



KWR 2025.022 | Juli 2025

Grenswaarden Waterkwaliteit

Voorspellen van plantprestaties op basis
van online waterkwaliteitsmetingen

Samenwerkingspartners

TOPSECTOR
TUINBOUW & UITGANGSMATERIALEN

 **WATER & MARITIME**
WIS IN DE CHALLENGES. SMART SOLUTIONS

 **Kijk| Kennis in je Kas**

 **Glastuinbouw Nederland**
morgen groeit vandaag

 **CONTROL IN FOOD & FLOWERS**

KWR

 **Normec**

Plantum

sendot

 **Vivent**
Biosignals

 **LET'S GROW**
COM

AGRONA
BIOLOGICAL CROP PROTECTION & WATER TREATMENT

TOMAT WORLD


Colofon

Grenswaarden Waterkwaliteit

Voorspellen van plantprestaties op basis van online waterkwaliteitsmetingen

KWR 2025.022 | Juli 2025

Opdrachtnummer

404010

Projectmanager

Joep van den Broeke

Samenwerkingspartners

Normec Groen Agro Control, Plantum NL, Sendot, Letsgrow.com, Vivent Biosignals, Agrona, Stichting Kennis in je Kas, Glastuinbouw Nederland, Stichting Tomatoworld, STOWA, Stichting Control in Food & Flowers

Auteur(s)

Xin Tian, Mohamad Zeidan, Ammar Riyadh, Mollie Torello, Ruud Kaarsemaker, Joep van den Broeke

Kwaliteitsborger(s)

Peter van Thienen

Verzonden naar

Begeleidingscommissie TKI Grenswaarden Waterkwaliteit

Deze activiteit is mede gefinancierd met PPS-financiering uit de Toeslag voor Topconsortia voor Kennis en Innovatie (TKI's) van het Ministerie van Economische Zaken en Klimaat en de resultaten zijn openbaar. Dit project werd mede mogelijk gemaakt door een financiële bijdrage vanuit het Stimuleringsbudget Glastuinbouw. Kijk voor meer informatie op www.stowa.nl/glastuinbouw. Werkwijzen, rekenmodellen, technieken, ontwerpen van proefinstallaties, prototypen en door KWR gedane voorstellen en ideeën alsmede instrumenten, waaronder software, die in het onderzoeksresultaat zijn opgenomen, zijn en blijven het eigendom van KWR. Ook alle rechten die voortvloeien uit intellectueel en industrieel eigendom, alsmede de auteursrechten, blijven bij KWR berusten en derhalve eigendom van KWR.

Keywords

Greenhouse, Plant Growth, Water Quality, Data Science

Jaar van publicatie
2025

Meer informatie
Joep van den Broeke
T +31 (0)30 6069658
E joep.van.den.broeke@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

T +31 (0)30 60 69 511
E info@kwrwater.nl
I www.kwrwater.nl

KWR

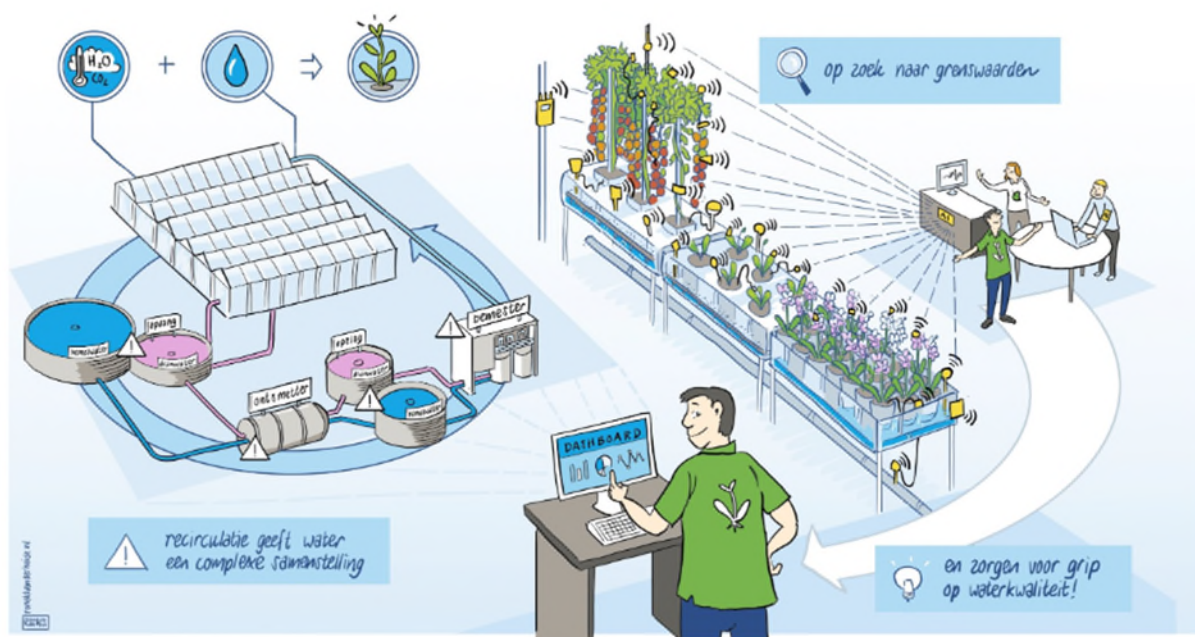
Juli 2025 ©

Alle rechten voorbehouden aan KWR. Niets uit deze uitgave mag - zonder voorafgaande schriftelijke toestemming van KWR - worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enig andere manier.

Samenvatting

Water is één van de belangrijkste productiefactoren in de glastuinbouw. Het is echter niet onbeperkt voorradig en ook andere sectoren hebben nood aan grote hoeveelheden water van goede kwaliteit. Daarnaast zijn er steeds strengere eisen aan het lozen van afvalwater. Dit heeft geleid tot een toename van de waterefficiëntie in de sector.; de bedrijfsvoering in de glastuinbouw is overgegaan een situatie waarin het drainwater geloosd (water in, water uit) naar een situatie waarbij het water zo lang mogelijk in de kas wordt gehouden door recirculatie. Dit brengt echter het risico met zich mee dat ongewenste stoffen, zoals natrium of ziekteverwekkers, zich ophopen. Zonder tijdige detectie kunnen deze ophopingen leiden tot groeivertragingen en verstopping van het watersysteem. Er is daarom behoefte aan een geautomatiseerd en betrouwbaar systeem dat real-time inzicht geeft in waterkwaliteit en plantgezondheid. Dit inzicht ontbreekt vooralsnog.

Het project “Grenswaarden Waterkwaliteit” beoogde het ontwikkelen van een waterkwaliteitsmodel op basis van plant-, water- en klimaatsensoren, op basis waarvan sturing op een gezonde teelt mogelijk wordt. Hiertoe werd in het project in praktijk- en proefkassen een combinatie van sensoren voor het meten van waterkwaliteit, kasklimaat en plantgezondheid geïnstalleerd. Door plant-, water- en klimaatsensoren te combineren, kan een ‘vingerafdruk’ van de waterkwaliteit worden bepaald en is het mogelijk de impact van waterkwaliteit op plantgezondheid te onderzoeken. Het figuur hieronder illustreert de aanpak in het project.



De doelstelling van het project Grenswaarden Waterkwaliteit schematisch weergegeven.

Om de relatie tussen waterkwaliteit en plantprestaties of plantgezondheid is gezocht naar parameters die met dezelfde frequentie als waterkwaliteitssensoren gemeten kunnen worden. Daarbij is gekozen voor het gebruik van Electron Transport Rate (ETR), een maat voor de efficiëntie van fotosynthese, en Plant Balance Index dat bio-elektrische signalen meet, welke een directe reactie stress in de plant weergeven. Beide parameters zijn met sensoren hoogfrequent te meten.

Met de geselecteerde sensoren zijn in een operationele omgeving meetwaarden verzameld bij Tomatoworld (tomatenteelt) en Levoplant (orchideënteelt). De op deze locaties verzamelde meetwaarden geven de reële variatie

in waterkwaliteit en klimaatomstandigheden weer. Echter, er was geen sprake van afwijkende waterkwaliteit en de waterkwaliteit kon niet actief negatief beïnvloed worden. Daarom is er ook in gecontroleerde omgeving, in klimaatkamers bij NORMEC, een aanvullende set experimenten uitgevoerd waarin de waterkwaliteit gevarieerd werd.

Met de verzamelde meetgegevens als input is vervolgens met DeepAR, een recurrent neural networks (RNN) gebaseerde benadering, onderzocht het mogelijk is om de gekozen plantgezondheidsparameters te voorspellen op basis van de water en klimaat parameters. Daarbij bleek de grote hoeveelheid verzamelde meetgegevens moeilijk te gebruiken voor een analyse met behulp van de gekozen data science benadering. Omdat er was gekozen voor de inzet van een groot aantal sensoren, werd het installeren, inregelen en onderhouden van de sensoren zeer complex. Dit was ook terug te zien in de dataset, die niet alleen complex, maar ook van onvoldoende kwaliteit en integriteit bleek te zijn om integraal te gebruiken in een data science aanpak. Op verschillende manieren is onderzocht of er relaties tussen de met behulp van sensoren gemeten waterkwaliteit en plantgezondheid vast te stellen waren. Zowel bij de metingen uitgevoerd onder praktijkomstandigheden, als bij de experimenten die in de klimaatkamer bij NORMEC zijn uitgevoerd, bleek het niet mogelijk om een directe correlatie tussen de met sensoren gemeten plantgezondheid en één individuele of een combinatie van waterkwaliteitsparameters vast te stellen.

Met behulp van machine learning is vervolgens onderzocht of er een tijdreeksmodel ontwikkeld kon worden met een voorspellende waarde voor plantgezondheid. Op basis van de beschikbare gegevens is het gelukt om een model met goede voorspellende waarde te creëren voor Plant Balance Index. Daarbij moet worden benoemd dat dit is gelukt door datasegmenten van voldoende goede kwaliteit te selecteren en dat het model is ontwikkeld op data die onder gecontroleerde en weinig variabele omstandigheden werd verzameld. Het ontwikkelde tijdreeksmodel dient gezien te worden als het bewijs dat met machine learning technieken relaties gevonden kunnen worden, en voorspellingen kunnen worden gedaan, die op basis van een klassieke data analyse niet gedaan kunnen worden. Het hier ontwikkelde model is hier een illustratie van. Echter, het model is nadrukkelijk geen vervanging voor het daadwerkelijk meten van de voorspelde grootte, PBI, en de voorspellende waarde onder operationele omstandigheden is niet getoetst.

Concluderend moet worden gesteld dat het op basis van de in Grenswaarden Waterkwaliteit verzamelde meetgegevens en het ontwikkelde model niet mogelijk is gebleken om grenswaarden te definiëren voor waterkwaliteit. Er werd *op basis van de verzamelde sensordata* geen causaal verband gevonden tussen waterkwaliteit, de impact van recirculatie en plantgezondheid.

Het werk beschreven in dit rapport laat zien dat het gebruik van een grote hoeveelheid aan sensoren niet per definitie ook een bruikbare dataset oplevert voor interpretatie en modelontwikkeling. Omdat iedere sensor betrouwbaar moet functioneren en hoogwaardige data dient te leveren, neemt de kans op problemen toe met een toenemend aantal sensoren. In Grenswaarden Waterkwaliteit is ingezet op het toepassen van een groot aantal relevant geachte sensoren. Met de ervaring opgedaan in het project met het installeren, inregelen en onderhouden van de sensoren, kan geconcludeerd worden dat het beter was geweest om in te zetten op een kleinere set aan sensoren waarvan de toegevoegde waarde vooraf goed beargumenteerd werd. Daarmee had de geleverde inspanning om het meetsysteem in te regelen efficiënter benut kunnen worden, en zou mogelijk een hoger datakwaliteit bereikt zijn, wat de interpretatie ten goede gekomen zou zijn.

Inhoud

Samenwerkingspartners	2
Colofon	3
Samenvatting	4
Inhoud	6
1 Inleiding	8
1.1 Aanleiding	8
1.2 Projectdoelstelling	9
1.3 Projectopbrengsten	9
1.4 Leeswijzer	10
2 Methoden en case studies	11
2.1 Literatuuroverzicht - hoe kunstmatige intelligentie helpt om de invloed van waterkwaliteit op plantengroei te begrijpen	11
2.1.1 Bepalen van plantgezondheid	11
2.1.2 Plantprestatie indicatoren	11
2.2 Meetlocaties en opzet experimenten	12
2.2.1 Opzet experimenten met waterbehandelingen	13
2.2.2 Parameter gemonitord voor modelontwikkeling	16
2.2.3 Berekening van ETR en correlatie met PAR	19
2.3 Tijdreeksmodellen	19
2.3.1 Architectuur van het Model	19
2.3.2 Model input en output	19
2.3.3 Model hyperparameter	20
2.3.4 Training, validatie en testen van het model	20
2.3.5 Prestaties van het ontwikkelde model (indicatoren)	20
2.3.6 Beoordeling van belang input parameters	20
2.4 Data voorbereiding voor het model	20
3 Resultaten en discussie	22
3.1 Case 1: Klimaatkamer bij NORMEC	22
3.1.1 Samenvatting datakwaliteit	22
3.1.2 Analyse correlatie tussen PAR en ETR binnen en tussen verschillende plantgroepen	22
3.1.3 Pulse-Amplitude Modulation gedurende de nacht (PAR = 0)	23
3.1.4 Resultaten behaald met de tijdreeks modellen	24
3.1.5 Belang van features	25

3.1.6	Dashboard	26
3.2	Case 2: Tomatoworld	27
3.2.1	Samenvatting datakwaliteit	27
3.2.2	Tijdreeksmodel	28
3.2.3	Optimalisatie van het tijdvenster	29
3.2.4	Belang van features	29
3.3	Case 3: LEVO	30
3.3.1	Kwaliteitsproblemen met de dataset	30
3.3.2	Tijdreeksen van gemeten parameters	30
4	Discussie en Conclusies	32
4.1	Algemene bevindingen	32
4.2	Online metingen en tijdreeksmodel	33
4.3	Aanbevelingen	34
5	Dankwoord	36
6	Referenties	37

1 Inleiding

1.1 Aanleiding

Water is één van de belangrijkste productiefactoren in de glastuinbouw. Het is echter niet onbeperkt voorradig en ook andere sectoren hebben nood aan grote hoeveelheden water van goede kwaliteit. Daarnaast zijn er steeds strengere eisen aan het lozen van afvalwater. Deze ontwikkelingen leiden tot een toenemende waterefficiëntie in de sector; de glastuinbouw maakt momenteel een overgang mee van een doorstroombedrijf (water in, water uit) naar een bedrijfsvoering waarbij het water zo lang mogelijk in de kas wordt gehouden. Zo wordt het gietwater dat door de planten niet wordt opgenomen, het drainwater, verzameld en hergebruikt in de teelt. Daartoe wordt het opgevangen, verzameld, gefilterd, ontsmet, opgeslagen, gemengd met vers water en gevoed met meststoffen. Water en meststoffen worden hergebruikt, wat een kostenbesparing op kan leveren.

Naast hergebruik van water zorgt recirculatie er echter ook voor dat stoffen die niet door planten worden opgenomen zich gaan ophopen in het water. Daarbij kunnen concentraties van stoffen die de plantgroei nadelig beïnvloeden, zoals natrium, oplopen tot een onwenselijk niveau. Daarnaast neemt door de langere verblijftijd van het water in het systeem ook de kans toe dat zich ziekteverwekkers ontwikkelen of dat biologische groei, en daardoor verstopping, optreedt in het watersysteem. Recirculatie is dus alleen duurzaam toe te passen als bewaakt wordt dat het water ook blijft voldoen aan de eisen van goed en gezond gietwater. Het is dus essentieel om bij recirculatie de kwaliteit van het gietwater goed te bewaken om problemen die de teelt of het watersysteem kunnen verstoren vroegtijdig te onderkennen en in de kunnen grijpen voordat dat ze zich tot een daadwerkelijk probleem ontwikkelen.

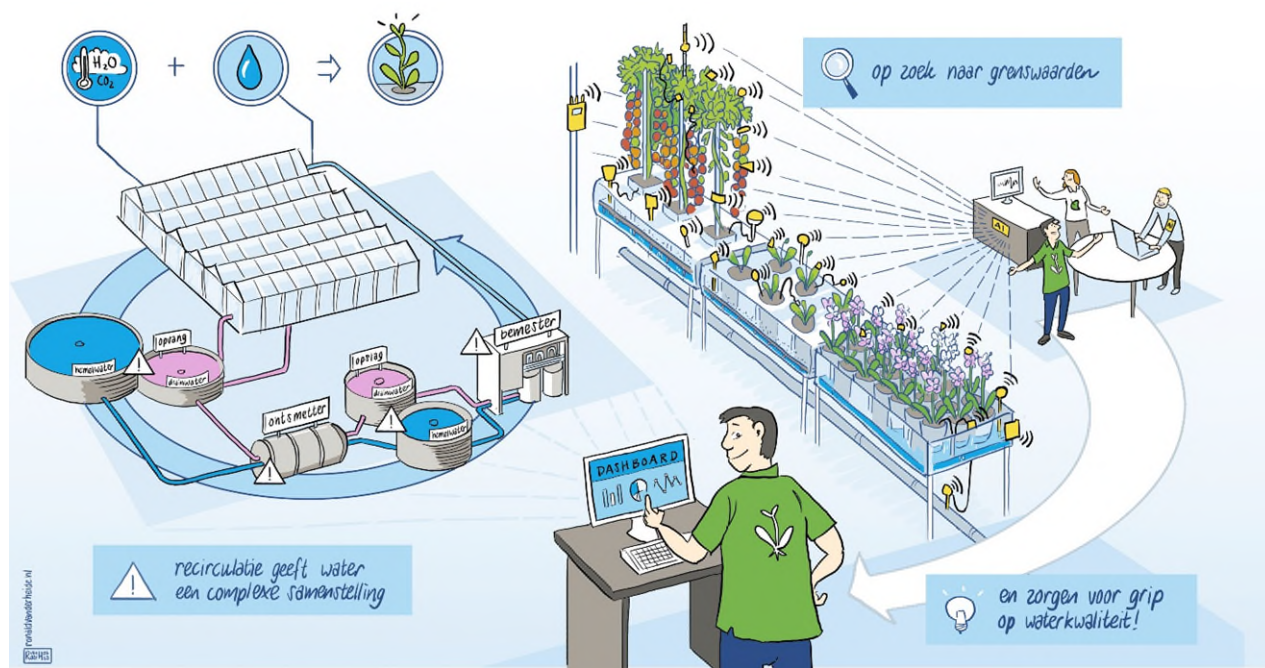
Het effect op de teelt van de aanwezige stoffen in combinatie met de verschillende microbiële processen in diverse compartimenten van het watersysteem is grotendeels onbekend. Waar voorheen de waterkwaliteit werd gezien als de aan-, of afwezigheid van nutriënten, leidt toenemend gebruik van recirculatie tot een verschuiving in het bewustzijn; niet alleen nutriënten maar de complexe samenstelling van het water is van belang. Hierbij gaat het om het samenspel tussen nutriënten, gewasbeschermings-, rem-, ontsmettingsmiddelen, organische stoffen, en diverse microbiologische parameters, waaronder de aanwezigheid van ziekteverwekkers en plagen. Voor het goed beheersen van het gietwatersysteem is het essentieel om te kunnen bepalen of de waterkwaliteit goed of slecht is met het oog op plantgezondheid en het operationeel houden van het watersysteem. Daarom is er de vraag vanuit de sector om meer en snel inzicht te krijgen in de waterkwaliteit en bijbehorende waterparameters met stuurmogelijkheden gericht op een gezonde en veilige teelt, vergelijkbaar met de reeds bestaande klimaatcomputer voor klimaat en belichting.

Voor een snelle en autonome meting van waterkwaliteit kan gebruikt gemaakt worden van waterkwaliteitssensoren. Er is een breed scala van zulke sensoren beschikbaar, en verschillende ervan zijn ook reeds ingezet of getest voor toepassing in de glastuinbouw (bijv. project TKI TU18084 – Waterkwaliteit Snel in Beeld). Dit laat zien dat sensoren weliswaar informatie over de watersamenstelling leveren, maar dat **geen enkele individuele sensor kan beoordelen of gietwater een goede kwaliteit heeft**. Dit komt omdat het totaal aan parameters bepaalt of de waterkwaliteit voldoende is om bedrijfszeker te worden hergebruikt, en veel van **de relevante parameters niet met sensoren te meten zijn en geen enkele sensor de verschillende parameters met elkaar kan verbinden**. Wel is bekend uit andere sectoren dat het mogelijk is om met combinaties van sensoren een ‘vingerafdruk’ van het water te meten en dat het mogelijk kan zijn om aan de hand van een dergelijke vingerafdruk uitspraken te doen over waterkwaliteit die op basis van de individuele sensoren niet mogelijk zouden zijn.

1.2 Projectdoelstelling

Het project Grenswaarden Waterkwaliteit had als doel om te onderzoeken of het mogelijk is om op basis van een combinatie van plant-, water- en klimaatsensoren te komen tot een beslissingsondersteunend meetsysteem waarin op basis van het monitoren van de waterkwaliteits-vingerafdruk aanwijzingen voor (bij)sturing van het recirculatieproces en waterbehandeling kunnen worden gegeven op basis van een set grenswaarden (Figuur 1). Daartoe was de doelstelling van het project om te komen tot een gebruiksvriendelijk beslissingsondersteunend systeem (dashboard) waarmee de eindgebruiker (tuinder) aan de hand van meetwaarden en grenswaarden inzicht heeft in de waterkwaliteit en ondersteuning krijgt bij de aansturing van recirculatie en waterbehandeling.

Daarnaast had het project ook als doelstelling om inzicht te verkrijgen in de toegevoegde waarde van sensoren voor waterkwaliteit, plantgroei en klimaat die individueel of gezamenlijk gebruikt kunnen worden voor het informeren van eindgebruikers over recirculatieprocessen en waterbehandeling. Als onderdeel daarvan zijn vooraf twee specifieke aandachtspunten vastgelegd in de projectdoelstellingen, te weten: 1) het ontwikkelen van een sensor ter vervanging van het kiemgetal als parameter voor microbiologische kwaliteit van gietwater en 2) het verkrijgen van de toegevoegde waarde van Next Generation Sequencing (NGS) bij het in kaart brengen van microbiologie, zowel gewenste als ongewenste organismen, in het watersysteem in de kas. Deze twee doelstellingen worden in de onderhavige rapportage niet besproken, zie sectie 1.3 voor een verwijzing naar de rapportages waarin deze resultaten te vinden zijn.



Figuur 1: De doelstelling van het project Grenswaarden Waterkwaliteit schematisch weergegeven.

1.3 Projectopbrengsten

Het project kende een looptijd van 01-01-2022 tot 31-03-2025. De bevindingen uit het project zijn samengevat in meerdere rapportages:

- 1) een overkoepelende samenvatting met daarin de belangrijkste bevindingen en conclusies (KWR 2025.025)

- 2) rapportages met betrekking tot de individuele meetcampagnes bij teeltbedrijven en klimaatkamerexperimenten, inclusief rapportage over de ontwikkelde sensor voor het meten van zuurstofverbruik (S20220001- 1, S20220001- 2, S20220001- 3, S20220001- 4, S20220001- 5).
- 3) een technische rapportage over de uitgevoerde experimenten en meetcampagnes en de daaraan gekoppelde modelontwikkeling (**deze rapportage, KWR 2025.022**)
- 4) een technische rapportage over de microbiologische analyses, in het bijzonder de Next Generation Sequencing activiteiten (KWR 2025.023)

1.4 Leeswijzer

Dit rapport presenteert een technische rapportage over de uitgevoerde experimenten en meetcampagnes en de daaraan gekoppelde modelontwikkeling. Na dit inleidende hoofdstuk 1 worden in hoofdstuk 2 de toegepaste methoden besproken. Daarbij wordt ingegaan op het meten van plantgezondheid op basis van met sensoren meetbare indicatoren (sectie 2.1), de experimentele opzet in het project (sectie 2.2), de methoden gebruikt om de tijdreeksmodellen te ontwikkelen (sectie 2.3) en de voorbereiding van de data gebruikt bij de modelontwikkeling (sectie 2.4). Hoofdstuk 3 bespreekt vervolgens de resultaten, waarbij achtereenvolgens aandacht wordt gegeven het werk in 3 casussen: klimaatkamerexperimenten bij NORMEC, meetcampagne bij Tomatoworld en meetcampagnes bij Levo Plant. Het rapport wordt afgesloten in hoofdstuk 4 met overkoepelende conclusies.

2 Methoden en case studies

2.1 Literatuuroverzicht - hoe kunstmatige intelligentie helpt om de invloed van waterkwaliteit op plantengroei te begrijpen

2.1.1 Bepalen van plantgezondheid

De afgelopen jaren heeft de integratie van kunstmatige intelligentie (AI) modellen in de landbouw veelbelovende resultaten opgeleverd bij het ondersteunen van telers in het beter begrijpen van waterkwaliteit en plantgezondheid (Oliveira en Silva, 2023). Het gebruik van AI biedt aanzienlijke mogelijkheden bij het voorspellen van plantprestaties op basis van waterkwaliteit, met name in systemen die afhankelijk zijn van gerecycled irrigatiewater. AI-modellen kunnen telers helpen om te identificeren wanneer de waterkwaliteit de plantprestaties kan beïnvloeden, zodat ze tijdig actie kunnen ondernemen voordat er aanzienlijke schade optreedt. Het doel van het in dit rapport beschreven onderzoek is om de relatie tussen de kwaliteit van gerecycled irrigatiewater en de fysiologische respons van planten in een kas te evalueren. Meer specifiek willen we bepalen of met sensoren meetbare variaties in de waterkwaliteit gekoppeld kunnen worden aan veranderingen in stress en groei van planten.

Plantengroei en plantprestaties worden traditioneel gemeten door het bepalen van natte en droge plantmassa. Een andere maatstaf is het totale gewicht en de kwaliteit van de geproduceerde vruchten. Dergelijke prestatie-indicatoren geven inzicht in de cumulatieve effecten van de groeiomstandigheden, maar daarbij is het niet mogelijk om een causaal verband vast te stellen tussen waterkwaliteit en plantengroei of -stress. Bovendien maakt het zeer lage aantal metingen dat met deze methoden kan worden verzameld, en het feit dat de informatie sterk gericht is op de omstandigheden van de plant aan het einde van zijn groeicyclus, dergelijke analysemethoden ongeschikt voor gebruik bij het bestuderen van de relatie tussen waterkwaliteit en plantprestaties.

Om data gestuurde methoden te kunnen toepassen om de invloed van gietwaterkwaliteit op plantprestaties te beoordelen en een model te ontwikkelen dat telers kan helpen bij het beheer van waterkwaliteit in recirculatiesystemen, is een andere maatstaf nodig voor de gezondheid en prestaties van planten. Hiervoor is namelijk een parameter voor plantprestaties nodig die met een vergelijkbare meetfrequentie als die gebruikt wordt voor het verzamelen van waterkwaliteitsgegevens. In dit project werden drie prestatiemetingen voor planten gebruikt die gemeten kunnen worden met real-time sensoren: Electron Transport Rate (ETR), Pulse Amplitude Modulation (PAM) en Plant Balance Index (PBI). Deze parameters worden kort geïntroduceerd in paragraaf 2.1.2.

Om de relatie tussen waterkwaliteit en plantengroei te onderzoeken, werden data gestuurde modellen toegepast. De gebruikte methoden worden geïntroduceerd in paragraaf 2.3.

2.1.2 Plantprestatie indicatoren

De gezondheid en prestaties van planten worden traditioneel gemonitord met behulp van meetgegevens zoals fotosynthese efficiëntie, die vaak wordt bepaald aan de hand van chlorofylfluorescentie of Electron Transport Rate (ETR) (Genty et al., 1989). Hoewel metingen van chlorofylfluorescentie soms een nuttige maat zijn voor de fotosynthetische prestaties van planten, ligt hun echte kracht in hun vermogen om informatie te geven die op andere manieren niet direct beschikbaar is. Fluorescentie kan met name inzicht geven in het vermogen van een plant om milieustress te verdragen en in de mate waarin die stress het fotosyntheseapparaat heeft beschadigd (Maxwell en Johnson, 2000). Deze meetgegevens bieden nuttige inzichten in de efficiëntie van fotosynthese onder verschillende omgevingsomstandigheden, maar zijn beperkt in hun vermogen om vroege tekenen van stress bij

planten te detecteren, omdat ze voornamelijk veranderingen in het fotosyntheseapparaat weerspiegelen (Baker, 2008).

Pulse Amplitude Modulation (PAM) fluorometrie wordt veel gebruikt om chlorofylfluorescentie te meten en fotosynthetische activiteit te beoordelen, met name de efficiëntie van Fotosysteem II (PSII). Overdag kunnen PAM-metingen inzicht geven in de ETR en het vermogen van de plant om lichtenergie om te zetten in chemische energie (Genty et al., 1989). PAM-waarden 's nachts bieden echter een ander perspectief en dienen als een indicatie voor het herstel van de plant en stressniveaus wanneer de fotosynthese inactief is (Schreiber, 2004). 's Nachts fotosynthetiseren planten niet actief, maar ondergaan ze essentiële fysiologische processen zoals ademhaling en herstel van stress overdag. Het meten van de nachtelijke PAM-waarden, met name de minimale fluorescentie en de maximale fluorescentie, geeft inzicht in de algehele hersteltoestand van de plant. Lage waarden van PAM-fluorescentie 's nachts duiden vaak op een plant onder stress of een plant die niet goed herstelt van de stress die overdag is ervaren. Omgekeerd suggereren stabiele of hoge nachtelijke PAM-waarden dat de plant efficiënt herstelt van de stress overdag en haar fysiologische balans behoudt (Schreiber, 2004, Glowacka et al., 2023). Een belangrijke parameter die wordt waargenomen tijdens nachtelijke PAM-metingen is de maximale quantumopbrengst van de fotochemie van PSII, die de efficiëntie van de energieomzetting binnen PSII weergeeft. Deze waarde blijft meestal relatief constant onder gezonde omstandigheden, rond 0,83 voor de meeste planten. Een afname van de maximale quantumopbrengst 's nachts kan duiden op de accumulatie van stress in de plant, mogelijk als gevolg van factoren zoals een suboptimale waterkwaliteit, onevenwichtige voedingsstoffen of extreme omgevingsomstandigheden (Murchie en Lawson, 2013). Het 's nachts monitoren van deze veranderingen zorgt voor een vroegtijdige waarschuwing van stress voordat het zichtbaar wordt in het uiterlijk van de plant of de groeiprestaties beïnvloedt. In de context van onderzoeken in kassen, zoals dit experiment, zijn nachtelijke PAM-waarden bijzonder nuttig om te evalueren hoe goed planten herstellen van de stress die wordt opgelegd door verschillende waterbehandelingen.

Voor een snellere detectie van stress is een meer geavanceerde indicator, de Plant Balance Index (PBI), gebruikt. Wanneer planten gestrest zijn of veranderingen in de omgeving waarnemen, ontstaan er kleine elektrische signalen door de beweging van geladen deeltjes in de vaten van de planten. De PBI detecteert deze bio-elektrische signalen en meet zo een directe reactie op de gezondheid van de plant en de dynamiek van stress. PBI opnames van het elektrofysiologische signaal van planten worden gemaakt met meerkanaals Vivent PhytSigns apparaten. Elk apparaat gebruikt elektrodeparen om de elektrofysiologische signalen in een plant te meten: een actieve elektrode in een hoger deel van de plantenstengel en een referentie-elektrode in het onderste deel van de stengel. De ingangsimpedantie van het apparaat is van de orde van vele MW en de meetfrequentie is 500 Hz. Per PhytSigns-apparaat worden acht planten op deze manier gemonitord. De verzamelde informatie wordt samengevoegd en geïnterpreteerd met behulp van op machine learning gebaseerde algoritmen van Vivent, en de PBI is de output van deze meet- en analyseprocedure. Een meer gedetailleerde beschrijving van de meting met behulp van PhytSigns is te vinden in Najdenovska et al., 2021a, Tran et al., 2019, en Najdenovska et al., 2021b.

Het is aangetoond dat PBI vroegtijdige tekenen van stress kan detecteren, vaak voordat de symptomen zich visueel manifesteren, waardoor het een waardevol hulpmiddel is voor preventief ingrijpen in het kasbeheer. In combinatie met real-time sensorsystemen kan deze geavanceerde meetmethode telers real-time inzicht geven in zowel de gezondheid van planten als de waterkwaliteit, zodat ze irrigatie- en voedingsstrategieën kunnen aanpassen als reactie op veranderende omstandigheden.

2.2 Meetlocaties en opzet experimenten

Om de effecten van de waterkwaliteit op de plantprestaties te onderzoeken, werden meetcampagnes uitgevoerd in op drie locaties in Nederland, namelijk in de kas bij Stichting Tomatoworld, in een kas bij Levoplant en in een klimaatkamer bij Normec.

2.2.1 Opzet experimenten met waterbehandelingen

2.2.1.1 Klimaatkamer Normec

Binnen zijn in de klimaatkamer bij Normec experimenten uitgevoerd met tomaten en orchideeën (Phalaenopsis). De experimentele opstelling is ontworpen om de effecten van verschillende waterbehandelingen op de groei en prestaties van deze planten te onderzoeken.

Specifiek is in de klimaatkamer gewerkt met zes plantkarren, gerangschikt in drie behandelingsgroepen (A, B en C) met twee karren per groep (Figuur 2). Deze configuratie maakt replicatie van behandelingen mogelijk en betrouwbare vergelijkingen tussen verschillende waterkwaliteitscondities en fotosynthetisch actieve straling (PAR). Paragraaf 2.2.2 geeft een overzicht van alle gemeten parameters in de klimaatkamer.

Op deze locatie is ook gebruik gemaakt van continue bewaking van de waterkwaliteit met behulp van de Sendot Water Module (zie ook figuur 3), waarvan er 1 is ingezet om ieder van de toegepaste waterbehandelingsexperimenten te bewaken:

- Behandeling A (referentie): bekkenwater met voedingsoplossing, geen beluchting.
- Behandeling B: 10% oud water, 45% gedesinfecteerd water, 45% bekkenwater, geen beluchting.
- Behandeling C: hetzelfde als behandeling B, maar met Agrona nanobubble beluchting.



Figuur 2: De klimaatkamer en layout van het experiment bij Normec (bron: Normec).

De mengsels voor behandeling A – C werden drie dagen voor eerste gebruik als gietwater bereid. Deze periode is gebaseerd op experimenten (Kaarsemaker et al, 2025) die een piek in microbieel zuurstofverbruik lieten zien 2-4 dagen na het mengen. Dit is uitgelegd als een snel verbruik van in het water aanwezige voedingsstoffen voor micro-organismen in het water. Na deze fase van groei is de watersamenstelling stabiel(er) en kan het enkele weken achtereen gebruikt worden als gietwater.

In het experiment is gebruik gemaakt van 150 tomatenplanten geënt op een Maxifort onderstam, verdeeld over de drie behandelingsgroepen. Er wordt een eb/vloed irrigatiesysteem gebruikt, waarbij water handmatig aan de bakken wordt toegevoegd en het teveel na vijf minuten wordt afgevoerd. De irrigatie wordt uitgevoerd wanneer het gewicht van het blok onder een bepaalde drempel komt, die wordt aangepast naarmate de planten groeien.

2.2.1.2 Tomatoworld

Tomatoworld (TW), gevestigd te Honselersdijk, dient als een wereldwijd platform voor het bevorderen van kennis en het tonen van innovatie in duurzame tomatenteelt. Het biedt een unieke omgeving voor onderzoek, onderwijs

en kennisuitwisseling om de efficiëntie en productiviteit van moderne tuinbouwpraktijken te verbeteren. De instelling integreert geavanceerde technologieën met wetenschappelijke methoden om uitdagingen in de sector aan te pakken, zoals het optimaliseren van het gebruik van hulpbronnen en het verbeteren van de kwaliteit van gewassen.

Bij de aanvang van het project is er voor gekozen om een zeer uitgebreide set aan parameters met sensoren te monitoren. Om ervaring op te doen met de operationele aspecten van installatie, data verzameling en overdracht en onderhoud, is er voor gekozen om in jaar 1 van het project het meetsysteem te installeren bij Tomatoworld. Dit had als bijkomend voordeel dat er ook enige ruimte was om te variëren met de irrigatie van de planten. Er werd gekozen voor de zeer uitgebreide set aan sensoren, omdat dit een zo compleet mogelijke dataset zou genereren, en omdat dit de mogelijkheid zou bieden om de toegevoegde waarde van verschillende, deels innovatieve, sensoren voor water- en plantparameters te kunnen bestuderen. Een totaaloverzicht van de geïnstalleerde sensoren is te vinden in Tabel 1. Voorbeelden van de installatie van de sensoren worden gegeven in Figuur 3.



Figuur 3: Voorbeelden van sensorinstallaties bij Tomatoworld in 2022. Van links naar rechts: Sendot mobiele waterkwaliteitsmeetopstelling, Gremon Trutina plantgewicht meting, Sendot Leafclip fotosynthese sensor.

Zoals uit tabel 1 valt op te maken, is er een groot en divers aantal sensoren geselecteerd. Voor het in dit rapport beschreven werk is een selectie gemaakt uit de beschikbare meetgegevens. Daarbij is ervoor gekozen om alleen te werken met de meetgegevens die in 2023 zijn verzameld bij Tomatoworld, omdat in 2022 de datakwaliteit van verschillende sensoren nog onvoldoende constant was (leerfase in het project). Daarnaast is er voor gekozen om slechts een selectie van parameters te beoordelen. Deze selectie is toegelicht in sectie 2.2.2. De gekozen selectie is tot stand gekomen in overleg met de experts van SCFF, waarbij de focus is gelegd op waterkwaliteitsparameters en de parameters die de grootste korte termijn impact op fotosynthese en PBI (gekozen maat van plantprestaties) hebben. De sensoren werden geïnstalleerd in 4 rijen planten (rijen 210-214) in het midden van de kas. In het jaar 2023 hebben de planten in de gemonitorde rijen geen speciale behandeling ontvangen, anders dan dat er vanaf augustus middels nanobubbles (Argona) de zuurstofconcentratie in het irrigatiewater werd verhoogd.

Tabel 1: Overzicht van in jaar 1 bij Tomatoworld geïnstalleerde sensoren.

Sensor (leverancier)	Gemeten grootheid	Aantal sensoren geplaatst
SnelWaterKwaliteit-unit (Sendot)	Waterkwaliteit: pH, EC, Redoxpotentieel, temperatuur, opgelost zuurstof, UV-fluorescentie, algen, troebelheid, zuurstof verbruik	1*
Fotosynthese sensor (Sendot)	Fotosynthese efficiëntie, fotosynthetisch actieve straling	3
PhytSigns (Vivent Biosignals)	Plant biosignalen	3
PhytoClip Leaf (2Grow)	Bladdikte	4
PhytoStem (2Grow)	Stengeldikte, sapstroom	12
Planttemperatuur	Temperatuur	2
Teros 12 (METER)	EC, watergehalte en temperatuur in substraat	3
Trutina (Gremont Systems)	Plant gewicht	2
CARA MET (SenseNL)	EC, watergehalte en temperatuur in substraat	6
Aquabalance (Hogendoorn Growth Management)	Gewicht substraat, drain volume	1
Pyranoladder	Fotosynthetisch actieve straling	1
Pyranometer (Kipp & Zonen)	Netto inkomende straling	1
Pyrgeometer	Netto uitgaande straling	1
PAR sensor	fotosynthetisch actieve straling	2
Zuurstofsensor (Sendot)	Zuurstofgehalte in substraat	3
Aspirator Box	luchttemperatuur, luchtvochtigheid	3
Gearsense (Gearbox Innovations)	beeldanalyse	1
Themoview 48	warmtebeeldcamera	1

*: in jaar 2 zijn 2 aanvullende meetopstellingen geleverd

2.2.1.3 Levoplant

Op locatie van Levoplant (LEVO) in Honselersdijk zijn metingen verricht in een kas gewijd aan de teelt van Phalaenopsis orchideeën. Als belangrijke speler in de sierteeltindustrie maakt LEVO gebruik van geavanceerde technologieën en nauwkeurige omgevingscontrole om de groei en bloei van deze populaire planten te optimaliseren (Figuur 4).



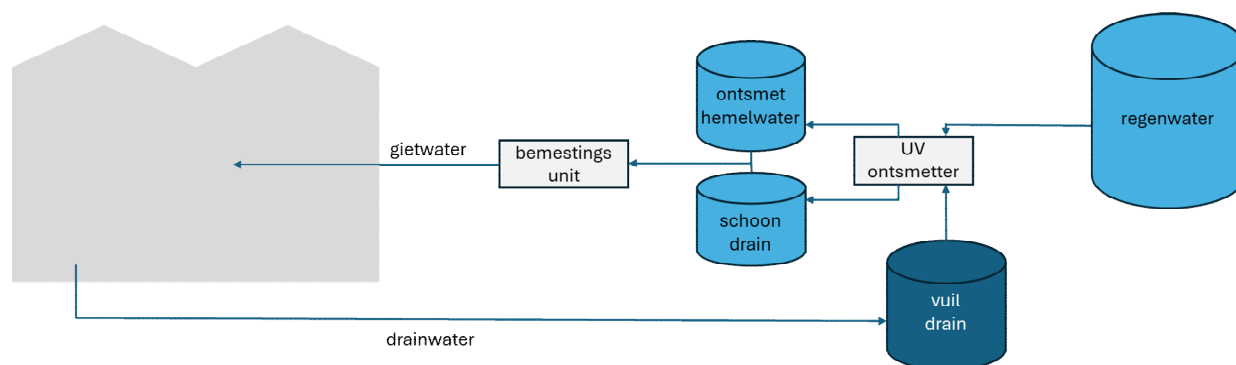
Figuur 4: Kas bij Levoplant. (foto: KWR)

Het Phalaenopsis kweekproces bij LEVO wordt gekenmerkt door een nauwgezet beheerde groeicyclus van vier fasen, die ongeveer 44 weken duurt vanaf het moment dat de planten op de locatie arriveren. Tijdens elke fase worden temperatuur, blootstelling aan licht en waterbeheer zorgvuldig gemanaged om een optimale ontwikkeling van de planten te garanderen:

- Initiële groeifase (29 weken bij 29°C)
- Afkoelingsfase (7 weken bij 19°C met verhoogde blootstelling aan licht)
- Ontwikkelingsfase van de bloemstengel (6 weken)
- Bloeifase (2 weken)

LEVO's faciliteit beschikt over ultramoderne automatiseringssystemen, waaronder klimaatbeheersing, het volgen van planten en sorteertechnologieën. Daarnaast is watermanagement een essentieel onderdeel van LEVO's activiteiten. LEVO maakt gebruik van een geavanceerd waterrecirculatiesysteem, waarbij opgevangen regenwater wordt gecombineerd met behandeld water dat opnieuw wordt gecirculeerd. Het aandeel gerecirculeerd water varieert per groeifase, met een conservatiever gebruik in de vroege, gevoelige fasen en een verhoogd gebruik in latere fasen.

LEVO ondervindt uitdagingen die te maken hebben met de waterkwaliteit en de invloed daarvan op de plantengroei, vooral in de vroege ontwikkelingsstadia. Deze problemen, die mogelijk verband houden met het recirculatieproces of de duur van de wateropslag, bieden mogelijkheden voor onderzoek en optimalisatie. Het watersysteem is hieronder kort weergegeven in figuur 5.



Figuur 5: Schematische weergave waterrecirculatiesysteem LEVO.

2.2.2 Parameter gemonitord voor modelontwikkeling

Om de prestaties van de planten te bestuderen, werd een reeks omgevings- en fysiologische parameters gevolgd tijdens de experimenten op de drie locaties. Factoren zoals zuurstofconcentratie, elektrische geleidbaarheid, pH en redoxpotentiaal helpen bij het beoordelen van de geschiktheid van het water en de groeiomgeving voor de ontwikkeling van planten. Daarnaast bieden parameters zoals kooldioxide, temperatuur, PAR en ETR een beter begrip van de fotosynthetische prestaties en energieomzettingsefficiëntie van planten. Geavanceerde meetgegevens zoals UV-fluorescentie, algenconcentratie en PBI geven de aanwezigheid van organisch materiaal, biologische activiteit en stressreacties van planten aan. Tabel 2 geeft een overzicht van de gemeten parameters die ook zijn meegewogen in de model ontwikkeling.

Tabel 2: Overzicht van gemeten parameters meegewogen in de modelontwikkeling beschreven in deze rapportage.

Parameter	eenheid	afkorting	Rol in relatie tot plantgezondheid en -prestaties
Zuurstof concentratie (in substraat)	%	O ₂	Beïnvloedt de beschikbaarheid van zuurstof voor het wortelsysteem en opname van voedingsstoffen
Zuurstof concentratie (in oplossing)	%	DO	Beïnvloedt de beschikbaarheid van zuurstof voor het wortelsysteem
Electrische geleidbaarheid	μS/cm	EC	Geeft het totaal aan opgeloste ionen aan (voedingsstoffen + zouten)
pH	-	pH	Bepaalt de beschikbaarheid van voedingsstoffen (bijv., Oplosbaarheid van ijzer)
Koolstofdioxide	ppm	CO ₂	Essentieel voor fotosynthese
Redox Potentiaal	mV	ORP	Geeft de oxidatiereductietoestand aan (aërobe/anaërobe omstandigheden)
Algen concentratie	mg/L	Alg	Indicator van biologische activiteit; concurreert om voedingsstoffen
UV Fluorescentie	RFU	UVF	Meet organisch materiaal (tryptofaan/humuszuren)
Temperatuur	°C	Tem	Beïnvloedt de oplosbaarheid van zuurstof en microbiële activiteit
Bodemvocht	%		Beïnvloedt de beschikbaarheid van zuurstof voor het wortelsysteem
Gewicht	kg		Geeft het watergehalte in het substraat aan en kan worden gebruikt om de biomassatoename te berekenen
Electron Transport Rate	μmol/m ² /s	ETR	Als maat voor fotosynthetische prestaties berekend met PAR om inzicht te geven in de efficiëntie van de plant bij het omzetten van licht in chemische energie.
Photosynthetically Active Radiation	μmol/m ² /s	PAR	Licht energie beschikbaar voor fotosynthese en heeft een bereik van 400-700 nm.
Plant Balance Index	-	PBI	Real-time indicator van de gezondheid van planten en stressreacties
Pulse Amplitude Modulation	%	PAM	Maat voor de effectieve quantum yield (Y) van de fotosynthese. ¹

¹ Het PAM-fluorometriepincipe is gebaseerd op een lichtpuls van 1 μs (lage intensiteit, niet-actinisch) die wordt gesynchroniseerd met een lock-in versterker. Hierdoor kunnen effectieve kwantumrendementsbepalingen worden uitgevoerd in (zon)licht, aangezien de lock-in versterker alle signalen verwijdert die niet geassocieerd worden met het lock-in signaal. Na aanpassing aan het donker kan de minimale fluorescentiewaarde (Fo) van het monster worden gemeten. Als het monster vervolgens wordt blootgesteld aan een verzadigende lichtpuls, zodat alle reactiecentra de energie van een foton dat door hun antennes is geabsorbeerd hebben gevangen (reactiecentra zijn gesloten), kan de maximale hoeveelheid fluorescentie (Fm) worden gemeten. Het verschil tussen deze twee extreme waarden is de variabele fluorescentie (Fv). Fv/Fm geeft een maat voor de maximale fotochemische efficiëntie van PSII. Bij actinische belichting van het monster kan de fluorescentieopbrengst variëren tussen de twee extreme waarden Fo en Fm. De kortstondige fluorescentieopbrengst, F, kan kort voor de verzadigende puls worden gemeten, waarbij de maximale fluorescentieopbrengst, Fm', wordt bereikt. De effectieve opbrengst Y(II) van de fotochemische energieomzetting van PSII onder verlichting wordt berekend als (Fm'-F)/Fm'. Bron:

<https://www.wur.nl/en/show/pulse-amplitude-modulation.htm>

Tabel 3 toont deze gemeten parameters op elke locatie. Merk op dat er verschillen zijn in de specifieke parameters die op elke locatie werden gemeten. Zo werd de algenconcentratie (Alg) in alle kassen gemeten, maar was de monitoring van kooldioxide (CO₂) uniek voor LEVO. De periodes en kwaliteit van de gegevensverzameling varieerde ook. Bij TW is over een lange periode over een vrijwel volledige dataset met een minimaal aantal ontbrekende meetwaarden gerealiseerd, terwijl bij LEVO voor verschillende parameters grotere hoeveelheden data ontbraken, met name voor Alg, ORP en PBI. Daarnaast vertoonde deze dataset veel irregulariteiten (veranderingen meetinterval, inconsistente timestamps, niet gesynchroniseerde meettijden) wat de analyse sterk bemoeilijkt. De dataset verzameld in de klimaatkamer van Normec bevatte ook hiaten, met name voor O₂ en UVF, en vertoonde lichte variaties in meetintervallen. Details over de gegevenskwaliteit van de locaties worden gerapporteerd in hoofdstuk 3.

Tabel 3: Overzicht van de gemeten parameters per locatie.

	Waterkwaliteit	Omgevingsparameters	Plant prestatie indicator	Opmerkingen (frequentie, data kwaliteit, etc.)
Tomatoworld	Alg O ₂ UVF Tem EC pH ORP	PAR, bodemvocht, blokgewicht	PBI, ETR	Meetfrequentie: iedere 5 minutes tussen 17 Feb 2023 en 21 Okt 2023 (uitgezonderd PBI alleen beschikbaar tussen 17 Feb 2023 en 20 Jul 2023). Weinig ontbrekende meetwaarden.
LEVO	Alg ORP O ₂ CO ₂ EC pH UVF	PAR, gewicht	PBI, ETR	Meetfrequentie: iedere 5 minutes tussen 1 Nov 2022 en 1 Jul 2023). Ontbrekende data in het bijzonder voor Alg, ORP, en PBI.
Klimaatkamer NORMEC	Alg ORP O ₂ EC pH UVF Tem	PAR, pot gewicht	PBI, ETR	Meetfrequentie iedere 5 minutes tussen 4 Apr 2024 en 22 Jun 2024, met variaties in meetinterval (bijv. Die van PAR en PAM waren niet consistent tussen de plant groepen A, B en C, wat resulteerde in problemen met het ontwikkelen van tijdreeksmodellen). Er zijn blokken data van meerdere weken die ontbreken voor O ₂ en UVF.

Er is in het project ook gewerkt met een nieuw ontwikkelde opstelling om het zuurstofverbruik in een watermonster te bepalen. De resultaten van deze meting zijn echter afgekeurd voor de tijdreeksmodellering vanwege onvoldoende gegevenskwaliteit. Details over deze opstelling zijn te vinden in het rapport KWR 2025.025.

2.2.3 Berekening van ETR en correlatie met PAR

Er werd eerst een correlatieanalyse uitgevoerd tussen fotosynthetisch actieve straling (PAR) en elektronentransportsnelheid (ETR) om de relatie tussen lichtbeschikbaarheid en fotosynthetische efficiëntie in de planten onder verschillende waterkwaliteitscondities te bepalen. Daarbij wordt opgemerkt dat ETR, afgeleid van PAR- en chlorofylfluorescentiemetingen (PAM), een directe meting is van het vermogen van de plant om lichtenergie om te zetten in chemische energie. Door de correlatie tussen deze twee parameters te kwantificeren, wilden we bepalen in welke mate en in welke verhouding PAR de fotosynthetische activiteit beïnvloedt en of waterbehandeling deze relatie beïnvloedt. De ETR werd berekend uit de PAR-waarden met vergelijking (1), volgens Genty et al. (1989):

$$ETR = Y(II) \times PAR \times 0.84 \times 0.5 \quad (1)$$

waarbij PAR de fotosynthetisch actieve straling is ($\mu\text{mol}/\text{m}^2/\text{s}$), 0,84 de gemiddelde verhouding van het licht dat door planten wordt geabsorbeerd, 0,5 de gemiddelde verhouding van Fotosysteem II en Y(II) de effectieve opbrengst van de fotochemische energieomzetting van PSII, zoals beschreven in voetnoot 1.

Om de relatie tussen PAR en ETR onder verschillende waterkwaliteitscondities verder te analyseren, werden zowel analyses binnen als tussen de groepen planten in het klimaatkamerexperiment uitgevoerd. De analyse binnen de groep (type waterbehandeling) beoordeelde hoe de correlatie tussen PAR en ETR varieert binnen elke groep (A, B en C), wat inzicht verschafte in mogelijke variaties binnen de groep in fotosynthetische efficiëntie. Ondertussen vergeleek de analyse tussen de groepen (waterbehandelingen) de correlatietrends over de drie groepen om te evalueren of waterbehandeling deze relatie verschillend beïnvloedt in verschillende omstandigheden. Door deze analyses uit te voeren, wilden we de consistentie van PAR-ETR interacties kwantificeren en bepalen in welke mate omgevingsfactoren de fotosynthetische respons in elke groep vormgeven.

2.3 Tijdreeksmodellen

In dit onderzoek werden tijdreeksmodellen gebruikt om de prestaties van planten te voorspellen op basis van waterkwaliteitsparameters. Het primaire doel was het voorspellen van de Plant Balance Index (PBI), een fysiologische indicator van de gezondheid van planten, met behulp van historische sensorgegevens. Er is gekozen voor het modelleren van PBI, en niet ETR, om de volgende redenen:

- De kwaliteit van de verzamelde meetgegevens van PBI is beter dan die van ETR;
- PBI is een van de expliciete indicatoren die kortstondige stress in planten laat zien, er kan daarom verwacht worden dat PBI reageert op snelle of kortstondige variaties in waterkwaliteit;
- ETR vertoont een zeer sterke correlatie met PAR. Eerste analyse lieten zien dat PAR verreweg de dominante inputparameter was bij het bepalen van ETR.

2.3.1 Architectuur van het Model

Het DeepAR-model (Salinas et al, 2020), een op recurrent neural networks (RNN's) gebaseerde benadering met Long Short-Term Memory (LSTM)-lagen (Sherstinsky, 2020). LSTM-lagen zijn een gespecialiseerde variant van RNN's die zijn ontworpen om langetermijnafhankelijkheden in sequentiële gegevens te verwerken. Ze worden gebruikt om temporele afhankelijkheden vast te leggen en probabilistische voorspellingen te genereren.

De architectuur van het model bestaat uit meerdere gestapelde LSTM-lagen gevolgd door een volledig verbonden uitvoerlaag die probabilistische voorspellingen genereert. Het model voorspelt zowel puntschattingen als betrouwbaarheidsintervallen.

2.3.2 Model input en output

Rekening houdend met de gegevenskwaliteit (dichtheid en ontbrekende hiaten) en relevantie is er voor gekozen om als **input** voor het model gebruikt te maken van de belangrijke waterkwaliteitsparameters zoals pH, opgeloste

zuurstof, algenconcentratie, temperatuur, UV-fluorescentie, elektrische geleidbaarheid (EC) alsmede fotosynthetisch actieve straling (PAR), terwijl de *uitvoer* de PBI is..

2.3.3 Model hyperparameter

De prestaties van het model werden verfijnd door het aantal LSTM-lagen, verborgen eenheden per laag, de batchgrootte en de leersnelheid aan te passen. In deze studie werd het model geconfigureerd met twee LSTM-lagen die elk 50 verborgen eenheden bevatten. De Adam optimizer (Arunachalam et al, 2022) werd gebruikt met een leersnelheid van 0,001 om een optimale convergentie tijdens de training te garanderen.

2.3.4 Training, validatie en testen van het model

Om robuuste prestaties te garanderen, is het model ontwikkeld via drie stappen, nl. een trainings-, validatie- en testproces. De dataset werd verdeeld in een training set bestaande uit van 80% van de geselecteerde meetwaarden, een validatie set bestaande uit 10% van de geselecteerde meetwaarden voor het afstemmen van de hyperparameters en een testset bestaande uit de resterende 10% van de meetwaarden voor de out-of-bag evaluatie. De 'sliding window'-benadering werd gebruikt om waarnemingen uit voorbije N-stappen in het model in te voeren, waardoor het effectief temporele patronen kan leren. Verschillende vensterformaten van N werden onderzocht om de voorspellende nauwkeurigheid te optimaliseren.

2.3.5 Prestaties van het ontwikkelde model (indicatoren)

De prestaties van het model worden geëvalueerd met behulp van een aantal maatstaven, waaronder de correlatiecoëfficiënt (R^2) en de gemiddelde absolute fout (mean absolute error - MAE). De correlatiecoëfficiënt (R^2) wordt gebruikt om te meten hoe goed het model de variantie in de waargenomen PBI verklaart, waarbij waarden dicht bij 1 duiden op een sterk voorspellende relatie tussen waterkwaliteitsparameters en de gezondheid van planten. Daarnaast wordt de MAE berekend om de gemiddelde afwijking tussen voorspelde en werkelijke PBI-waarden te kwantificeren. Een hogere R^2 en een lagere MAE wijzen samen op een model met sterke voorspellende prestaties, dat de complexe temporele afhankelijkheden tussen inputkenmerken en de plantenprestaties effectief vastlegt.

2.3.6 Beoordeling van belang input parameters

Om de voorspellingen van het model te interpreteren en het relatieve belang van verschillende waterkwaliteitsparameters te begrijpen, wordt SHapley Additive exPlanations (SHAP) toegepast (Lundberg en Lee, 2017). Dit is een op speltheorie gebaseerde methode die wordt gebruikt om de uitvoer van machine-learning modellen te verklaren door de bijdrage van elk kenmerk aan individuele voorspellingen te kwantificeren. SHAP biedt zowel lokale verklaringen (voor afzonderlijke voorspellingen) als globale inzichten (over hele datasets) terwijl het model agnostisch blijft. Daarom kwantificeren SHAP-waarden, in ons geval, de bijdrage van elk inputkenmerk aan de modeluitvoer.

2.4 Data voorbereiding voor het model

Het datavoorbereidingsproces voor het model omvat meerdere stappen, waaronder het opschonen van gegevens, voorbewerking en transformatie om ervoor te zorgen dat de invoer geschikt is voor deep learning. Eerst wordt de dataset geladen en wordt de kolom 'DateTime' geconverteerd naar een datetime-indeling voor een goede temporele afstemming. Belangrijke numerieke kolommen met betrekking tot de waterkwaliteit ('Algen', 'Zuurstof', 'UV-fluorescentie', enz.) worden gecontroleerd op nulwaarden, die worden vervangen door NaN om interpolatie te vergemakkelijken. Ontbrekende waarden worden dan lineair geïnterpoleerd, gevolgd door voorwaartse en achterwaartse opvulling om volledigheid te garanderen. Vervolgens worden categorische kenmerken zoals 'Plant Type' one-hot gecodeerd om ze om te zetten in numerieke vorm, zodat ze geschikt zijn voor invoer in het model. De numerieke kenmerken worden vervolgens gestandaardiseerd tussen 0 en 1 om consistente schaling van kenmerken te garanderen, terwijl de kolom 'DatumTijd' wordt behouden zonder schaling. De verwerkte gegevens

worden gestructureerd in tijdsequenties, waarbij een rolling window-benadering wordt toegepast om input-outputsequenties voor het model te genereren. De dataset wordt opgesplitst in trainings-, validatie- en testsets en geconverteerd naar PyTorch tensors voor efficiënte batchverwerking.

3 Resultaten en discussie

3.1 Case 1: Klimaatkamer bij NORMEC

3.1.1 Samenvatting datakwaliteit

Over het algemeen is de dataset van hoge kwaliteit met een consistente frequentie en meestal volledige gegevens (zie tabel 4), afgezien van enkele ontbrekende waarden in UV-fluorescentie en variabiliteit in sommige parameters (bijv. zuurstof).

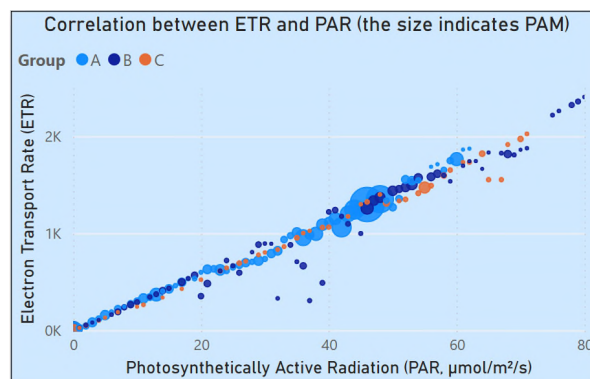
Tabel 4: Samenvatting van de data verzameld in de klimaatkamer bij NORMEC.

	Count of data points	Mean	Standard Deviation	Min	Max	Missing Value Ratio (%)
EC	69120	3.1	1.1	0.0	4.9	0.0
pH	69120	6.0	0.6	3.4	7.3	0.0
Algen	69120	0.3	0.8	-0.2	5.0	0.0
Opgelost Zuurstof	68810	24.3	13.1	2.7	65.2	0.4
UV fluorescentie	49041	-83.2	49.0	-163.0	160.2	29.0
Temperatuur	69120	19.1	0.9	17.5	21.7	0.0
Redox potentiaal	69120	381.6	36.4	185.5	455.4	0.0
PBI	65542	0.7	0.1	0.0	1.0	5.2
PAR	14809	37.4	16.8	0.0	80.0	0.0
PAM	14809	66.5	8.8	0.0	82.8	0.0
ETR	14809	1032.2	488.6	0.0	2419.2	0.0

3.1.2 Analyse correlatie tussen PAR en ETR binnen en tussen verschillende plantgroepen

Figuur 6 toont de correlatie tussen ETR en PAR binnen de drie meetgroepen (A, B en C), waarbij de puntgrootte de PAM-intensiteit weergeeft. De figuur laat een sterke positieve correlatie zien. Met andere woorden, ETR neemt toe met stijgende PAR-niveaus in alle drie de groepen. Opvallend is dat grotere punten (hogere PAM-waarden) geclusterd zijn rond PAR-waarden in het midden bereik ($\sim 30\text{--}50 \mu\text{mol}/\text{m}^2/\text{s}$), wat duidt op een grotere variabiliteit in ETR onder gematigde lichtomstandigheden. De Pearson coëfficiënten (r) tussen PAR en ETR zijn respectievelijk 0,98, 0,95 en 0,98 voor de groepen A, B en C, wat duidt op een consistente relatie tussen de beschikbaarheid van licht en de efficiëntie van het elektronentransport in de verschillende experimenten.

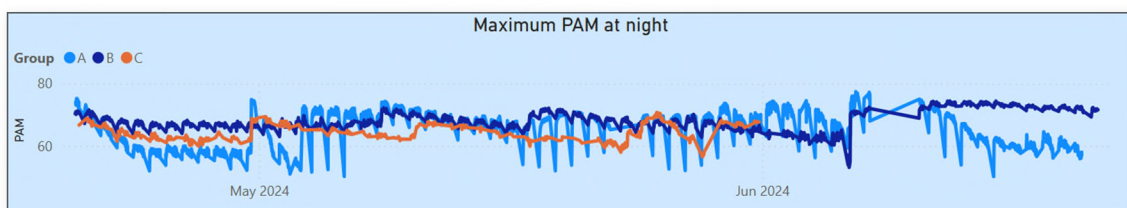
Aangezien deze drie plantgroepen A, B en C ieder met een andere kwaliteit gietwater werden geïrrigeerd, en er geen verschil tussen de groepen is waar te nemen met betrekking tot de ETR in relatie tot de beschikbare PAR, kan worden geconcludeerd dat er geen meetbare invloed van de waterkwaliteit op de fotosynthese-efficiëntie, als uitgedrukt in ETR, kan worden vastgesteld.



Figuur 6: Correlatie tussen de snelheid van het elektronentransport en de fotosynthetisch actieve straling. De grootte van de stippen geeft de grootte van de pulsamplitudemodulatie (PAM) weer. Drie experimenten uitgevoerd in de klimaatkamer van NORMEC worden weergegeven in A, B en C met verschillende kleuren.

3.1.3 Pulse-Amplitude Modulation gedurende de nacht (PAR = 0)

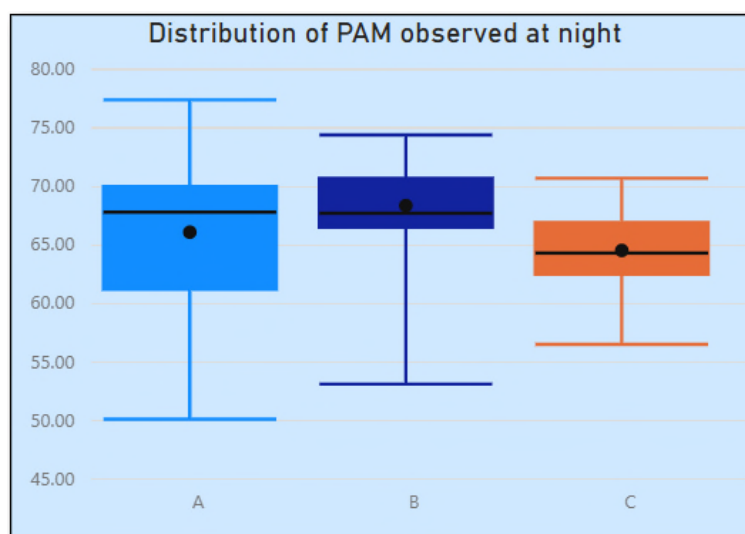
Figuur 7 toont een tijdreeks van nachtelijke PAM-waarden over verschillende weken in 2024, die variaties in herstellpatronen van planten weergeeft. Groep A (lichtblauw) had aanvankelijk de hoogste PAM-waarden, maar van begin mei tot juni wordt een dalende trend waargenomen. Dit suggereert dat planten in deze groep een toenemend niveau van nachtelijke stress ondervonden of een geleidelijke afname in hun herstelefficiëntie. Groep B (donkerblauw) vertoonde een stabiel patroon met matige fluctuaties, wat duidt op een consistentere maar licht dalende herstelltrend. Groep C (oranje) vertoonde de meest stabiele PAM-waarden, wat suggereert dat deze groep het meest uniforme nachtelijke herstel had, met een lagere efficiëntie. Er moet opgemerkt worden dat in elke plantengroep slechts 1 individuele plant uitgerust was met een bladklem-sensor om PAR en PAM te meten. Aangezien PAR en PAM correleren met zowel de omgevingsomstandigheden als met de sensorposities tijdens de experimenten (bijv. afstand tussen sensor en lichtbron, hoek van het blad) en deze niet werden geregistreerd, kan er geen oorzaak-gevolgrelatie worden vastgesteld tussen de verschillen in de gemeten PAM-waarden tussen de groepen A, B en C. Dit wordt onderstreept door het feit dat er tussen de planten visueel, en wat betreft plantgewicht, nauwelijks verschillen op te merken waren terwijl de PAM waarden voor de 3 groepen, als gevisualiseerd in figuur 7, een substantieel verschil laten zien. Dit suggereert dat andere variabelen dan de plantgezondheid/plantprestaties een belangrijke rol spelen in de resultaten verzameld met de PAM/PAR sensoren gebruikt in dit experiment.



Figuur 7: Tijdreeks van de maximale waarde van de Puls Amplitude Modulatie per nacht. Drie experimenten uitgevoerd in de klimaatkamer van NORMEC worden weergegeven in A, B en C met verschillende kleuren.

Figuur 8 geeft een statistisch overzicht van de nachtelijke PAM-waarden, door de spreiding en variabiliteit weer te geven. Groep A vertoonde de hoogste mediane PAM-waarde (rond 69), wat aangeeft dat de planten in deze groep over het algemeen een sterker nachtelijk herstel vertoonden. Het de interkwartielafstand en de uitgebreide whiskers van de boxplot geven echter een substantiële variatie tussen de gemeten waarden aan, wat overeenkomt met de waargenomen fluctuaties in de tijdreeks. De gemonitorde plant in groep B had een vergelijkbare maar iets lagere mediaan dan groep A, met een strakkere verdeling, wat duidt op een stabielere maar iets lager nachtelijk herstel dan groep A. Groep C vertoonde de laagste mediaan PAM-waarden en de kleinste variatie, wat impliceert dat de gemonitorde plant in deze groep het meest consistente herstel hadden in vergelijking met de andere twee groepen. Bovendien illustreren de zwarte stippen die de gemiddelde waarden weergeven ook dat de groepen A en B een scheefgroei vertoonden naar hogere PAM-waarden, terwijl het gemiddelde van groep C lager was.

gepositioneerd, wat impliceert dat de planten in groep C meer consistente maar gematigde stressniveaus hadden. Zoals hierboven beschreven, werd elke groep planten door één sensor gevolgd en we waarschuwen dat het trekken van conclusies uit het bovenstaande een over-interpretatie kan zijn van variaties veroorzaakt door de plaatsing van de sensoren en de prestaties van individuele planten in plaats van de gevolgen van de verschillende waterkwaliteit van behandelingen A, B en C. Dit wordt bevestigd dat ondanks de beschreven verschillen, de mediaan en gemiddelde waarden voor de drie groepen nauwelijks van elkaar verschillen, en zoals in figuur 7 te zien gedurende verschillende perioden vrijwel identiek waren, en daar waar ze van elkaar verschilden er geen consistente trend in de verhouding tussen A, B en C te duiden is. Het leggen van een verband tussen plantgezondheid en ETR, en in het bijzonder tussen waterkwaliteit zoals door de waterbehandelingen in groepen A, B en C toegepast en de plantgezondheid, is op basis van de meetgegevens uit het klimaatkamerexperiment niet mogelijk.



Figuur 8: Boxplots met de distributie van nachtelijk PAM waarden van de plantgroepen behandeld met waterkwaliteit A, B en C.

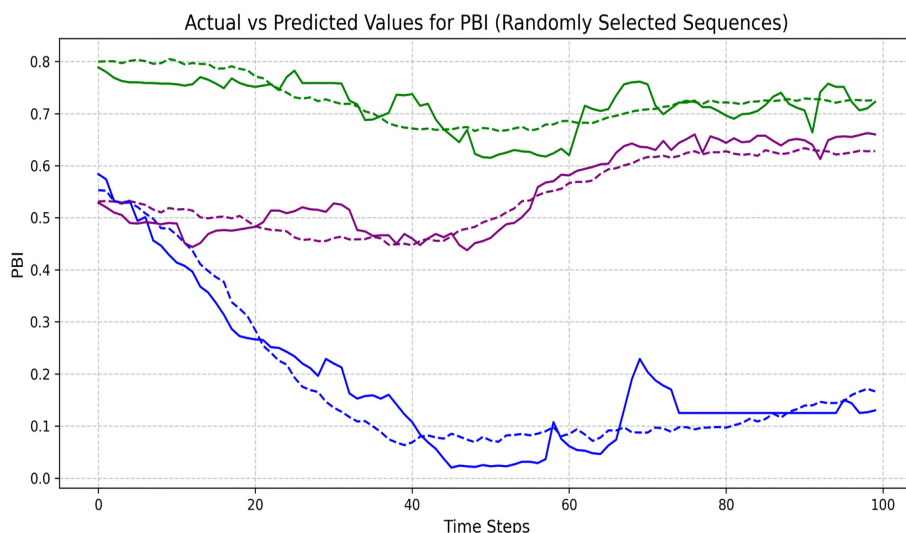
3.1.4 Resultaten behaald met de tijdreeks modellen

Voor de evaluatie van de prestaties van het model bij het voorspellen van PBI-waarden werd de correlatiecoëfficiënt (R^2) berekend op 0,93, wat aangeeft dat het model een aanzienlijk deel van de variantie in PBI kon verklaren. Verder werd de gemiddelde absolute fout (MAE) berekend op 0,0309, wat de gemiddelde grootte van de afwijking tussen voorspelde en waargenomen waarden weergeeft. Deze prestatiecijfers, beoordeeld op de validatiedataset, tonen de effectiviteit van het ontwikkelde tijdreeksmodel in het modelleren van de complexe interacties tussen waterkwaliteitsparameters en de gezondheid van planten, terwijl temporele trends nauwkeurig worden vastgelegd. Figuur 9 geeft een visuele vergelijking tussen waargenomen en voorspelde PBI-waarden voor drie willekeurig geselecteerde reeksen. In deze figuur geven de ononderbroken lijnen de waargenomen waarden aan, terwijl de stippellijnen de modelvoorspellingen weergeven.

De 'sliding window'-benadering die werd gebruikt in het tijdreeksmodel was cruciaal voor het vastleggen van de evoluerende patronen binnen de dataset, waardoor het model de dynamische interacties tussen waterkwaliteitsparameters en plantprestaties kon leren. Dankzij deze methodologie kon het model 96 tijdstappen (= 8 uur) vooruit betrouwbare voorspellingen doen over de PBI voor de specifieke dataset gebruikt in de modelontwikkeling. Dit demonstreert het potentieel van machine learning in het maken van modellen voor echte toepassingen in gecontroleerde landbouwomgevingen, waar nauwkeurige voorspellingen precisielandbouwpraktijken kunnen ondersteunen. Echter, een validatie van het model in de praktijk en onder een bredere set aan omstandigheden was niet meer mogelijk gedurende de looptijd van het project. Zo is het voorspellende vermogen bij snelle veranderingen in de input parameters (waterkwaliteit) niet uitvoerig getoetst. Te

verwachten valt dat in zulke gevallen de voorspellingskracht van het tijdreeksmodel, voor kortere tijdstermijnen, minder goed zal functioneren.

Het hier beschreven resultaat moet dan ook gezien worden als een demonstratie van de potentie van tijdreeksmodellen en niet als een in de praktijk toepasbaar model.

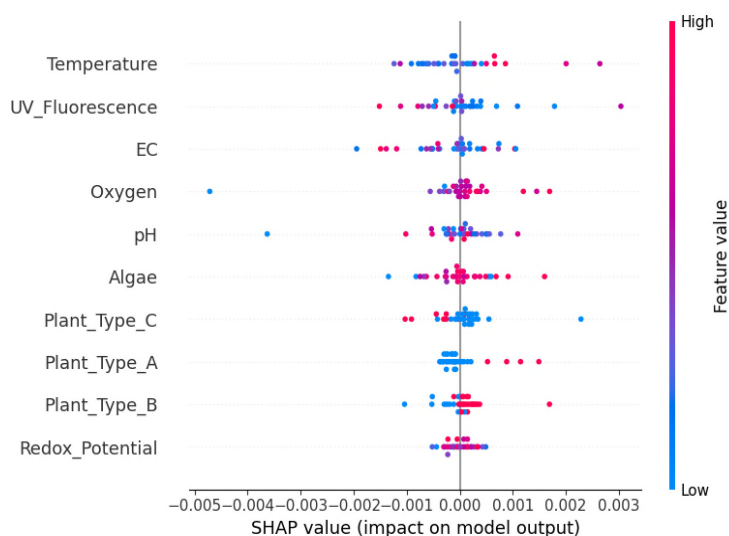


Figuur 9: Vergelijking tussen voorspelde PBI en waargenomen PBI op basis van geselecteerde voorspellingen. De ononderbroken lijnen geven de waargenomen waarden aan, terwijl de stippellijnen de modelvoorspellingen weergeven..

3.1.5 Belang van features

Om de zeggingskracht en duiding van de waarde van voorspellingen die het tijdreeksmodel doet te begrijpen, werden SHapley Additive exPlanations (SHAP) gebruikt om het gewicht van verschillende waterkwaliteitsparameters in de PBI voorspelling te kwantificeren. De SHAP-samenvattingsplot (Figuur 10) toont een gerangschikte verdeling van het belang van de verschillende input parameters. Temperatuur en UV-fluorescentie vertonen de hoogste SHAP-waarden, wat aangeeft dat ze een aanzienlijke invloed hebben op de PBI-voorspellingen. De kleurgradiënt in figuur 10, die de waarden van de kenmerken weergeeft, suggereert dat hogere temperaturen en meer UV-fluorescentie worden geassocieerd met positieve bijdragen aan PBI, wat hun rol in het bevorderen van plantengroei in de experimenten kan impliceren. Aan de andere kant vertonen kenmerken zoals pH, elektrische geleidbaarheid (EC) en algen een bredere spreiding van SHAP-waarden, wat meer variabele effecten op de uitvoer van het model impliceert. Deze variabiliteit suggereert complexe, contextafhankelijke interacties tussen deze parameters en de prestaties van de plant. De categorische variabelen die de invloed van de drie verschillende waterbehandelingen A, B en C, en daarmee de verschillende waterkwaliteiten, vertegenwoordigen vertonen relatief lage SHAP-waarden, wat impliceert dat verschillen tussen behandelingsexperimenten niet leiden tot significante effecten op PBI-voorspellingen in vergelijking met andere omgevingsfactoren. Afgezien van de rangschikking moet het diagram worden gelezen door te kijken naar de spreiding van SHAP-waarden voor elk kenmerk. Een grote spreiding in SHAP-waarden geeft aan dat het kenmerk een significante en variabele invloed heeft op de voorspellingen van het model. Deze variabiliteit suggereert dat de invloed van het kenmerk op PBI-voorspellingen kan veranderen afhankelijk van de context of omstandigheden van de gegevens.

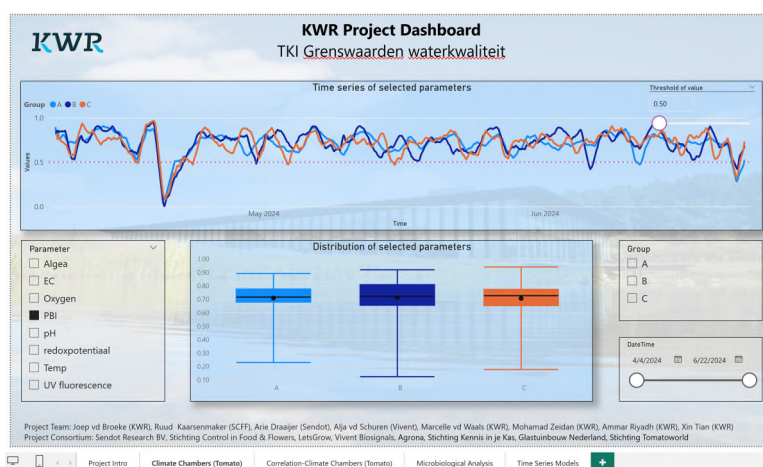
Samenvattend laat de SHAP-analyse de voornaamste parameters die bijdragen aan de voorspellingen zien, en suggereert dat omgevingsparameters (bijv. temperatuur) een sterkere invloed hebben op PBI dan variaties geïntroduceerd door de verschillende waterbehandelingen.



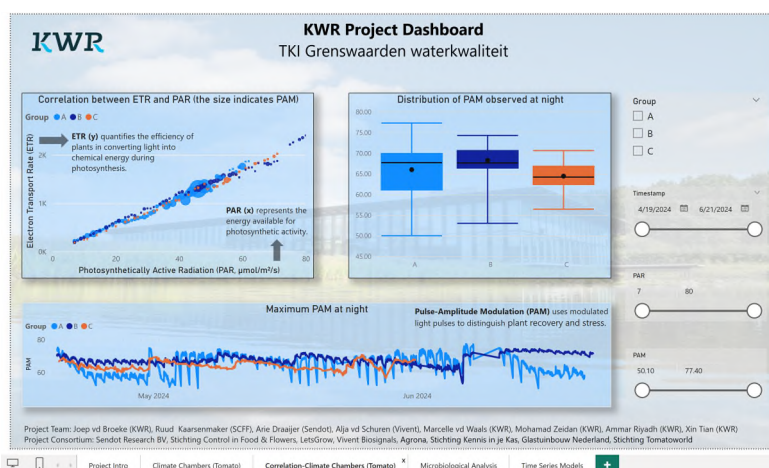
Figuur 10: SHAP gebruikt om het belang van inputparameters in het tijdreeksmodel te schatten.

3.1.6 Dashboard

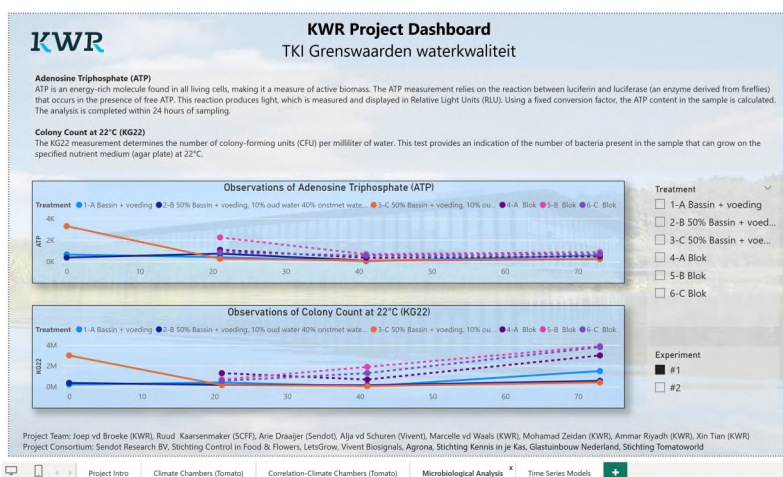
Op basis van de resultaten verkregen in de klimaatkamer-experimenten bij NORMEC is een interactief dashboard ontwikkeld om tijdreeksen, correlaties en modelresultaten te visualiseren. Aan de hand van dit dashboard kan een gedetailleerde analyse van de belangrijkste waterkwaliteitsparameters gemaakt worden (Figuur 11 tot Figuur 14). Het dashboard integreert meerdere interactieve componenten, waaronder een tijdreeksgrafiek om de variaties van acht parameters in de tijd te volgen, een verdelingsgrafiek voor vergelijkende statistische inzichten over verschillende experimentele groepen, en selectiehulpmiddelen om relevante parameters te filteren. Dit vergemakkelijkt de identificatie van trends, overschrijdingen van drempelwaarden en verschillen in de prestaties van verschillende plantgroepen en/of waterbehandelingen. Daarnaast presenteert het dashboard ook de resultaten van correlatie, microbiologische analyses (apart gerapporteerd rapport KWR 2025.023) en de tijdreeksmodellen. We merken op dat het dashboard niet publiek beschikbaar is gesteld en alleen data uit een statische dataset (gedownload uit het MyLetsgrow platform, zie ook KWR 2025.025) visualiseert. Het ontwikkelde dashboard is daarmee een voorbeeld van hoe informatie over waterkwaliteit en plantgezondheid in een geïntegreerde omgeving, die ook analyse mogelijkheden aanbiedt, weergegeven kunnen worden.



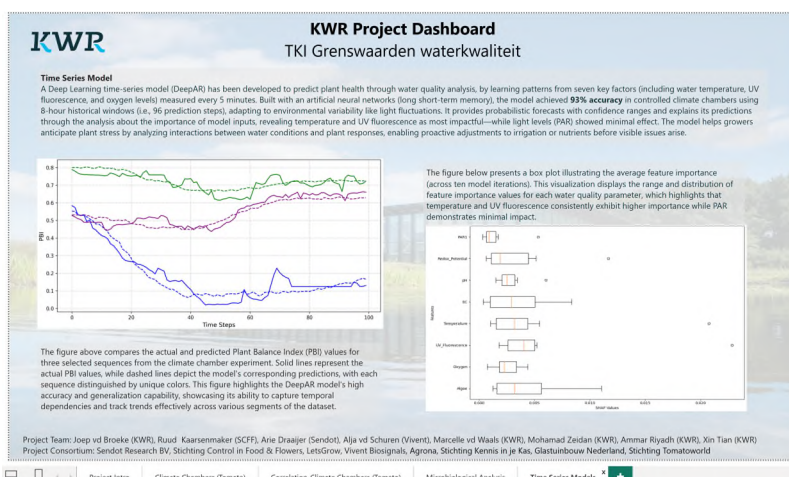
Figuur 1: Screenshot 1 van het dashboard, dat ontworpen is om tijdreeksen van gemeten parameters voor de klimaatkamer bij NORMEC te tonen.



Figuur 2: Screenshot 2 van het dashboard, dat is ontworpen om de correlatie tussen PAR en ETR te tonen voor de klimaatkamer in NORMEC.



Figuur 3: Screenshot 3 van het dashboard, dat ontworpen is om microbiële analyses te tonen voor de klimaatkamer bij NORMEC.



Figuur 4: Screenshot 4 van het dashboard, dat ontworpen is om tijdreeksmodellen voor de klimaatkamer in NORMEC te tonen.

3.2 Case 2: Tomatoworld

3.2.1 Samenvatting datakwaliteit

In deze dataset zijn de waterkwaliteitsparameters vrijwel complete (uitgezonderd redox potentiaal). De meetreeksen voor PAR en ETR laten substantiële onderbrekingen zien (~40% van de datapunten ontbreekt). Dit

heeft een negatieve impact op de analyse met betrekking tot fotosynthese. Een overzicht van de dataset is gegeven in tabel 5.

Tabel 5: Samenvatting van de data verkregen bij Tomatoworld in 2023.

	Count	Mean	Standard Deviation	Min	Max	Missing Value Ratio (%)
Algen	63215	0.0	0.1	-0.2	0.4	2.8
Opgelost zuurstof	65046	24.8	12.3	0.5	63.3	0.0
UV_Fluorescentie	65046	-91.1	38.9	-172.1	162.5	0.0
Temperatuur (water)	65046	20.5	2.6	12.8	36.0	0.0
Oxygen2	65046	6.6	9.1	0.5	50.0	0.0
Temperatuur (lucht)	65046	28.0	3.5	18.8	43.7	0.0
EC	63346	2.2	1.2	0.0	7.2	2.6
pH	64630	5.3	1.1	1.4	10.7	0.6
Redox Potentiaal	58888	326.4	103.8	0.0	1444.6	9.5
PBI	45297	0.7	0.2	0.0	1.0	30.4
PAR	36490	447.1	406.4	0.7	1927.9	43.9
ETR	36490	125.8	114.4	0.2	542.5	43.9

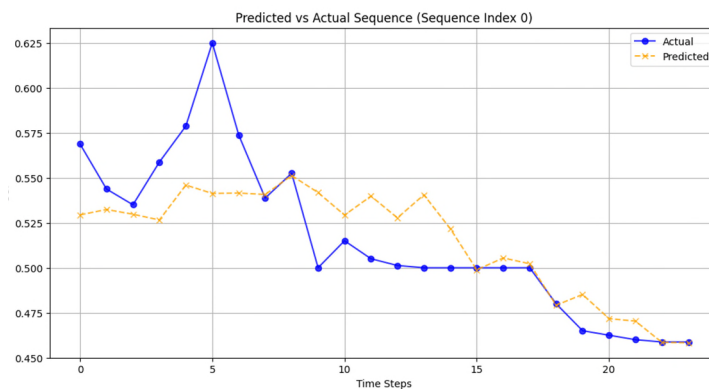
3.2.2 Tijdreeksmodel

N.B. - opgemerkt moet worden dat, hoewel de metingen bij Tomatoworld werden verzameld in 2023, dus vóór de klimaatkamerexperimenten beschreven in sectie 3.1 (Klimaatkamerexperimenten NORMEC in 2024), de gegevens van Tomatoworld werden beoordeeld nadat het modelleringswerk voor de klimaatkamerexperimenten NORMEC was afgerond. Dit had te maken met de volgorde waarin de datasets beschikbaar kwamen voor analyse. Omdat in de casus NORMEC reeds een keuze was gemaakt om alleen tijdreeksmodellering voor PBI onderzoeken, is deze beperking ook toegepast in de casus Tomatoworld. De nadruk lag hierbij op het valideren van de aanpak in de klimaatkamer met betrekking tot tijdreeksmodellering en niet op het ontwikkelen van een nieuw model.

De prestatie indicatoren laten een R^2 van 0,84 en een MAE van 0,0450 zien. Dit is slechter is dan de voorspellende waarde behaald voor de meetreeksen uit de NORMEC klimaatkamer. Dit verschil is waarschijnlijk te wijten aan de grotere omgevingsvariabiliteit in de kas van Tomatoworld, zoals licht- en vochtigheidsschommelingen. Bovendien kan de meenemen van een variabele PAR waarde (deze was immers constant in de klimaatkamer en bij Tomatoworld betreft het variabele instraling van zonlicht) andere interacties hebben geïntroduceerd die in het tijdreeksmodel ontwikkeld in de klimaatkamer niet aanwezig waren.

Figuur 15 toont een vergelijking van de werkelijke en voorspelde reeksen voor PBI op basis van het tijdreeksmodel dat is gevalideerd voor de kas in Tomatoworld. Een belangrijke observatie is dat het model weliswaar effectief de algemene neerwaartse trend in PBI volgt, maar geen plotselinge pieken en snelle veranderingen voorspelt die in de werkelijke tijdreeksen worden waargenomen, vooral in de eerste tijdstappen. Dit suggereert dat het model de korte termijn fluctuaties onvoldoende weergeeft, mogelijk als gevolg van een afvlakkingseffect dat inherent is aan probabilistische voorspellingsbenaderingen zoals die in deze studie zijn toegepast, namelijk DeepAR. Bovendien komen de voorspelde waarden naar latere tijdstappen toe meer overeen met de werkelijke waarden, wat suggereert dat het model beter in staat is om lage PBI-waarden te voorspellen. Dit gedrag suggereert dat hoewel het tijdreeksmodel goed generaliseert in dynamische omgevingen, het vermogen om snelle fluctuaties vast te

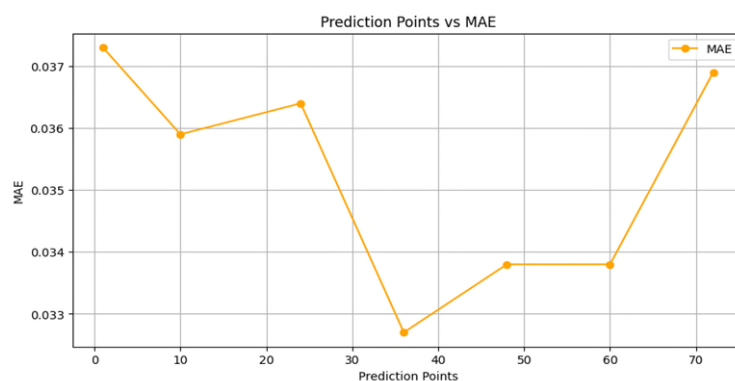
leggen beperkt kan zijn, waarschijnlijk als gevolg van de omgevingsvariabiliteit in de kasomgeving en de invloed daarvan op de plantprestaties.



Figuur 15 Vergelijking tussen voorspelde PBI en waargenomen PBI op basis van geselecteerde voorspellingen. De ononderbroken lijnen geven de waargenomen waarden aan, terwijl de stippellijnen de modelvoorspellingen weergeven.

3.2.3 Optimalisatie van het tijdvenster

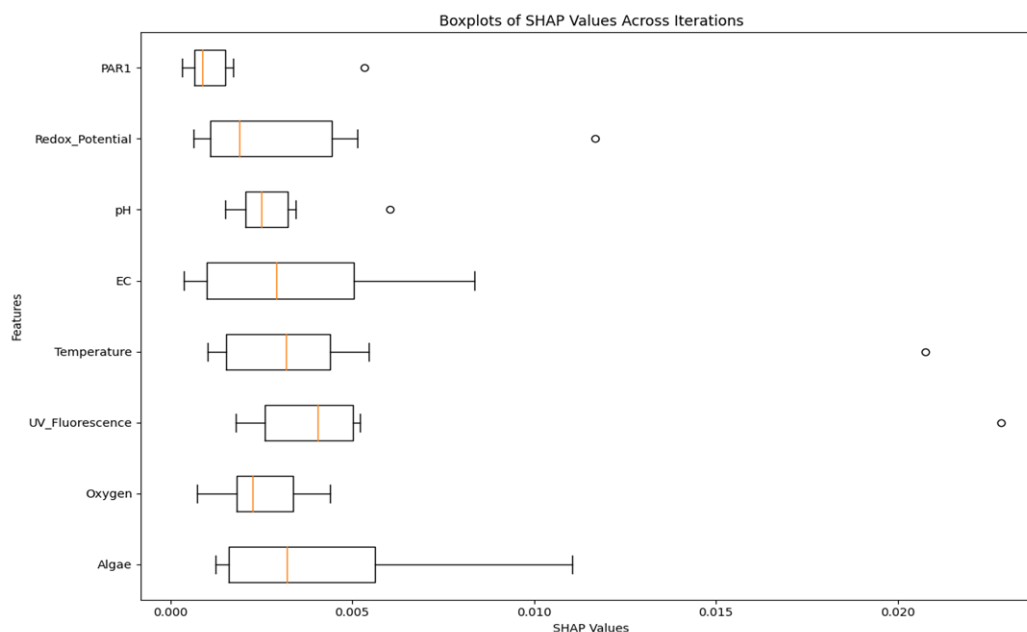
Om het resultaat te verbeteren, hebben we ook de invloed getest van de lengte van het tijdvenster dat bepaald welke historische gegevens worden gebruikt als invoer in het tijdreeksmodel. We hebben dezelfde modelarchitectuur aangehouden en geëxperimenteerd met verschillende venstergrootten van 1, 10, 24, 36, 48, 60 en 72 stappen. Met andere woorden, input met een lengte van deze stappen wordt ingevoerd in de tijdreeksmodellen om te beoordelen hoe de historische contextlengte de voorspellingsnauwkeurigheid beïnvloedt, waarbij bij venstergrootte 1 slechts de meest recente waarde wordt gebruikt als input en bij venstergrootte 72 de laatste 72 voorgaande meetpunten. Zoals weergegeven in Figuur 16, vonden we dat de optimale venstergrootte 36 tijdstappen (d.w.z. 3 uur) zou moeten zijn, omdat dit resulteerde in de laagste MAE. Dit suggereert dat kortere historische vensters effectiever waren in het vastleggen van relevante patronen in de kasomgeving zonder overmatige ruis of irrelevante gegevens uit het verleden te introduceren, met name in het geval van de kas van Tomatoworld.



Figuur 16: Omvang van het tijdvenster voor input data vs. Mean Absolute Error (MAE).

3.2.4 Belang van features

Figuur 17 geeft een statistisch overzicht van SHAP-waarden over meerdere iteraties voor de casus Tomatoworld, waarbij zowel de mediane invloed als de variabiliteit van elk kenmerk op PBI-voorspellingen wordt benadrukt. Ten eerste vertonen UV-fluorescentie, temperatuur en EC de hoogste SHAP-waarden, wat hun significante invloed bevestigt, terwijl pH, redoxpotentiaal en zuurstof ook bijdragen, maar in mindere mate. Eigenschappen met een lagere impact, zoals PAR, behouden relatief kleine SHAP-waarden, wat duidt op een incidentele maar minder belangrijke invloed.



Figuur 5: SHAP toegepast om het gewicht van features in het tijdreeksmodel te bepalen.

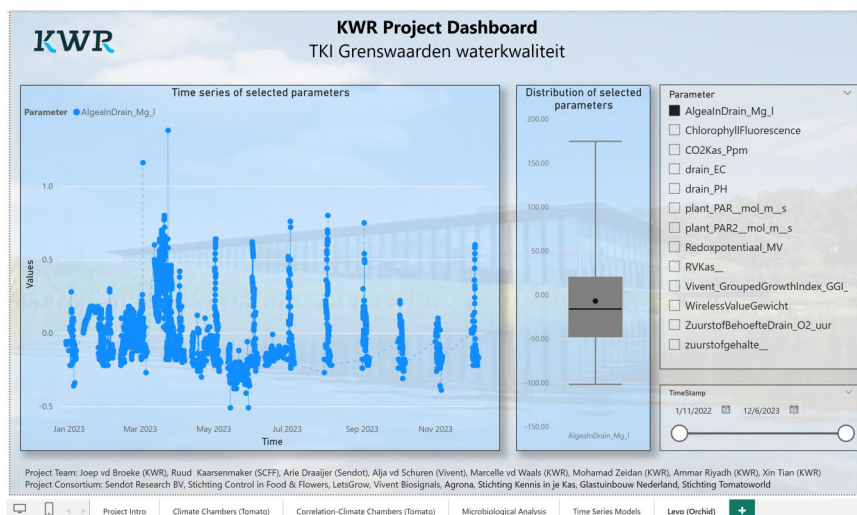
3.3 Case 3: LEVO

3.3.1 Kwaliteitsproblemen met de dataset

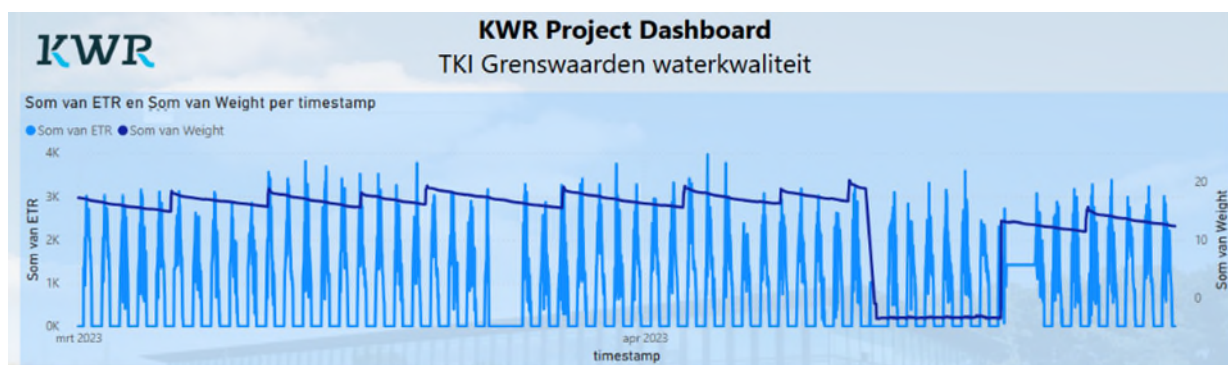
De dataset van LEVO Plant bevat een aantal kritieke problemen met de gegevenskwaliteit die een uitgebreide analyse in de weg staan. Met name het gewicht van de plant (WirelessValueGewicht) en het zuurstofverbruik (ZuurstofBehoeftedrain) vertonen een zeer hoog percentage ontbrekende waarden, met respectievelijk 66,7% en 97,6% ontbrekende gegevens. Bovendien vormt de aard van de dataset een aanzienlijke uitdaging voor het trainen van een tijdreeksmodel vanwege ondermeer de aanwezigheid van discontinue metingen, wisselende meetfrequenties voor sommige parameters, en gebrek aan synchronisatie tussen metingen uitgevoerd met verschillende sensorsystemen. Daarnaast werden ook veelvuldig herhalingen van meetwaarden in de datasets aangetroffen, wat wijst op problemen bij het uit MyLetsGrow downloaden van de meetgegevens. Aangezien tijdreeksmodellen vertrouwen op opeenvolgende afhankelijkheden en trends in de tijd, introduceren bovenstaande problemen in de gegevens inconsistenties waardoor het model geen zinvolle patronen kan leren. Daarom richten we ons in het onderstaande deel van dit hoofdstuk voornamelijk op de tijdreeksen van waargenomen parameters gemeten bij LEVO.

3.3.2 Tijdreeksen van gemeten parameters

Door de tijdreeksen in een dashboard weer te geven (figuur 18), worden de discontinuïteiten in meetfrequentie en ontbrekende gegevenspunten benadrukt. Door verschillende parameters uit het menu te selecteren, kunnen gebruikers het dashboard echter nog steeds gebruiken om de tijdreeksen en de verdeling van de geselecteerde parameters te bekijken, waardoor een gedetailleerd onderzoek van trends, variaties en hiaten in de gegevens mogelijk wordt (figuur 19).



Figuur 6: Screenshot 5 van het dashboard, dat duidelijk de discontinue aard van de data verzameld bij Levo laat zien.



Figuur 19: Screenshot 6 van het dashboard, waarin ETR (lichtblauw) en gewicht (donkerblauw) worden uitgezet tegen de tijd. Dit maakt het mogelijk om te bepalen er een afhankelijkheid is tussen deze twee parameters.

4 Discussie en Conclusies

4.1 Algemene bevindingen

Het project had als doelstelling om te onderzoeken of het monitoren van de waterkwaliteit met sensoren, in het bijzonder een combinatie van sensoren die een samen waterkwaliteits-vingerafdruk meten, aanwijzingen voor (bij)sturing van het recirculatieproces en waterbehandeling kan geven. Een vervolgvraag was of het mogelijk is op basis van dergelijke informatie grenswaarden vast te stellen voor goede kwaliteit van gietwater bij recirculatie. In dit rapport is beschreven dat er grote hoeveelheden meetgegevens zijn verzameld met behulp van waterkwaliteitssensoren. Daarnaast zijn er grote hoeveelheden aan omgevings- en plantgegevens verzameld die bedoeld waren om de invloed van waterkwaliteit op plantgezondheid en -producties te kunnen bepalen. Om deze relaties te vinden is gebruikt gemaakt van machine learning.

Op basis van de in dit rapport beschreven analyses moet worden geconcludeerd dat er voor de vraagstelling zoals hierboven benoemd geen eenduidig antwoord is gevonden. Allereerst is er gemeten onder praktijkomstandigheden bij Tomatoworld en LEVO. In de praktijksituatie zijn er verschillende variabelen, waaronder PAR en omgevingstemperatuur, die een grote invloed hebben op de fotosynthese. De waterkwaliteit is op deze locaties niet actief gevarieerd (anders dan gedurende een langere periode toevoegen van nanobubbles bij Tomatoworld) omdat er geen risico genomen kon worden om de groei van de planten negatief te beïnvloeden. Daardoor is er geen duidelijk causaal verband te vinden tussen waterkwaliteit, de impact van recirculatie en plantgezondheid op basis van de verzamelde sensordata. Dit laatste is een belangrijk punt, want zoals wordt beschreven in de andere eindrapportages bij dit project, kunnen op basis van handmatige analyses en microbiologische analyses in sommige gevallen wel verschillen in waterkwaliteit, en impact op plantprestaties, worden waargenomen. Zie hiervoor ook de rapportages S20220001 en KWR 2025.023.

Ook bij de experimenten die in de klimaatkamer bij NORMEC zijn uitgevoerd zijn meetgegevens verzameld. Bij deze experimenten is de waterkwaliteit wel actief beïnvloed door het mengen van water uit verschillende onderdelen van een watersysteem bij een tomatenteler. Daarnaast is actieve beluchting met nanobubbles toegepast. Ook in deze situatie bleek het niet mogelijk om op basis van de verzamelde sensorische waterkwaliteitsmetingen een significante correlatie tussen verschillende waterkwaliteiten, gecreëerd door de 3 verschillende behandelingen, en de plantgezondheid en -producties te vinden. Noch tussen één individuele waterkwaliteitsparameter, noch met een eenvoudige combinatie van waterkwaliteitsparameters, kon een direct relatie vastgesteld worden.

Met behulp van machine learning (DeepAR) is vervolgens onderzocht of er een tijdreeksmodel ontwikkeld kon worden met een voorspellende waarde voor plantgezondheid. Op basis van de beschikbare gegevens is het gelukt om een model met voorspellende waarde te creëren voor PBI. Daarbij moet worden benoemd dat dit is gelukt door datasegmenten van voldoende goede kwaliteit te selecteren. Daarnaast is het model ontwikkeld op data die onder gecontroleerde en weinig variabele omstandigheden werd verzameld – de planten zullen naar alle waarschijnlijkheid weinig stress hebben ondervonden. De daadwerkelijke waarde van dit model onder variabele omstandigheden en bij condities die stress in planten veroorzaken zou getest moeten worden. Hiervoor ontbrak in het project de tijd.

Op basis van de verzamelde meetgegevens en het ontwikkelde model was het nog niet mogelijk om grenswaarden te definiëren voor waterkwaliteit.

Het ontwikkelde tijdreeksmodel dient gezien te worden als het bewijs dat met machine learning technieken relaties gevonden kunnen worden, en voorspellingen kunnen worden gedaan, die op basis van een klassieke data-analyse niet gedaan kunnen worden. Het hier ontwikkelde model is hier een illustratie van. Echter, het model is nadrukkelijk geen vervanging voor het daadwerkelijk meten van de voorspelde grootte, PBI, en de voorspellende waarde onder operationele omstandigheden is niet getoetst.

Daarnaast laat het werk beschreven in dit rapport zien dat het gebruik van een grote hoeveelheid aan sensoren om een grote verscheidenheid aan parameters te monitoren niet per definitie ook een bruikbare dataset oplevert voor interpretatie en modelontwikkeling. Omdat iedere sensor betrouwbaar moet functioneren en hoogwaardige data dient te leveren, neemt de kans op problemen toe met een toenemend aantal sensoren. In het hier beschreven project is ingezet op het toepassen van een zo groot mogelijk aantal van de relevant geachte sensoren. Met de ervaring opgedaan in het project met het installeren, inregelen en onderhouden van de sensoren, kan geconcludeerd worden dat het beter geweest zou zijn om in te zetten op een kleinere set aan sensoren waarvan de toegevoegde waarde vooraf goed beargumenteerd werd. Daarmee zou de geleverde inspanning om het meetsysteem in te regelen efficiënter benut zijn. Meer is niet altijd beter.

4.2 Online metingen en tijdreeksmodel

De resultaten met het tijdreeksmodel voor PBI laten zien dat sensortechnologie en data science toegepast kunnen worden bij het onderzoeken van de relatie tussen waterkwaliteit en de gezondheid van planten in de glastuinbouw. Door online metingen van de waterkwaliteit te combineren met fysiologische indicatoren voor planten (PBI, PAM, ETR) in drie verschillende kweekomgevingen, namelijk een klimaatkamer bij NORMEC, tomatenteelt in een kas bij Tomatoworld en orchideeënteelt in een kas bij Levoplant, hebben we onderzocht of er meer inzicht gekregen kan worden in hoe variaties in de waterkwaliteit de groeiprestaties en stressreacties van planten beïnvloeden. De belangrijkste bevindingen zijn als volgt:

1. Waarde van nachtelijke PAM metingen als indicatie voor herstel van planten

Nachtelijke PAM-metingen in de klimaatkamer bij NORMEC laten zien dat er verschil is tussen de planten die met verschillende kwaliteit gietwater zijn behandeld in het herstel tijdens perioden waarin de fotosynthese inactief is. Op basis van de PAM metingen wordt gesuggereerd dat planten met hogere en stabielere nachtelijke PAM-waarden overdag beter herstellen van stress. De verschillen zijn echter klein, en kan niet worden vastgesteld of het om een significant verschil gaat omdat de variatie in de metingen binnen de verschillende groepen planten groter is dan het verschil tussen de groepen onderling.

2. Toepassing en herbruikbaarheid van tijdreeksmodellen

Het was mogelijk om met behulp van DeepAR en LSTM-gebaseerde tijdreeksmodellen met waterkwaliteits- en klimaatparameters als input in de cases NORMEC en Tomatoworld een voorspelling te doen voor een fysiologische stressindicator voor planten (PBI). In de NORMEC klimaatkamer behaalde het model bijvoorbeeld een hoge determinatiecoëfficiënt ($R^2 \approx 0,93$) en een lage gemiddelde absolute fout (MAE $\approx 0,031$). Het presteerde ook goed in Tomatoworld, hoewel het minder goed was in het vastleggen van extreme kortetermijnpieken in PBI.

Er moet worden opgemerkt dat het tijdreeksmodel niet volledig operationeel is, aangezien het alleen is getest met een aantal gegeven omstandigheden en niet kan omgaan met omstandigheden die sterk afwijken van historische metingen. Voor gebruik onder andere omstandigheden is hertraining van het model met dezelfde architectuur maar met nieuw verzamelde gegevens essentieel om de voorspellende nauwkeurigheid te behouden. Verschillen in de omstandigheden in kassen (bijv. planttype, klimaat,

waterbehandeling, irrigatiestrategie) vereisen dat het model periodiek wordt bijgewerkt om rekening te houden met bedrijfsspecifieke variaties in de loop van de tijd.

Er wordt opgemerkt dat het ontwikkelde model slechts aantoont dat met behulp van tijdreeksmodellen parameters voorspeld kunnen worden op basis van combinaties van input parameters, ook als er niet altijd een duidelijk directe relatie is tussen 1 of meer van deze input parameters en de gemodelleerde meetwaarde. In het voorbeeld in dit project is gekozen voor PBI om deze mogelijkheid te onderzoeken.

Het ontwikkelde model is nadrukkelijk geen vervanging voor het direct meten van PBI, wat een directe indicatie is voor plantfysiologische processen.

3. Randvoorwaarden voor het ontwikkelen van een tijdreeksmodel

Om het DeepAR-model effectief te gebruiken voor het voorspellen van de prestatie van planten op basis van waterkwaliteitsparameters, moet aan een aantal minimumvereisten worden voldaan. Ten eerste is het essentieel om continu gegevens van hoge kwaliteit te verzamelen, waaronder belangrijke waterkwaliteitsparameters zoals pH, opgeloste zuurstof, algenconcentratie, temperatuur, UV-fluorescentie, elektrische geleidbaarheid (EC) en fotosynthetisch actieve straling (PAR). De gegevens moeten met regelmatige tussenpozen worden verzameld om consistentie in de tijd en volledigheid te garanderen. Daarnaast heeft het model voldoende historische gegevens nodig om temporele afhankelijkheden vast te leggen en nauwkeurige voorspellingen te genereren. Hiervoor zijn voorbewerkingsstappen nodig zoals het opschonen van gegevens, interpolatie van ontbrekende waarden en standaardisatie van numerieke kenmerken. De architectuur van het model, dat LSTM-lagen (Long Short-Term Memory) bevat, moet worden verfijnd door hyperparameters zoals het aantal LSTM-lagen, verborgen eenheden per laag, batchgrootte en leersnelheid aan te passen om optimale prestaties te garanderen. Ten slotte moet de dataset worden verdeeld in een trainings-, validatie- en testset om de prestaties van het model te evalueren en de robuustheid ervan te garanderen.

4. Parameters met de hoogste voorspellende waarde in tijdreeksmodellen PBI

De SHAP-analyses laten zien dat temperatuur, UV-fluorescentie het meest bijdragen aan het verklaren van PBI-variatie in de tijdreeksmodellen ontwikkeld voor de cases klimaatkamers NORMEC en Tomatoworld. Daarnaast zijn andere parameters, zoals ORP en pH, ook in verschillende modellen bovengemiddeld belangrijk. Uit de tijdreeksmodellen komt echter geen duidelijk patroon naar voren van waterkwaliteitsparameters die consequent een hoge voorspellende waarde hebben.

5. Optimalisatie van de historische context (window size)

Bij Tomatoworld bood een historische venstergrootte van 36 tijdstappen (ongeveer drie uur) de beste balans voor de nauwkeurigheid van de voorspelling: te korte vensters konden trends niet goed voorspellen, terwijl te lange vensters onnodige ruis introduceerden.

4.3 Aanbevelingen

Op basis van de bevindingen en analyses in dit rapport kunnen verschillende aanbevelingen worden gedaan om het waterkwaliteitsbeheer en de plantenprestaties in de glastuinbouw te verbeteren:

1. **Integratie van real-time monitoringsystemen:** Het implementeren van real-time waterkwaliteitsbewakingssystemen, zoals die van Sendot en Vivent, kan waardevolle gegevens opleveren. Dit stelt telers in staat om tijdig aanpassingen te doen aan irrigatie- en bemestingsstrategieën, waardoor een optimale gezondheid en groei van de planten wordt gegarandeerd. Het gebruik van een betrouwbaar en robuust gegevensbeheersysteem is echter ook

van groot belang, om er voor te zorgen dat de verzamelde gegevens op een eenvoudig toegankelijke en makkelijk te interpreteren manier aan telers worden aangeboden. Zowel het uitlezen van grotere aantallen sensoren als het presenteren van de resultaten verdienen aandacht.

2. **Complexiteit van het werken met combinaties van grote aantallen sensoren:** Meer is niet altijd beter. Er dient een weloverwogen keuze gemaakt te worden met betrekking tot te installeren sensoren. Iedere sensor vraagt aandacht voor installatie, integratie en onderhoud. Met een toenemend aantal sensoren neemt de complexiteit van deze opgave toe. In dit project was de complexiteit te hoog, waardoor de kwaliteit en daarmee bruikbaarheid, van de resultaten in het gedrang is gekomen.
3. **Gebruik van geavanceerde plantengezondheidsindicatoren:** Door gebruik te maken van geavanceerde indicatoren voor de gezondheid van planten, zoals de Plant Balance Index (PBI) en Pulse Amplitude Modulation (PAM) fluorometrie, kan stress bij planten vroegtijdig worden gedetecteerd. Deze indicatoren geven inzicht in de fysiologische toestand van planten.
4. **Gebruik van datagestuurde modellen om de prestaties van planten te begrijpen:** Het ontwikkelen en gebruiken van tijdreeksmodellen aangedreven door deep learning, zoals het DeepAR tijdreeksmodel, kan helpen bij het voorspellen van de plantenprestaties op basis van historische waterkwaliteitsgegevens. Deze modellen kunnen telers mogelijk helpen bij het optimaliseren van waterbehandelingsprocessen en het efficiënter gebruik van hulpbronnen.
5. **Regelmatige kalibratie en validatie van sensoren:** Om de nauwkeurigheid en betrouwbaarheid van de sensorgegevens te garanderen, is het cruciaal om de sensoren die worden gebruikt voor het monitoren van de waterkwaliteit en de gezondheid van planten regelmatig te kalibreren en valideren. Dit zal helpen om de integriteit en kwaliteit van de gegevens te waarborgen die nodig is voor het ontwikkelen van voorspellende modellen en het zeker stellen van de effectiviteit van de beslissingsondersteunende systemen.

5 Dankwoord

Hierbij willen auteurs van deze rapportage Groen Agro Control, in het bijzonder Arian Bardak en Faezeh Nami, Sendot, in het bijzonder Arie Draaijer, en Vivent Biosignals, in het bijzonder Alja van der Schuren, voor hun bijdrage aan het uitvoeren van de meetcampagnes, het beschikbaar maken van de verzamelde data en de interpretatie van de resultaten. Ook willen we Tomatoworld en Levoplant bedanken voor het ter beschikking stellen van hun locatie voor het uitvoeren van metingen onder praktijkomstandigheden. Daarnaast we willen Margreet Schoenmakers bedanken voor ondersteuning vanuit Glastuinbouw Nederland. Dit project werd uitgevoerd in samenwerking met SCFF en werd mogelijk gemaakt door financiële ondersteuning van de glastuinbouw sector Stichting Kennis in je Kas en vanuit de Topsectoren Tuinbouw en Uitgangsmaterialen, Topsector Water en Maritiem, en STOWA.

6 Referenties

Arunachalam, K., Thangamuthu, S., Shanmugam, V., Raju, M., & Premraj, K. (2022). Deep learning and optimisation for quality of service modelling. *Journal of King Saud University - Computer and Information Sciences*, 34(8), 5998–6007. <https://doi.org/10.1016/j.jksuci.2022.01.016>

Baker, N.R., 2008. Chlorophyll Fluorescence: A Probe of Photosynthesis In Vivo. *Annu. Rev. Plant Biol.* 59, 89–113. <https://doi.org/10.1146/annurev.arplant.59.032607.092759>

Beatrix, W., 1996. Besluit van 12 maart 1996, houdende regels voor tuinbouwbedrijven met bedekte teelt (Besluit tuinbouwbedrijven met bedekte teelt milieubeheer).

Gent, M.P.N., Short, M.R., 2012. Effect on Yield and Quality of a Simple System to Recycle Nutrient Solution to Greenhouse Tomato. *horts* 47, 1641–1645. <https://doi.org/10.21273/HORTSCI.47.11.1641>

Genty, B., Briantais, J.-M., Baker, N.R., 1989. The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta (BBA) - General Subjects* 990, 87–92. [https://doi.org/10.1016/S0304-4165\(89\)80016-9](https://doi.org/10.1016/S0304-4165(89)80016-9)

Głowacka, K., Kromdijk, J., Salesse-Smith, C.E., Smith, C., Driever, S.M. and Long, S.P. (2023), Is chloroplast size optimal for photosynthetic efficiency?. *New Phytol*, 239: 2197-2211. <https://doi.org/10.1111/nph.19091>

Grattan, S.R., Grieve, C.M., 1998. Salinity–mineral nutrient relations in horticultural crops. *Scientia Horticulturae* 78, 127–157. [https://doi.org/10.1016/S0304-4238\(98\)00192-7](https://doi.org/10.1016/S0304-4238(98)00192-7)

Grewal, H.S., Maheshwari, B., Parks, S.E., 2011. Water and nutrient use efficiency of a low-cost hydroponic greenhouse for a cucumber crop: An Australian case study. *Agricultural Water Management* 98, 841–846. <https://doi.org/10.1016/j.agwat.2010.12.010>

Lundberg, S., & Lee, S.-I. (2017). *A Unified Approach to Interpreting Model Predictions*. <https://doi.org/10.48550/ARXIV.1705.07874>

Maxwell, K., Johnson, G.N., 2000, Chlorophyll fluorescence—a practical guide, *Journal of Experimental Botany*, 51, 659–668,

Murchie, E.H., Lawson, T., 2013. Chlorophyll fluorescence analysis: a guide to good practice and understanding some new applications. *Journal of Experimental Botany* 64, 3983–3998. <https://doi.org/10.1093/jxb/ert208>

Najdenovska, E., Dutoit, F., Tran, D., Rochat, A., Vu, B., Mazza, M., Camps, C., Plummer, C., Wallbridge, N., Raileanu, L.E., 2021a. Identifying General Stress in Commercial Tomatoes Based on Machine Learning Applied to Plant Electrophysiology. *Applied Sciences* 11, 5640. <https://doi.org/10.3390/app11125640>

Najdenovska, E.; Dutoit, F.; Tran, D.; Plummer, C.; Wallbridge, N.; Camps, C.; Raileanu, L. Classification of Plant Electrophysiology Signals for Detection of Spider Mites Infestation in Tomatoes. *Appl. Sci.* **2021b**, 11, 1414. <https://doi.org/10.1038/s41598-019-53675-4>

Oliveira, R.C.D., Silva, R.D.D.S.E., 2023. Artificial Intelligence in Agriculture: Benefits, Challenges, and Trends. *Applied Sciences* 13, 7405. <https://doi.org/10.3390/app13137405>

Raudales, R.E., Toro-Herrera, M.A., Fisher, P.R., Boldt, J.K., Altland, J.E., 2024. Paclobutrazol Residues in Recirculated Water in Commercial Greenhouses. *hortte* 34, 198–203. <https://doi.org/10.21273/HORTTECH05367-23>

Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting*, 36(3), 1181–1191. <https://doi.org/10.1016/j.ijforecast.2019.07.001>

Schreiber, U., 2004. Pulse-Amplitude-Modulation (PAM) Fluorometry and Saturation Pulse Method: An Overview, in: Papageorgiou, G.C., Govindjee (Eds.), *Chlorophyll a Fluorescence, Advances in Photosynthesis and Respiration*. Springer Netherlands, Dordrecht, pp. 279–319. https://doi.org/10.1007/978-1-4020-3218-9_11

Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 404, 132306. <https://doi.org/10.1016/j.physd.2019.132306>

Tran, D.; Dutoit, F.; Najdenovska, E.; Wallbridge, N.; Plummer, C.; Mazza, M.; Raileanu, L.E.; Camps, C. Electrophysiological assessment of plant status outside a Faraday cage using supervised machine learning. *Sci. Rep.* **2019**, 9, 17073. <https://doi.org/10.3390/app11041414>

Van Os, E.A., Gieling, T.H. and Lieth, J.H., 2019. Technical equipment in soilless production systems. In *Soilless culture* (pp. 587-635). Elsevier.