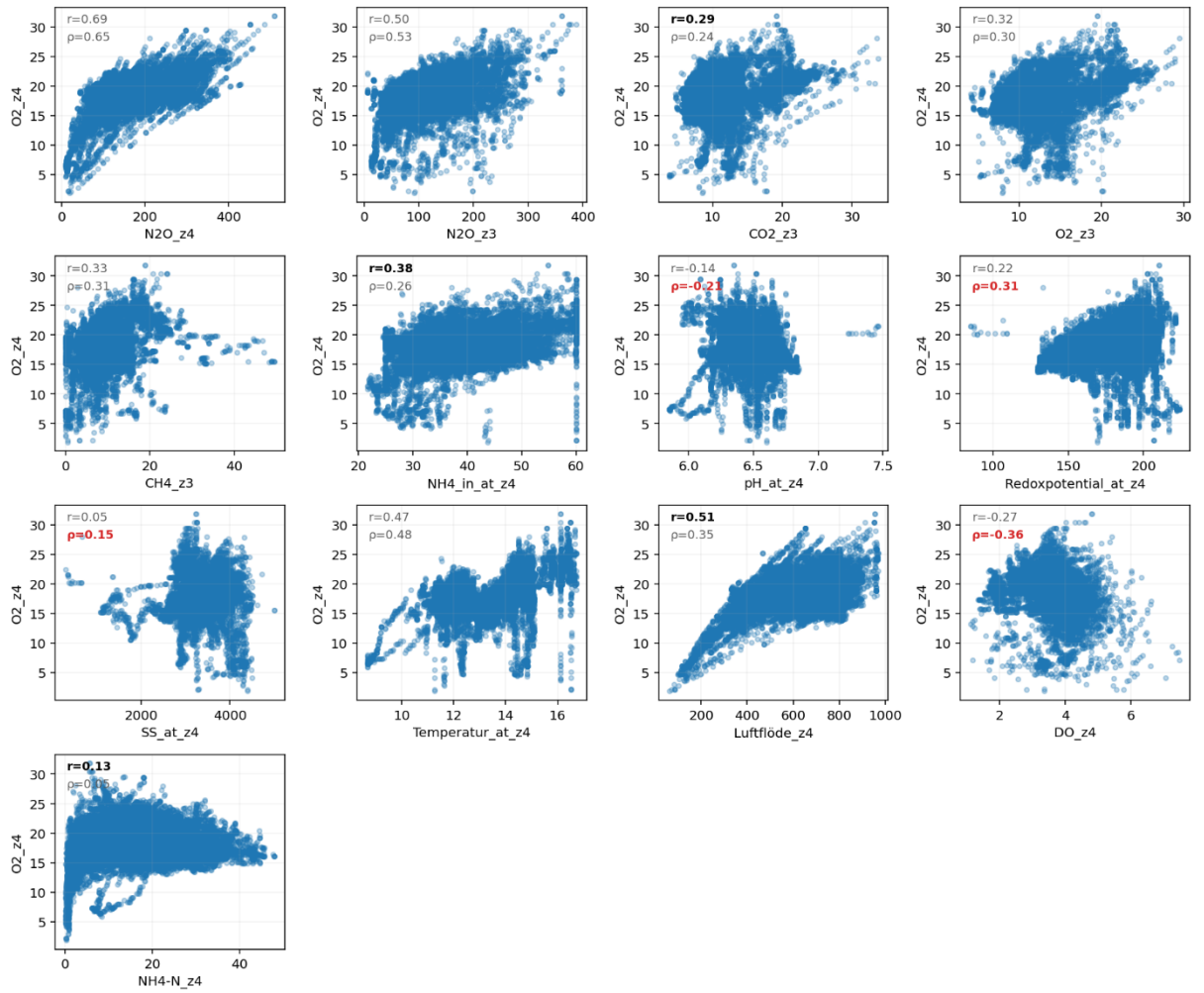
Zon 2 - Scatter mot O₂_z2 (n=20000)



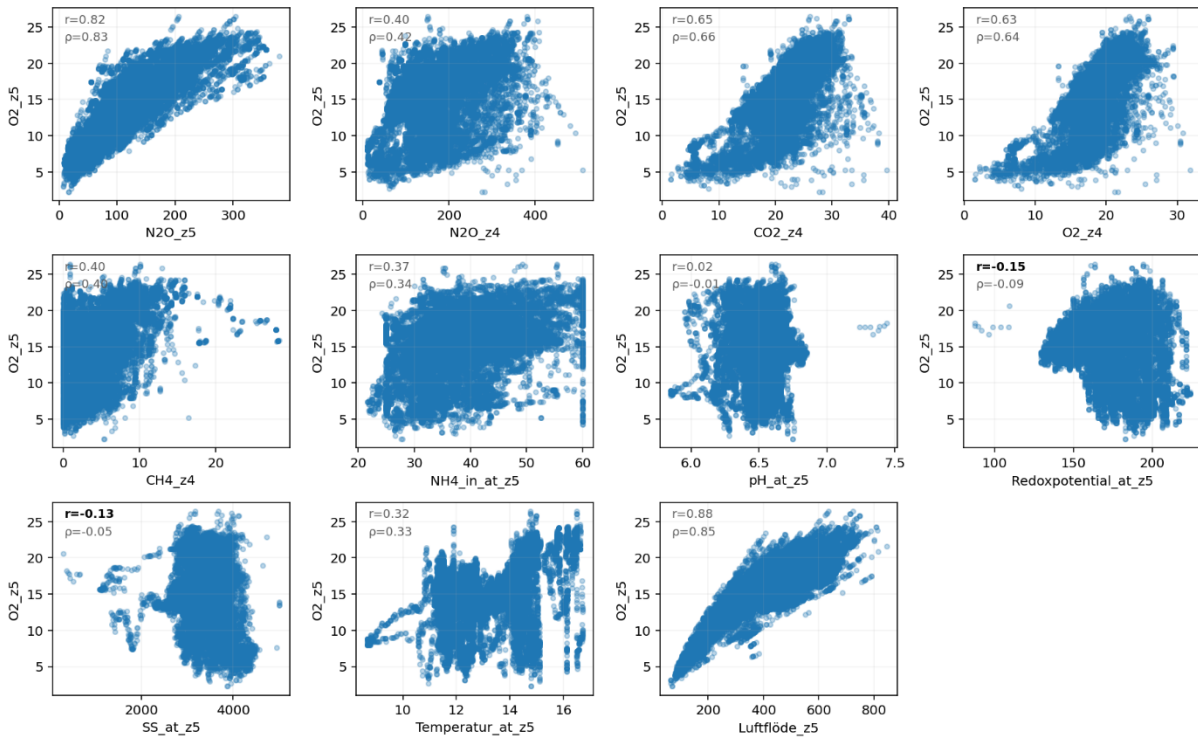
Zon 3 - Scatter mot O2_z3 (n=20000)



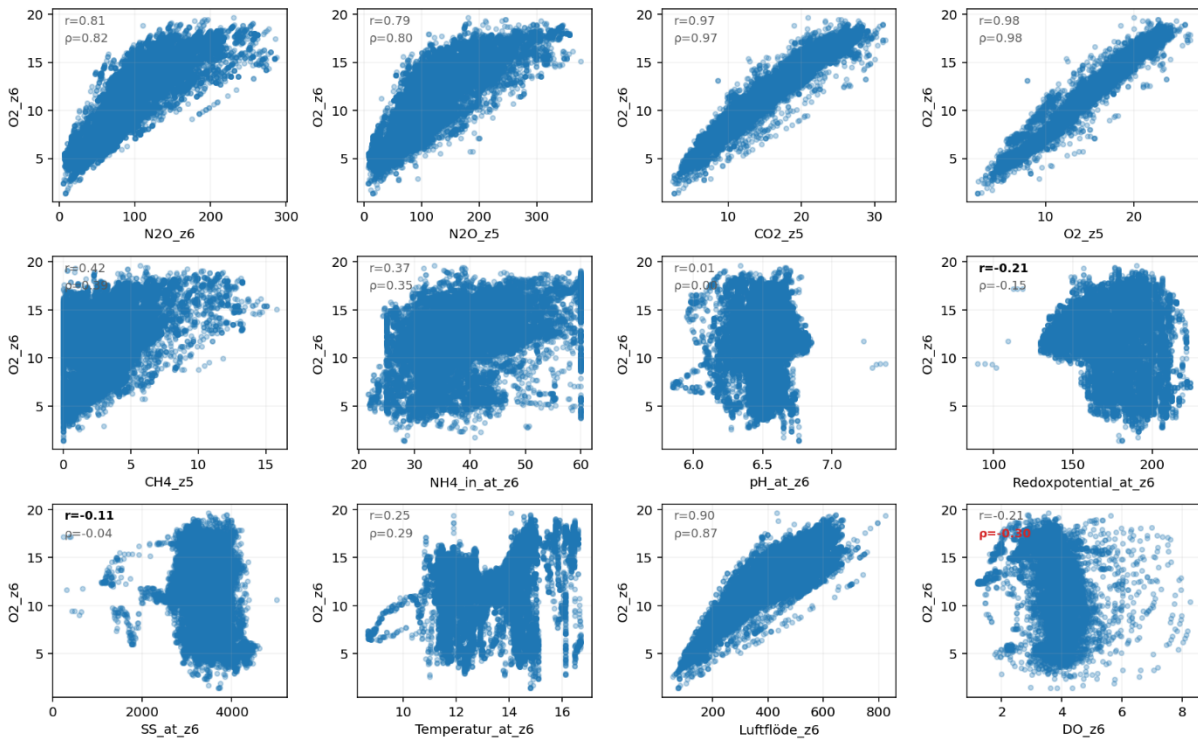
Zon 4 - Scatter mot O2_z4 (n=20000)



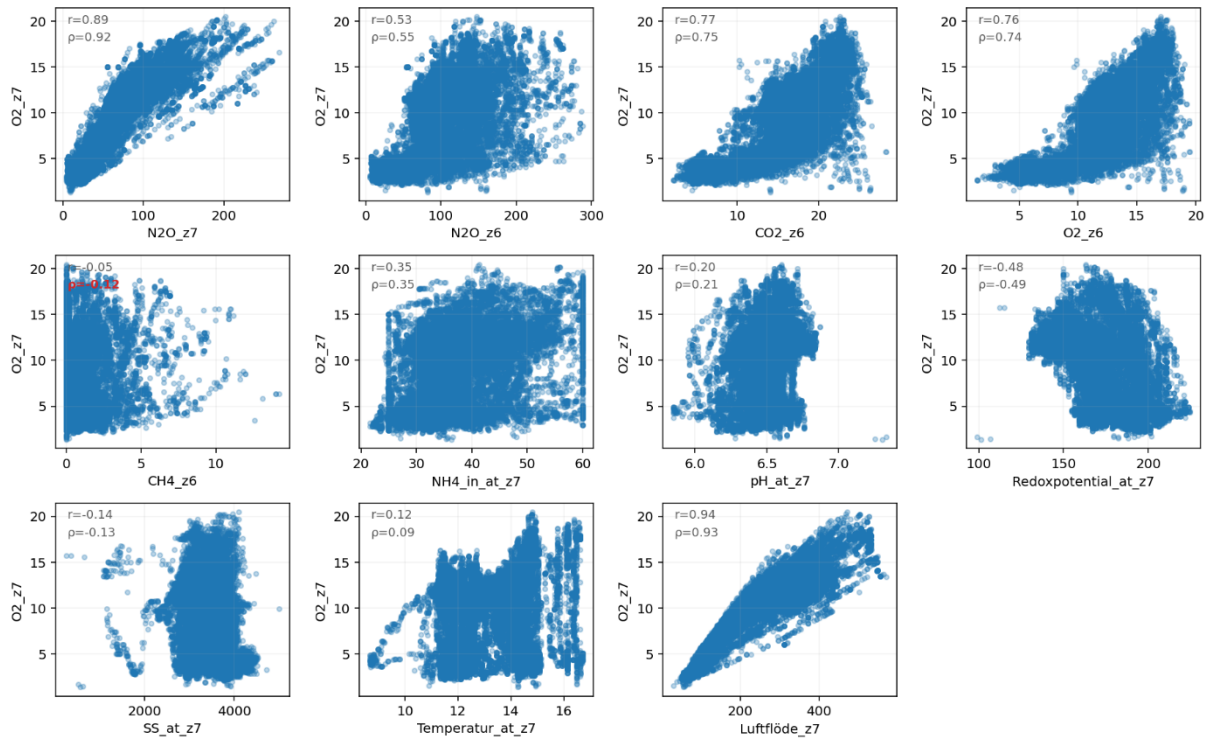
Zon 5 - Scatter mot O2_z5 (n=20000)



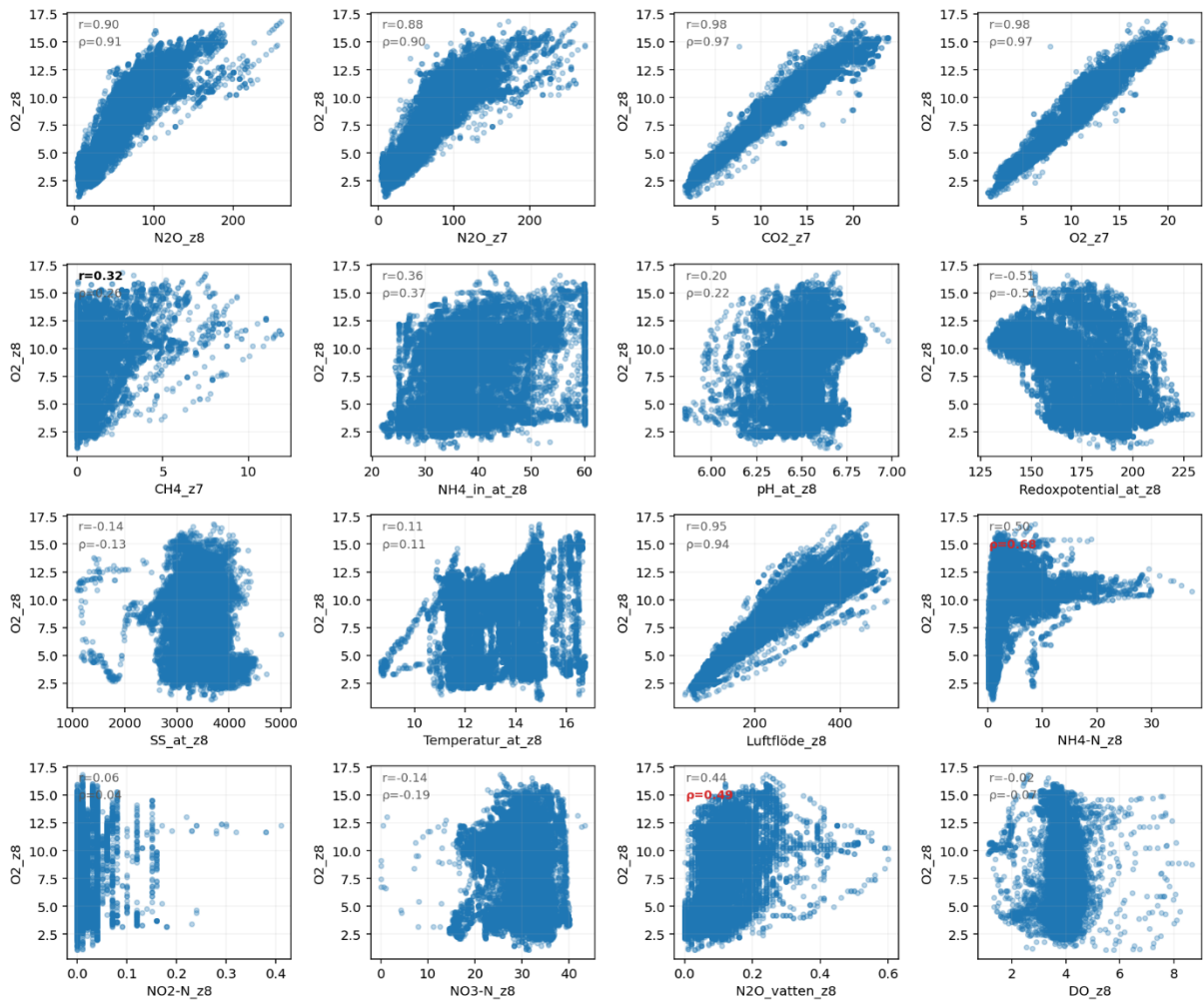
Zon 6 - Scatter mot O2_z6 (n=20000)



Zon 7 - Scatter mot O2_z7 (n=20000)

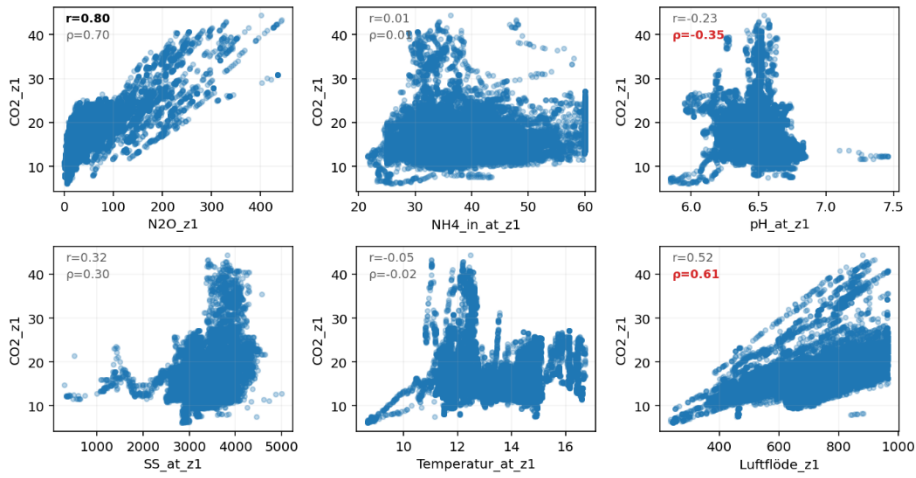


Zon 8 - Scatter mot O2_z8 (n=20000)

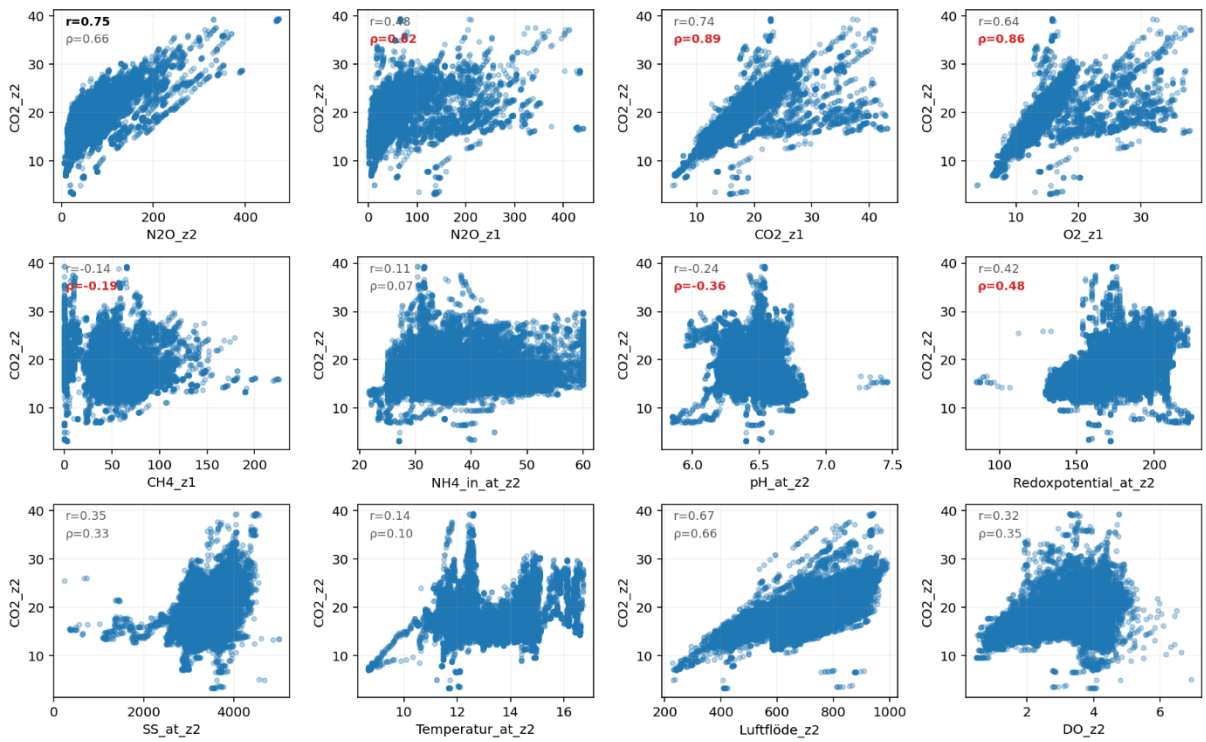


c) Korrelation mot CO₂

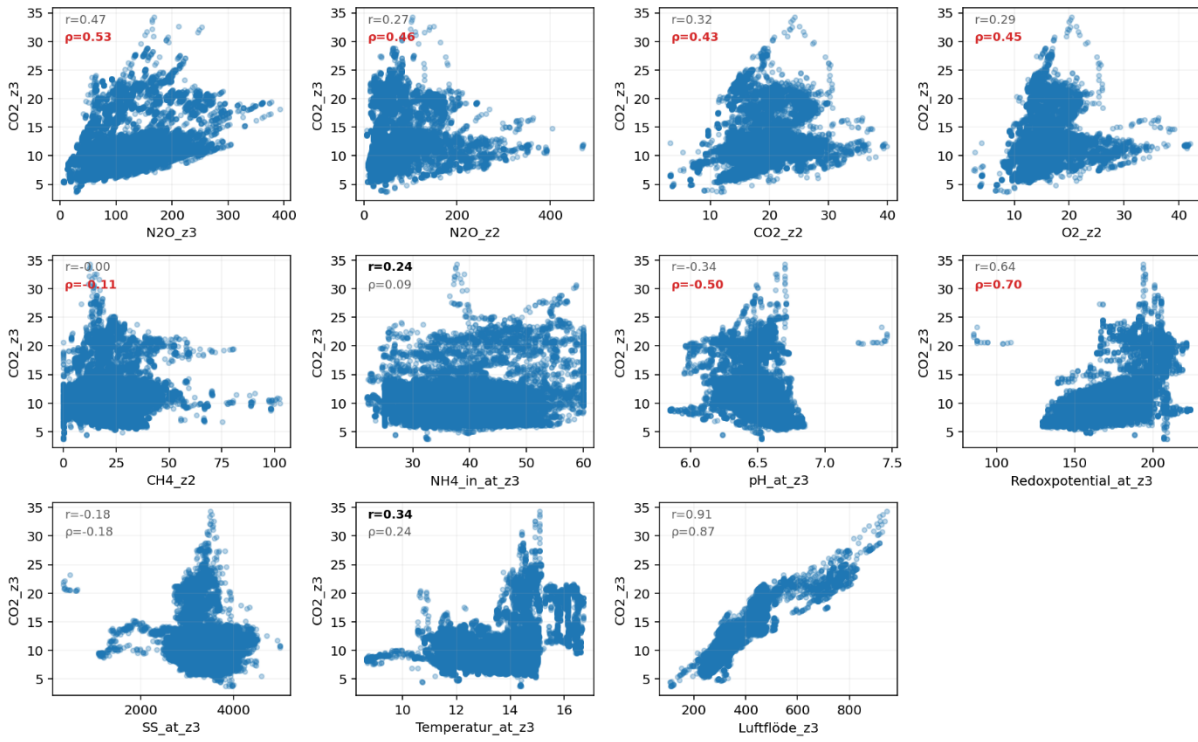
Zon 1 - Scatter mot CO₂_z1 (n=20000)



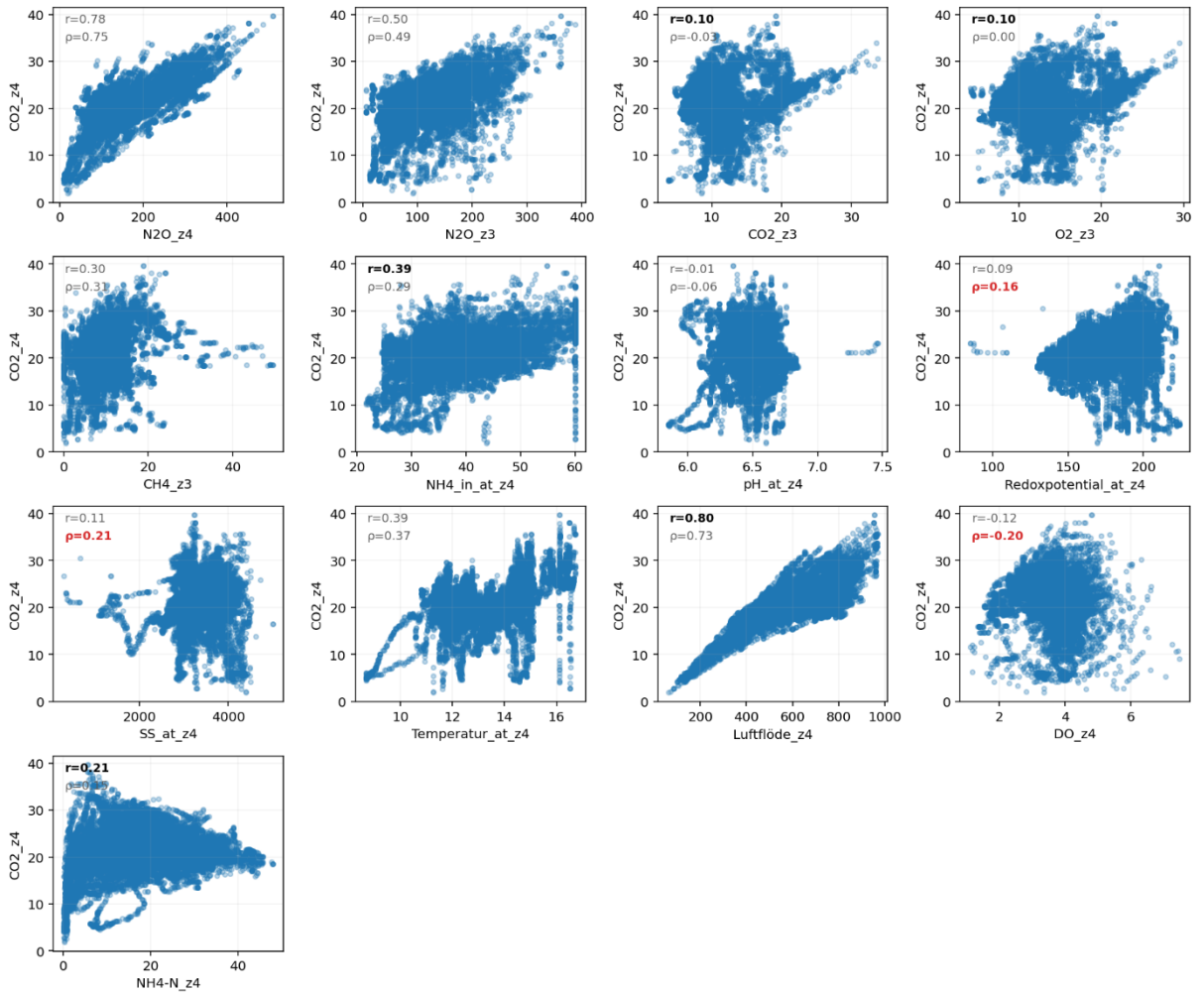
Zon 2 - Scatter mot CO₂_z2 (n=20000)



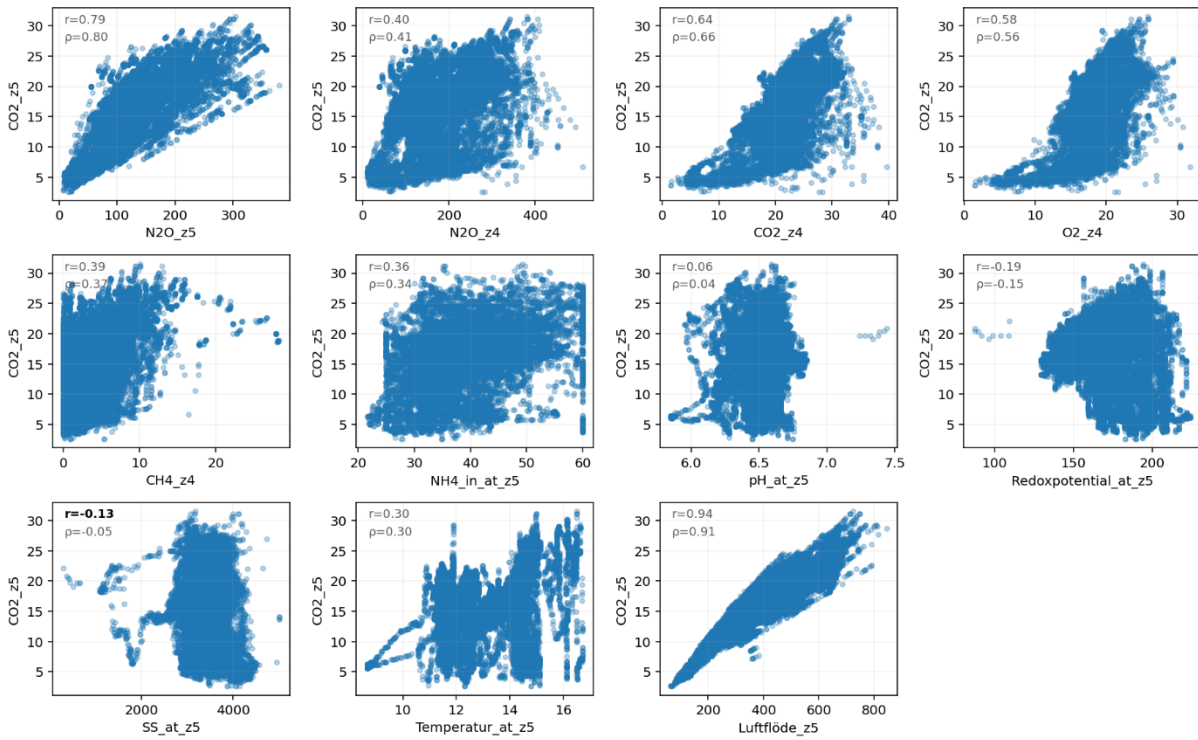
Zon 3 - Scatter mot CO2_z3 (n=20000)



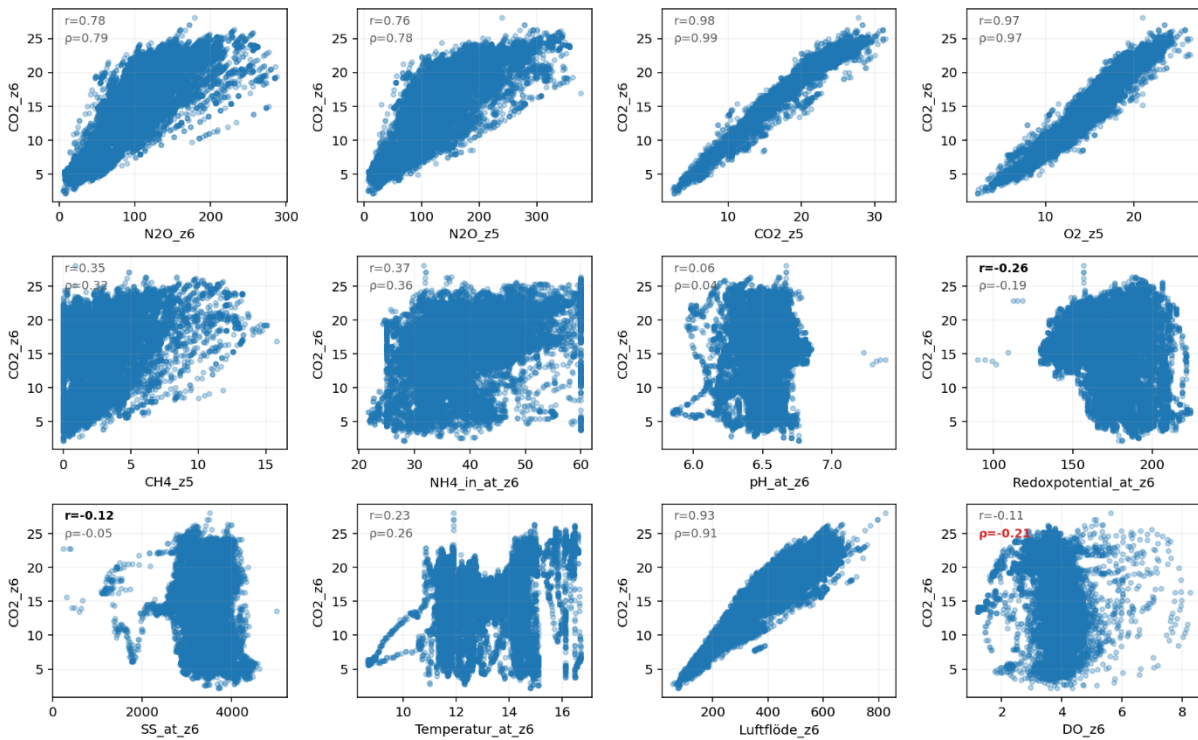
Zon 4 - Scatter mot CO2_z4 (n=20000)



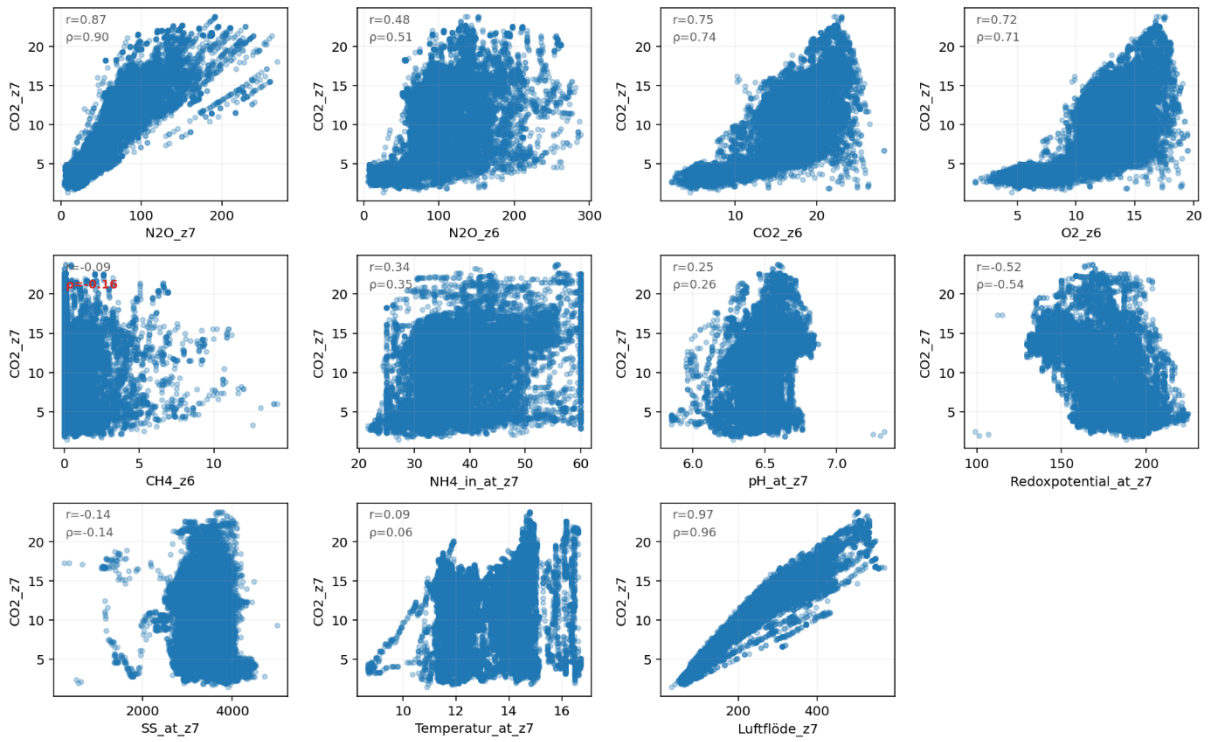
Zon 5 - Scatter mot CO2_z5 (n=20000)



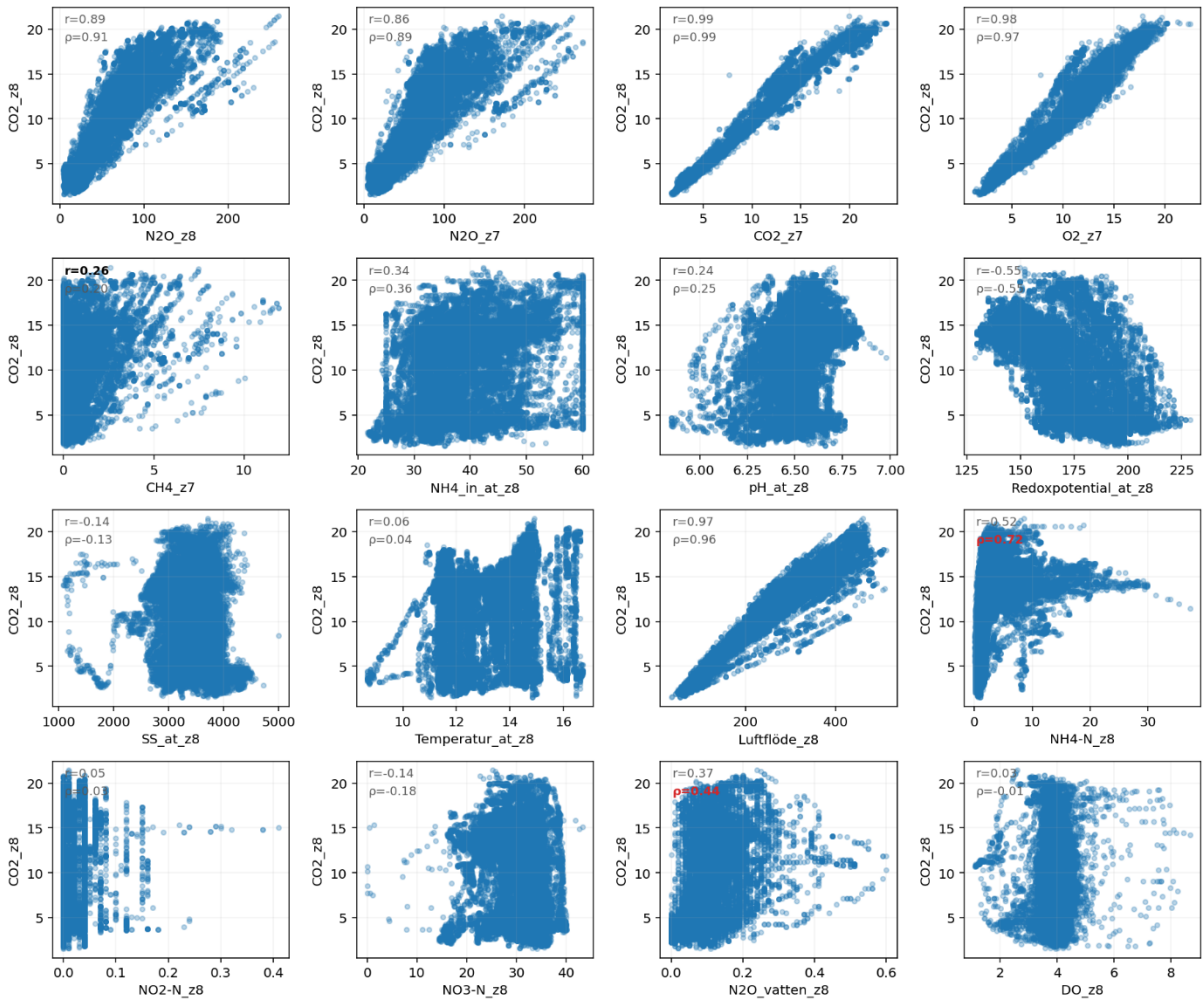
Zon 6 - Scatter mot CO2_z6 (n=20000)



Zon 7 - Scatter mot CO2_z7 (n=20000)

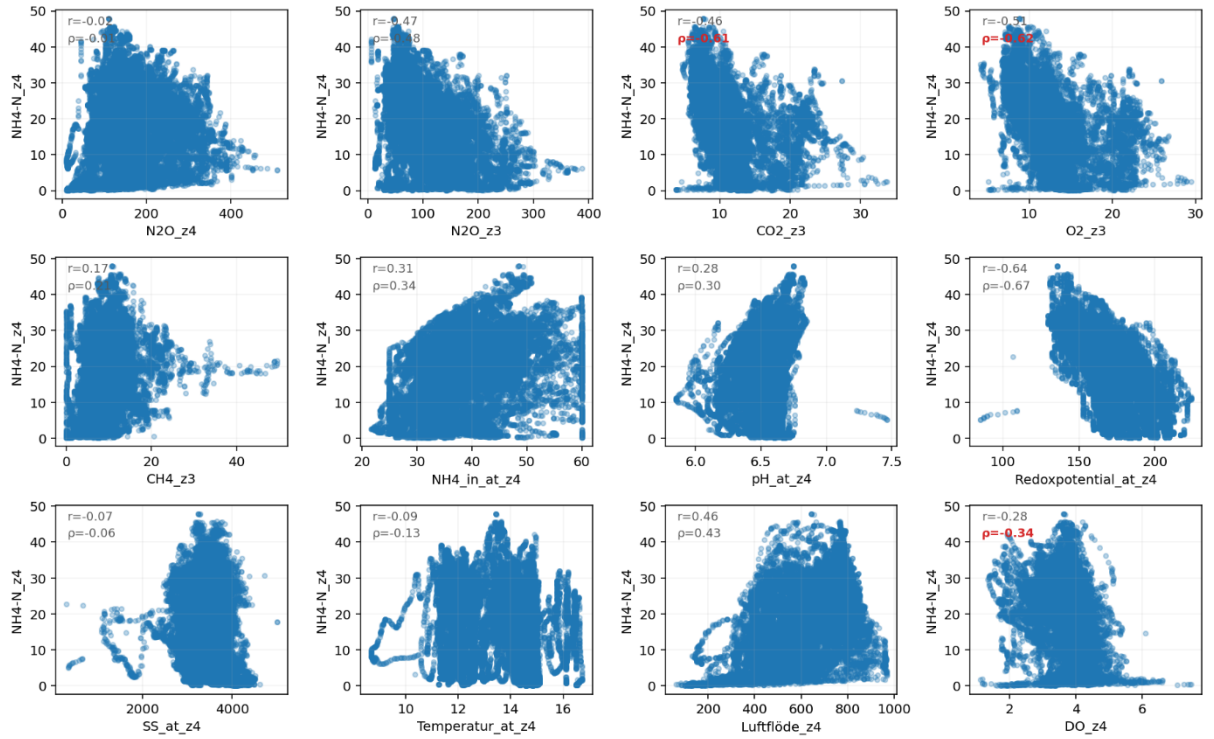


Zon 8 - Scatter mot CO2_z8 (n=20000)

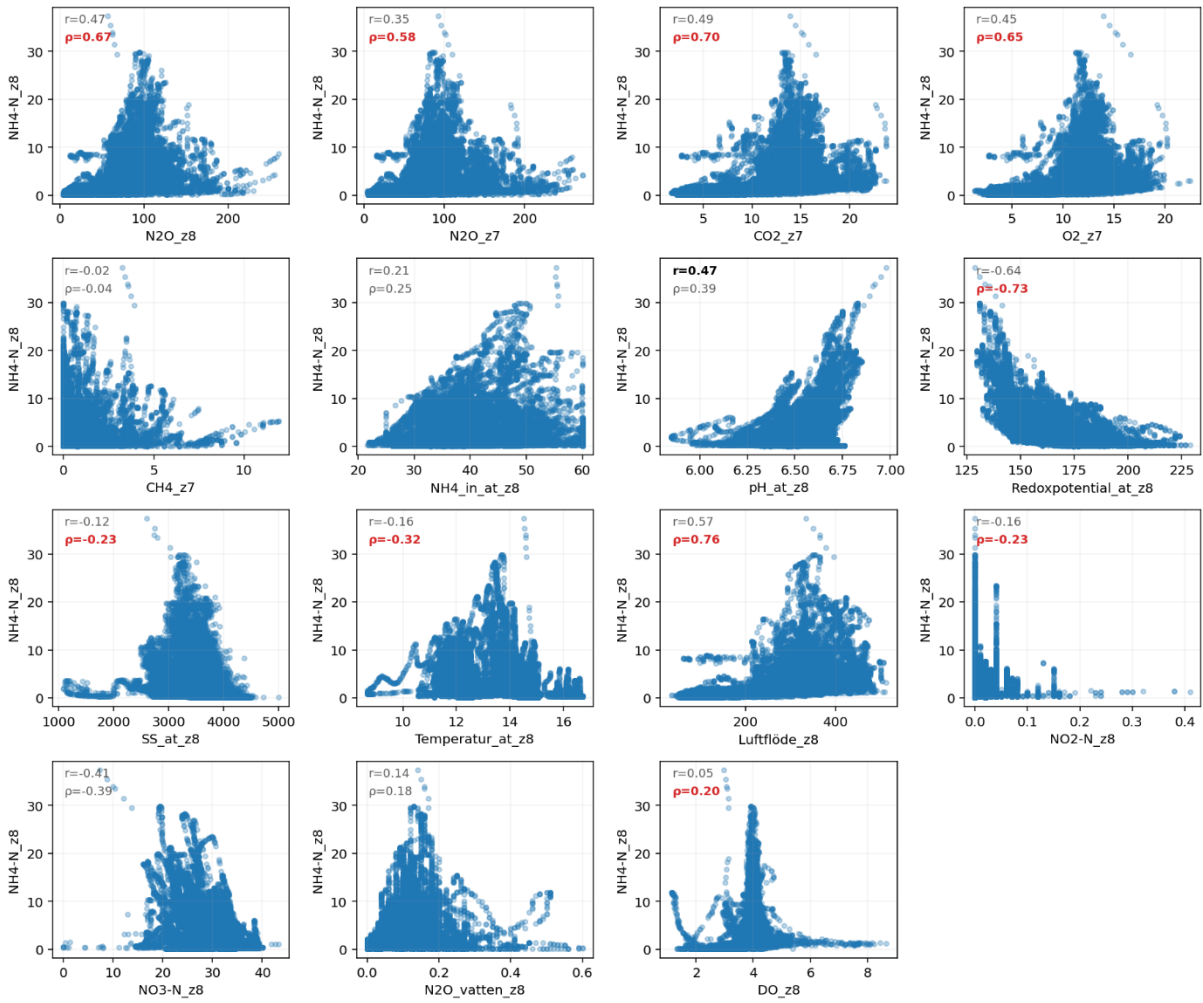


d) Korrelation mot NH₄-N

Zon 4 - Scatter mot NH₄-N_{z4} (n=20000)

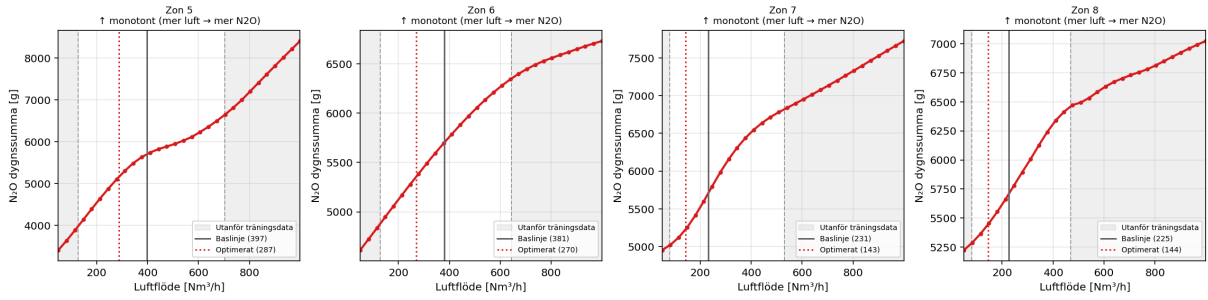


Zon 8 - Scatter mot NH4-N_z8 (n=20000)

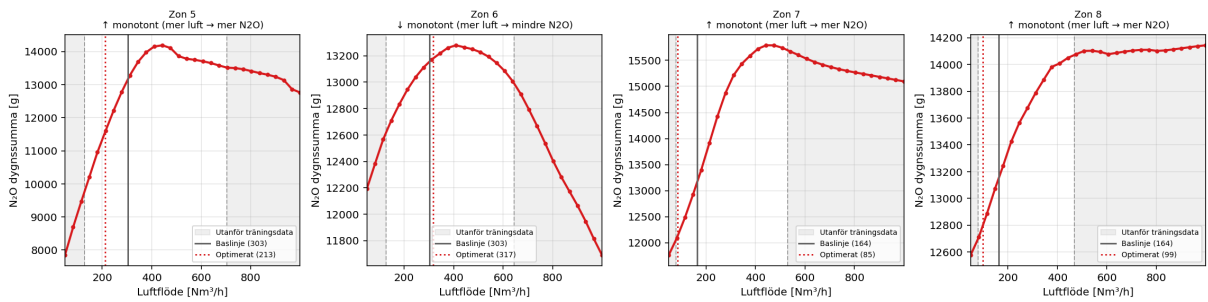


D. Känslighetsanalyser

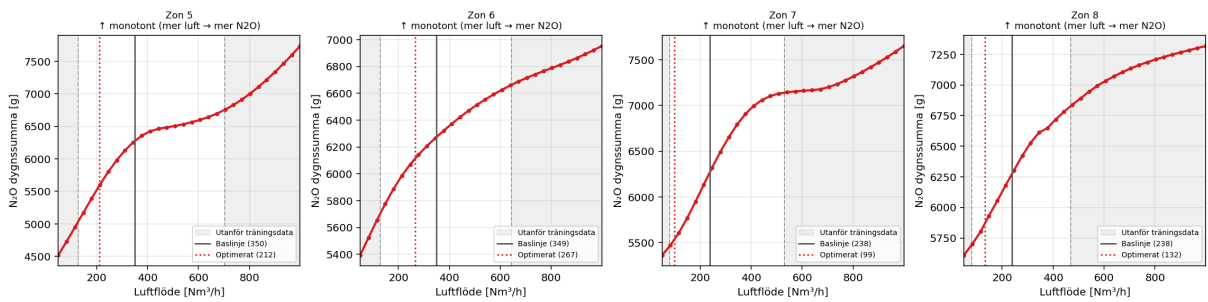
Sensitivitetstest: N2O vs luftflöde per zon (50-999 Nm³/h)
Grå = extrapolation utanför träningsdata - 2025-11-25



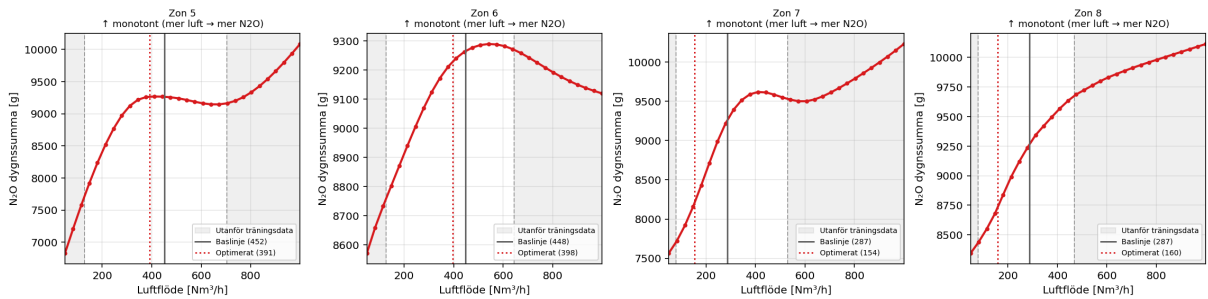
Sensitivitetstest: N2O vs luftflöde per zon (50-999 Nm³/h)
Grå = extrapolation utanför träningsdata - 2025-12-03



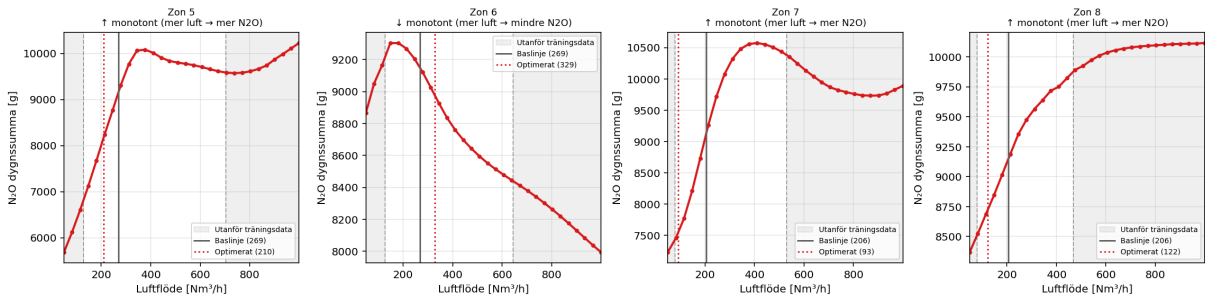
Sensitivitetstest: N2O vs luftflöde per zon (50-999 Nm³/h)
Grå = extrapolation utanför träningsdata - 2025-12-23



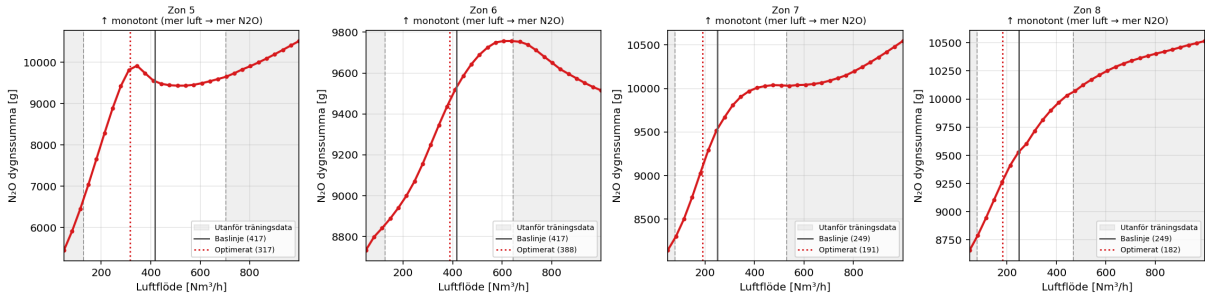
Sensitivitetstest: N2O vs luftflöde per zon (50-999 Nm³/h)
Grå = extrapolation utanför träningsdata - 2026-01-01



Sensitivitetstest: N2O vs luftflöde per zon (50-999 Nm³/h)
 Grå = extrapolation utanför träningsdata - 2026-01-09



Sensitivitetstest: N2O vs luftflöde per zon (50-999 Nm³/h)
 Grå = extrapolation utanför träningsdata - 2026-01-25



E. Jämförelse efter huvjustering – zon 2 och zon 5

