

A network diagram consisting of various sized light blue circles connected by thin white lines, set against a solid blue background. The circles vary in size and are scattered across the page, with some larger circles acting as hubs.

KWR 2022.126 | December 2022

**Begrip en beheersing
van onzekerheid in
conditiebepaling van
ondergrondse
leidingen**

Resultaten TKI Midas



Samenwerkingspartners



Rapport

Begrip en beheersing van onzekerheid in conditiebepaling van ondergrondse leidingen Resultaten TKI Midas

KWR 2022.126 | December 2022

Opdrachtnummer

403340

Projectmanager

Ir. P. (Petra) Holzhaus MSc, ir. F. (Frank) Oesterholt

Auteurs

KWR:	Spatial Insight:	Acquaint:	HDM Pipelines:
Dr. A. (Amitosh) Dash	Marc Felten	Frank van der Hulst	Dr. ir. S. (Sameera) Naib
Dr. X. (Xin) Tian	Jurjen den Besten	Werner Jousma	Henry van der Bij
Q. (Quan) Pan MSc			
Dr. ir. K. (Karel) van Laarhoven			
Dr. P. (Peter) van Thienen (eindredactie)			

Kwaliteitsborger

Dr. ir. E.J.M. (Mirjam) Blokker

Verzonden naar

Alle samenwerkingspartners

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. Werkwijzen, rekenmodellen, technieken, ontwerpen van proefinstallaties, prototypen door KWR ontwikkeld 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 intellectuele- en industriële eigendom, alsmede de auteursrechten, blijven bij KWR berusten en derhalve eigendom van KWR.

Keywords

Leidingnetten; leidingconditie; assetmanagement; ondergrond; leidingvernieuwing en -onderhoud; conditiemodellen.

Jaar van publicatie
2022

Meer informatie
Peter van Thienen
T 030 6069 602
E peter.van.thienen@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

December 2022 ©

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

Inleiding

Drinkwaterbedrijven en Waterschappen hebben steeds meer de wens om keuzen rond onderhoud en vernieuwing van hun leidingen te onderbouwen met gedetailleerde modellen voor leidingconditie. Het succesvolle gebruik van dergelijke modellen staat of valt met de beschikbaarheid van invoerdata die de toestand en situatie van het leidingbestand met voldoende detail en betrouwbaarheid beschrijft terwijl de beschikbaarheid van dergelijke data tot op heden beperkt is. De relevante modeldata kan worden verkregen door middel van inspecties, maar dit vergt een substantiële inspanning. Het is daarom van belang om inspecties zo veel mogelijk te prioriteren op de impact die ze kunnen hebben op beslissingen (de belangrijkste parameters het eerst meten) en om aanvullende data (bedrijfsdata of open data zoals bodemkaarten, wegenkaarten, KNMI data) zo goed mogelijk te benutten om de beschikbare directe metingen aan te vullen. Ook is het van belang om de mate van zekerheid (vaak wordt gesproken van betrouwbaarheid) bij deze parameters in de verwerking en het gebruik van de gegevens mee te nemen. Het voorliggende rapport beschrijft hoe in het project TKI Midas is gewerkt aan een verbeterde beschikbaarheid van data voor gebruik in leidingconditiemodellen en aan het inzichtelijk maken van het belang van individuele datacomponenten (en hun onzekerheid) voor de uitkomsten van deze modellen. Hierbij hebben de partners Acquaint, HDM Pipelines en Spatial Insight zich gericht de vergaring en kwalificering van beschikbare data, en KWR op de analyse en voortplanting van onzekerheid in deze data. Dit in samenspraak met eindgebruikers Brabant Water, Waterschapsbedrijf Limburg en Waterschap Zuiderzeeland om de aansluiting op de praktijkbehoeften te maximaliseren.

Conclusies en opbrengsten

Het project heeft een reeks bruikbare inzichten en concrete opbrengsten opgeleverd. Ten eerste is een uitputtend overzicht gegenereerd van de databehoeften voor conditiebepaling van ondergrondse leidingen. Daarbij zijn ook de beschikbaarheid van kwaliteit van de data geanalyseerd, en is een interactieve kaart gemaakt om deze aspecten inzichtelijk te maken. Vervolgens heeft een analyse van correlaties in omgevingsparameters die van invloed op de degradatie van leidingen kunnen zijn laten zien dat er geen surrogaatparameters aan te wijzen zijn die gebruikt kunnen voor parameters waarvoor geen (meet)waarden beschikbaar zijn. Dit is een negatief resultaat, maar niettemin belangrijk. Het dient de aanbeveling om over ca. 10 jaar deze analyse te herhalen met de data die dan beschikbaar zullen zijn.

Voor het kwantificeren van onzekerheid en de voortplanting daarvan in conditiemodellen is een methodiek uitgewerkt. Deze heeft concreet vorm gekregen in een dashboard. Met dit dashboard kan de voortplanting van onzekerheid inzichtelijk worden gemaakt, en kan worden beoordeeld welke parameters het meest bijdragen aan onzekerheid in conditiemodellen – dit zijn de parameters waar bij metingen of andere datavergeving de focus op zou moeten liggen, om de kwaliteit en bruikbaarheid van de uitkomsten van conditiemodellen zo veel mogelijk te verbeteren.

Inhoud

Samenwerkingspartners	1
Rapport	2
Samenvatting	3
Inhoud	4
1 Inleiding	6
1.1 Achtergrond	6
1.2 Doel	6
1.3 Leeswijzer	6
2 Parameters die de conditie van buizen beïnvloeden	7
2.1 Inleiding	7
2.2 Kwalitatief overzicht van invloedsfactoren	7
2.3 Overzicht van kwantitatief beschreven mechanismen	11
3 Dataintegratie en -kwaliteit	16
3.1 Dataintegratie	16
3.1.1 Gewenste data	16
3.1.2 Beschikbaarheid van data	16
3.1.3 Methoden	17
3.2 Analyse van de kwaliteit van en onzekerheden in databronnen	17
3.2.1 Planning en voorbereiding	17
3.2.2 Analyse kwaliteit databronnen	18
3.2.3 Implementatie en analyse	18
4 Correlatieanalyse	21
4.1 Inleiding	21
4.2 Methodologie	21
4.2.1 Stroomschema en stappen	21
4.2.2 Gebruikte programmeeromgeving	23
4.3 Resultaten	25
4.4 Conclusies en aanbevelingen	25
5 Kwantificering en propagatie van de onzekerheid	26
5.1 Hoe de eindgebruikers onder operationele omstandigheden gebruik kunnen maken van het Midas dashboard	26

5.2	Inleiding in onzekerheid	26
5.3	Onzekerheidspropagatie	27
5.3.1	Algemeen overzicht van onzekerheidspropagatie	27
5.4	Gevoeligheidsanalyse	30
5.4.1	Types gevoeligheidsanalyses	30
5.4.2	Visualisatie van de gevoeligheidsanalyse op het Midas- dashboard	31
5.4.3	Implementatie in Midas-dashboard	33
5.5	Onzekerheidspropagatie	33
6	Samenvatting en aanbevelingen voor eindgebruikers	35
6.1	Samenvatting	35
6.1.1	Databeschikbaarheid en -kwaliteit	35
6.1.2	Correlatieanalyse	35
6.1.3	Gevoeligheidsanalyse en onzekerheidspropagatie	35
6.2	Aanbevelingen voor praktische toepassing	36
6.3	Aanbevelingen voor onderzoek	36
7	Referenties	38
I	Invoerparameters voor modelberekeningen	40
II	Datavereisten en -beschikbaarheid	44
III	Databeschikbaarheid bij de verschillende betrokken partijen	64
IV	Koppelingen	65
V	Histogrammen datakwaliteitsbeoordeling	67
VI	Gebruiksvoorbeelden Midas-dashboard	73
VII	Resultaten correlatieanalyse	80

1 Inleiding

1.1 Achtergrond

Drinkwaterbedrijven en Waterschappen hebben steeds meer de wens om keuzen rond onderhoud en vernieuwing van hun leidingen te onderbouwen met gedetailleerde modellen voor leidingconditie. In recent gezamenlijk onderzoek van de Nederlands-Vlaamse drinkwaterbedrijven zijn dergelijke modellen ontwikkeld (zie bijvoorbeeld Wols, Moerman et al., 2015; Wols en Moerman, 2017; van Laarhoven et al, 2019). Daarbij is ook duidelijk geworden dat het succesvolle gebruik van dergelijke modellen staat of valt met de beschikbaarheid van invoerdata die de toestand en situatie van het leidingbestand met voldoende detail en betrouwbaarheid beschrijven. Ook is gebleken dat de beschikbaarheid van dergelijke data tot op heden beperkt is. De relevante modeldata kan worden verkregen door middel van inspecties, maar dit vergt een substantiële inspanning. Het is daarom van belang om inspecties zo veel mogelijk te prioriteren (zie ook Moerman et al., 2016) op de impact die ze kunnen hebben op beslissingen (de belangrijkste parameters het eerst meten) en om aanvullende data (bedrijfsdata of open data zoals bodemkaarten, wegenkaarten, KNMI data) zo goed mogelijk benutten om de beschikbare directe metingen aan te vullen. Ook is het van belang om de mate van zekerheid bij deze parameters in de verwerking en het gebruik van de gegevens mee te nemen.

1.2 Doel

Dit project heeft als doel de beschikbaarheid van data voor gebruik in leidingconditiemodellen te verbeteren en het belang van individuele datacomponenten (en hun onzekerheid) voor de uitkomsten van deze modellen inzichtelijk te maken. Het toepassen van dergelijke conditiemodellen staat of valt met het hebben van de juiste data. De uitkomsten van het project geven waterschappen en drinkwaterbedrijven de mogelijkheid om die data zo efficiënt mogelijk te verzamelen (alleen de belangrijke data kopen/meten) en inzicht in de benodigde betrouwbaarheid van inspectieresultaten. Door het gebruik van conditiemodellen te ondersteunen, geven de uitkomsten van dit project waterschappen en drinkwaterbedrijven toegang tot beter onderbouwde, nauwkeurigere inschattingen van de faalkansen van hun leidingen, die zij kunnen gebruiken voor de ondersteuning van hun beslissingen rond de vernieuwing van leidingen. Dit minimaliseert kosten door het vermijden vanodeloze investeringen (alleen de juiste leidingen vernieuwen) en doelgericht inspecteren en het minimaliseert de kosten en overlast van storingsen (storingsgevoelige leidingen worden het eerst vernieuwd). De uitkomsten bieden dienstverleners rondom assetbeheer van leidingen de mogelijkheid om conditiemodellen in de komende jaren op een gefundeerde manier in te gaan zetten bij het duiden van omgevingsdata en inspecties voor hun klanten, en om gericht inspecties te ontwikkelen die eventuele hiaten in data kunnen opvullen.

1.3 Leeswijzer

Dit rapport is als volgt opgebouwd. Om te beginnen geeft Hoofdstuk 2 een breed en uitputtend overzicht van interne en omgevingsparameters die van invloed (kunnen) zijn op de conditie van een leiding onder druk. Hoofdstuk 3 beschrijft de (bepaling van) de kwaliteit van en onzekerheden in bestaande en gebruikte databronnen. Vervolgens onderzoekt hoofdstuk 4 correlaties in de voor dit project bijeengebrachte data. Hoofdstuk 5 beschrijft hoe onzekerheid gekwantificeerd kan worden en hoe deze onzekerheid in modellen kan worden meegenomen. Ten slotte biedt Hoofdstuk 6 een samenvatting met praktijkgerichte aanbevelingen voor eindgebruikers.

2 Parameters die de conditie van buizen beïnvloeden

2.1 Inleiding

In grote lijnen kan het verband tussen de toestand van de leiding en een beïnvloedende factor op twee manieren worden bepaald:

1. Fysische/mechanische modellen – op basis van begrip van de fysische mechanismen wordt met behulp van een model de invloed van diverse factoren doorgerekend en gekwantificeerd. De opgelopen spanning in een leiding dient als proxy voor leidingconditie
2. Statistische modellen – een statistische relatie tussen het falen van leidingen en diverse factoren. Hier worden correlaties gezocht.

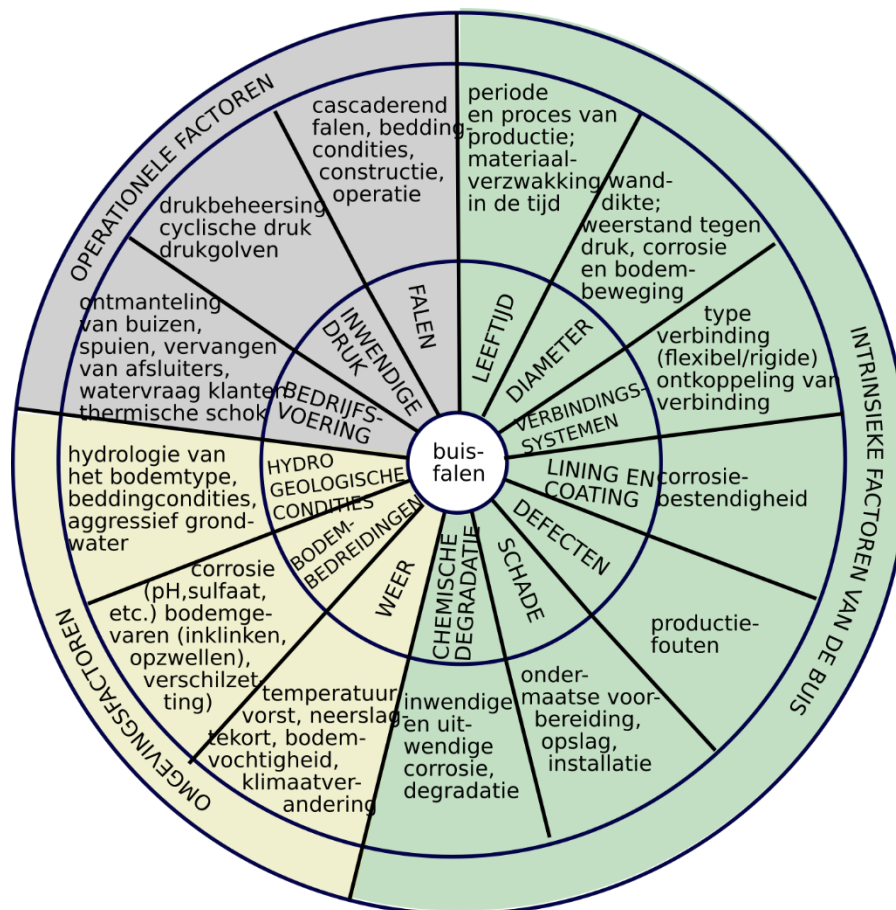
In de literatuur is gezocht naar modellen (en de bijbehorende noodzakelijke invoerparameters) die een expliciete wiskundige relatie hebben met de conditie van de buis, dus fysisch/mechanische modellen. Een voorname voorbeeld hiervan in KWR-onderzoek is COMSIMA (Wols, et al., 2015; Wols & Moerman, 2017; van Laarhoven, 2020). KWR heeft al vaak naar statistische modellen onderzoek gedaan en dat levert beperkt inzicht. Vandaar de behoefte om naar fysisch/mechanische modellen naar te kijken.

De beschouwde literatuur bestaat voornamelijk uit klassieke leerboeken en drie overzichtsartikelen die fysische modellen toevoegen die niet in bovengenoemde leerboeken zijn opgenomen. Het hier gepresenteerde overzicht is bedoeld om een overzicht te geven van bestaande modellen waarin een expliciet wiskundig verband tussen de buisconditie en een invoerparameter wordt gepresenteerd.

De uitkomsten van het literatuuronderzoek zijn in de volgende twee paragrafen beschreven, en zijn vervolgens in overzicht geplaatst in Tabel 5 (Bijlage I). In de tabel zijn de verschillende parameters opgenomen (rijen) die een rol spelen in één of meerdere van de verschillende kwantitatieve modellen (kolommen). Een aantal van de modellen voor verschillende aspecten van conditie komen samen in het conditiemodel Comsima aan de hand waarvan de voorbeelden in dit onderzoek zijn geïllustreerd. De deelmodellen/parameters die in Comsima een rol spelen zijn boven in de tabel geclusterd. Comsima fungeert als referentiemodel, omdat in dit onderzoek wordt het voor een uitgebreid gevoeligheidsanalyse en onzekerheidsvoortplanting gebruikt.

2.2 Kwalitatief overzicht van invloedsfactoren

Een recente overzichtsartikel (Barton, et al., 2019) biedt een uitgebreid overzicht van de talrijke factoren die van invloed zijn op het falen van buizen. Dit overzichtsartikel biedt vooral heuristische perspectieven en niet zozeer fysisch-wiskundige modellendelen. De invloedrijke factoren worden grofweg in drie typen ingedeeld: intrinsieke (buis), omgeving en operationeel (zie Figuur 1). Verder geven de auteurs een overzicht van de verschillende faalwijzen en mechanismen per materiaalsoort (zie Tabel 1).



Figuur 1: Overzicht van factoren die van invloed zijn op het falen van drinkwaterleidingen (naar Barton et al., 2019)

Tabel 1: Samenvatting van de faalwijzen en -mechanismen per materiaalsoort (naar Barton et al., 2019).

Materiaal	Mechanisme			Typische faalmodus
	Buis - intrinsiek	Omgeving	Operationeel	
Grijs gietijzer	Buisdiameter productiefouten grafitisering buisbescherming starre verbindingen aanleg en reparatie - incidentele schade	lage temperaturen vorst lage interne watertemperatuur sterk corrosieve bodems	materiaalmoetheid door cyclische druk drukgolven Operatie	rondbreuk falen van verbinding lengtebreuk chemische aantasting
Staal, nodulair gietijzer	Dunne buiswand productiefouten buisbescherming starre verbindingen	sterk corrosieve bodems	hoge druk materiaalmoetheid door cyclische druk	chemische aantasting falen van verbinding
AC	Buisdiameter productiefouten buisbescherming starre verbindingen	hoge temperaturen geringe regenval variërende bodemvochtigheid. Potentiële krimp en zwellen van klei en veen sterk corrosieve bodems	hoge druk materiaalmoetheid door cyclische druk	rondbreuk falen van verbinding chemische aantasting lengtebreuk
PVC	ondermaatse installatie - oplosmiddel opslag - UV-blootstelling productiefouten belasting - gevoeligheid voor puntbelasting	hoge temperaturen geringe regenval variërende bodemvochtigheid. Potentiële krimp en zwellen van klei en veen zandige bodems - uitspoeling	hoge druk materiaalmoetheid door cyclische druk	falen van verbinding lengtebreuk
PE	ondermaatse installatie van verbindingen	-	-	lengtebreuk falen van verbinding

Tabel 2: Criteria die de besluitvorming over assetbeheer kunnen beïnvloeden (naar Salehi et al., 2018)

Criteria	Hoofd-	Sub-	Code	Ondergeschikte subcriteria	criterium- waarde voor hogere prioriteit voor de buis	
Technische	Hydraulisch		01	Druk(m)	lager/hoger	
			02	Stroomsnelheid (m/s)	lager/hoger	
			03	Leeftijd (jaar)	ouder	
			04	lengte (m)	hoger	
			05	diameter (mm)	lager	
			06	diepte (m)	dieper/ ondieper	
			07	Wandruwheid	verschillend	
			08	Onkwetsbaarheid bij installatie levensduur		
	Eigenschap- pen waternet- werk	Mechanisch		09	buisseigen- schappen	
				10		gemak van onderhoud
				11	buis-materiaal	maximale externe belasting
				12		uitwendige en inwendige corrosie
				13		niet-drijvend
				14		hittebestendigheid
				15		aardbevingsbestendigheid
				16	Verbindings- meervoudig- heid (N)	hoger
17				faalfrequentie (N/km.jaar)	hoger	
18				Smaak (SI ¹) (mg/l)	hoger (als SI)	
19				Kleur (TCU)	dieper	
Eigenschappen water						20
	21	residueel chloor (mg/l)	lager/hoger			
	22	Corrosie (RSI ²) (pH)	lager/hoger (als RSI)			
	23	Kwaliteit (WQI ³) (mg/l)	lager (als WQ)			
	24	Elektrolytische corrosie (W.cm)	lager			
	25	pH (pH)	lager			
	26	Redoxpotentiaal (mV)	lager			
	Eigenschappen bodem					27
28				bodemtype	fijner	

<i>Criteria</i>	<i>Hoofd-</i>	<i>Sub-</i>	<i>Code</i>	<i>Ondergeschikte subcriteria</i>	<i> criterium- waarde voor hogere prioriteit voor de buis</i>
Niet- technische	Klimaat		29	gemak van opgraven	gemakkelijker
			30	type bedding (CF ⁴)	hoger (als CF)
			31	klimaatype (DMI ⁵)	hoger (als DMI)
			32	luchttemperatuur (°C)	lager/hoger
			33	klantdichtheid (N)	hoger
			34	stedelijke structuur (type)	verschillend in termen van de locatie in woon/commercieel/industrieel gebied
		Bevolking	35	omvang watervraag (l/s)	hoger
			36	levensduur klantgebouwen (jaar)	ouder
		Lokaal	37	padtype	verschillend in termen van pad- en bedekkingstype
			38	padbedekkings-type	
	Veiligheid		39	afstand tot bevolkingsconcentraties	korter
			40	afstand tot bovenstroomse drukreducerkleppen	langer
			41	afstand tot brandkranen	korter
			42	potentieel van biologische/hydraulische bedreiging	hoger (buisdiameter)

¹ Scholler Index (SI)

² Ryznar stabiliteitsindex (RSI)

³ Waterkwaliteitsindex (WQI)

⁴ Compactiefractir (CF)

⁵ De Martonneindex (DMI)

Verschillende parameters kunnen relevant zijn bij de beoordeling van de toestand van de buis (Salehi, et al., 2018). Deze factoren worden in Tabel 2 weergegeven. Het geeft een breed overzicht van mogelijke factoren die in overweging genomen moeten worden. Uiteraard kunnen niet alle genoemde criteria in fysisch-wiskundige modellen in acht genomen worden.

2.3 Overzicht van kwantitatief beschreven mechanismen

De invloed van diverse belastingen voor kwantitatieve conditie modellering zijn in diverse bronnen beschreven (Moser & Folkman, 2008; Rajani & Kleiner, 2001; St. Clair & Sinha, 2012; Wilson, et al., 2015). De belastingen op ondergrondse leidingen kunnen grofweg worden onderverdeeld in interne en externe belastingen. De interne belasting bestaat voornamelijk uit hydrostatische druk en drukpieken, terwijl de externe belasting wordt veroorzaakt door bodemdruk en/of oppervlaktebelasting.

A Hydrostatische druk

De meest gebruikte formule om de maximale spanning in een buis te schatten is de zogenaamde Barlow-formule (een reductie van Lamé's oplossing voor spanningen in een dikwandige cilinder). Een herformulering van deze formule wordt ook als vergelijking gebruikt door de International Standards Organization. Hiervoor is de volgende informatie nodig: **de inwendige druk, de uitwendige diameter van de buis** en de **wanddikte**.

B Drukpieken

Drukpieken kunnen grofweg worden ingedeeld in cyclische pieken of transiënte pieken (waterslag). Cyclische pieken leiden tot bezwijken door vermoeiing door langzame scheurgroei. Hiervoor zijn inputparameters voor langzame scheurgroei nodig: **initiële scheurgrootte, schommelingen in de spanningsintensiteitsfactor, fitparameters (bijvoorbeeld voor de wet van Paris)**

Voor waterslag wordt gewoonlijk de wet van Joukowsky gebruikt om de omvang van drukschommelingen te voorspellen. Hiervoor moet de golfsnelheid worden berekend, die afhangt van de eigenschappen van de buis, zoals **elasticiteitsmodulus, inwendige diameter** en **wanddikte**, alsmede van de eigenschappen van de vloeistof, zoals onder ander **elasticiteitsmodulus, dichtheid, hoeveelheid lucht**. Om te bepalen of waterslag al dan niet optreedt, moet de kritische tijd worden bepaald waarvoor bovendien de **buislengte** en de **effectieve sluitingstijd** van de afsluiter bekend moeten zijn. Als de wet van Joukowsky kan worden toegepast, moet ook de **grootte van de snelheidsverandering** bekend zijn.

C Bodemdruk

De Marston-belastingvergelijking is een algemeen erkende methode voor het voorspellen van grondbelastingen op ondergrondse leidingen. Bij plaatsing in een sleuf kan de belasting op een stijve buis worden bepaald met behulp van de **diepte van de ingraving, de sleufbreedte, de verhouding van het gewicht per eenheid grond, de Rankine-ratio** (verhouding van laterale tot verticale gronddruk) en de **wrijvingscoëfficiënt tegen de wanden van de sleuf** moeten bekend zijn. Deze worden gebruikt om belastingscoëfficiënten te berekenen die op hun beurt de belasting bepalen.

Bij plaatsing in een talud (wanneer de bovenkant van de leiding zich boven het natuurlijke grondoppervlak bevindt) kan de berekening van de belastingen omslachtig worden. Afhankelijk van de zettingsverhoudingen (die afhangen van de **samendrukking van de grond aan de zijkanten van de buis, de zetting van het natuurlijke grondoppervlak aan de zijkanten van de buis, de zetting van de fundering onder de buis, de doorbuiging van de bovenkant van de buis**) is er sprake van *projection conditions* (positieve zettingsverhouding) of *ditch conditions* (negatieve zettingsverhouding). Naast de invoerparameters die nodig zijn voor een stijve buis in een greppel, moeten ook de projectieverhouding (**verhouding tussen de verticale afstand van het natuurlijke grondoppervlak tot de bovenkant van de buis en de buisdiameter**) en de **hoogte van het vlak van gelijke zetting** bekend zijn.

Bij plaatsing in een tunnel of doorpersing in ongeroerde grond kan de Marston-tunnelbelastingvergelijking worden gebruikt. Daartoe moeten de **maximale tunnelbreedte** (of **buitendiameter van de buis** bij een doorpersing), een **belastingscoëfficiënt** (berekend zoals in voorgaande gevallen) en een **cohesiecoëfficiënt** bekend zijn.

Voor een flexibele buis die in de grond wordt geplaatst waarbij de zijvulling en de buis dezelfde stijfheid hebben, kan de Marston-Spangler-belastingvergelijking voor flexibele buizen worden gebruikt. Deze vergelijking levert doorgaans een ondergrens op voor de door de buis te ondergaan belasting. De belasting kan eenvoudig worden bepaald aan de hand van een **belastingscoëfficiënt, de verhouding van het gewicht per eenheid grond, de buisdiameter en de sleufbreedte**. Voor een flexibele buis biedt de schatting van de prismabelasting (gewicht van een verticaal prisma van grond over de buis) een meer realistische benadering. Hiervoor zijn alleen **de verhouding van het gewicht per eenheid grond** en de **diepte waarop de gronddruk wordt uitgeoefend** nodig.

D Belasting in de lengterichting

Buizen worden zelden alleen aan verticale grondbelasting blootgesteld. Waarnemingen van buisbreuken door de jaren heen suggereren dat andere krachten axiale buigspanningen in de buis veroorzaken. Bronnen van axiale buiging of balkwerking zijn onder meer niet-uniforme beddingondersteuning, differentiële zetting, grondbeweging door externe krachten (zoals aardbevingen of vorstschade). Voor asymptotische gevallen kan een expliciete wiskundige relatie worden gelegd tussen de toestand van de leiding en bepaalde invoerparameters met het zogenaamde Winkler-model (Hetenyi, 1946). Hiervoor moeten de **buigstijfheid van de pijpleiding**, het **zettingsprofiel in het groene veld**, de **coëfficiënt van de ondergrondreactie** en de **straal van de pijpleiding** bekend zijn.

E Wielbelasting (onder spanning)

De Boussinesq-oplossing wordt vaak gebruikt om de belasting op een ingegraven leiding door het grondverkeer te berekenen. Om de belasting per lengte op de buis te bepalen, de **geconcentreerde belasting**, een **impactfactor**, de **effectieve lengte van de buis**, de **belastingscoëfficiënt** (die verder afhangt van de **hoogte van de vulling vanaf de bovenkant van de buis tot het grondoppervlak** en de **buisdiameter**).

Om de op de buiswand overgedragen druk ten gevolge van wegen/spoorwegen te schatten, moeten de **geconcentreerde belasting aan het oppervlak boven de buis**, de **hoogte van de afdekking** en de **offset-afstand van de buis tot de lijn van de oppervlaktebelasting** bekend zijn. Er moet een minimale bodembedekking boven een leiding zijn om schade door belasting onder spanning te voorkomen.

F Bodemdaling

Het effect van bodemdaling wordt zelden in het ontwerp meegenomen. Gewoonlijk worden bodembewegingen toegepast op de ondergrondse leiding, hetzij via grondveren, hetzij door de leiding rechtstreeks te vervormen als een balk die de bodemcontouren volgt. Vervolgens worden de spanningen berekend. In een zeer vereenvoudigd geval, waarin een rechte ingegraven buis een trekkracht in de lengterichting ondervindt, moet men voor het schatten van de spanning in de buis de **elasticiteitsmodulus van de buis** kennen, de **trekkracht langs de as van de buis**, de **doorsnede van de buiswand** en de **wrijving tussen buis en grond in de lengterichting**.

G Belastingen door temperatuurstijging

In het scenario waarin de vloeistof in de pijpleiding warm of koud is ten opzichte van de omringende bodemtemperatuur, ontstaan spanningen doordat de uitzetting/inkrimping van de buis wordt beperkt door de omringende bodem. Voor lange stukken rechte pijpleidingen kan de langsspanning worden bepaald aan de hand van de **elasticiteitsmodulus van de buis**, de **thermische uitzettingscoëfficiënt**, de **maximale bedrijfstemperatuur**, de **installatietemperatuur**, de **Poisson-ratio** en de **omtrekspanning als gevolg van de vloeistofdruk**.

H Vorstbelasting

Er bestaan modellen om de vorstbelasting in een typische sleuf te voorspellen. De modellen vereisen gegevens zoals de **continue bevroeringsindex**, **eigenschappen van de aanvulgrond zoals porositeit, segregatiepotentieel, watergehalte in niet-bevroren toestand en thermische gradiënt aan het bevroeringsfront**, **vorstdiepte** en andere **gerelateerde variabelen om de vorstbelasting op een op een bepaalde diepte begraven leiding te voorspellen**. Deze modellen zijn complex en sommige invoerparameters zijn niet direct beschikbaar. Het is onduidelijk of de modellen zodanig zijn gevalideerd dat zij met vertrouwen kunnen worden gebruikt.

I Analyse buis-bodeminteractie

De klassieke manier om de interactie tussen buis en bodem te bestuderen bestaat erin de spanningen ten gevolge van de (i) interne druk, (ii) de compressie van de ring ten gevolge van externe belastingen en (iii) de spanning ten gevolge van de doorbuiging van de ring ten gevolge van statische en dynamische belastingen bij elkaar op te tellen. Verschillende van de bijdragende krachten zijn eerder in de tekst beschreven (en vormen tevens de basis voor COMSIMA) en de noodzakelijke invoerparameters worden niet herhaald. Dit is het zogenaamde Spangler-Watkins model waarbij de primaire belasting-weerstandsactie alleen in verticale richting plaatsvindt. Dit is geschikt voor middelgrote tot grote buisdiameters en ontoereikend voor buizen met een kleine diameter. Bovendien houdt dit model geen rekening met de invloed van de grond- en watertemperatuur.

Een longitudinaal model (Rajani, et al., 1996) houdt wel rekening met de invloed van veranderingen in interne/externe druk en temperatuursveranderingen. Voor dit model zijn de **elasticiteitsmodulus van de buis**, de **Poisson ratio van de buis**, de **thermische lineaire uitzettingscoëfficiënt van de buis**, de **interne druk**, het **temperatuurverschil**, de **elastische grondmodulus**, de **buisdiameter**, de **wanddikte**, de **reactiemodulus tussen buis en bodem**, en **constanten die functies zijn van grond- en buiseigenschappen** nodig.

J Resterende structurele weerstand

Om de druk te voorspellen waarbij een buis met een corrosieput zou bezwijken, zijn **de maximale diepte van het corrosiedefect**, **de dwarsdoorsnede van het metaal dat in het gecorrodeerde gebied verloren is gegaan**, **geprojecteerd op de lengte van de buis**, **de oorspronkelijke dwarsdoorsnede van het gecorrodeerde gebied**, **de maximale totale lengte van het corrosiedefect**, **de straal van de buis**, **de dikte van de buiswand**, **de treksterkte van de buis** en **de Folias-factor (houdt rekening met de bolling van de buis voordat deze bezwijkt bij inwendige druk)** nodig. Dit model (Kiefner & Vieth, 1989) is hoofdzakelijk ontwikkeld voor stalen buizen (in de olie- en gasindustrie) en het is onduidelijk of het model eenvoudig kan worden uitgebreid tot gietijzeren of zelfs nodulair gietijzeren buizen. Bovendien vereist dit model driedimensionale kenmerken van corrosieputten in de buis, die niet gemakkelijk beschikbaar zijn.

Een ander model (Rajani, et al., 2001) dat specifiek gericht is op leidingen van grijs gietijzer vereist een **geometrische factor die afhankelijk is van de afmetingen en de vorm van de corrosieput**, de **laterale afmeting van de corrosieput**, de **geschatte breuktaaiheid**, de **putdiepte**, **constanten** verkregen uit experimentele proeven en **de dikte van de buiswand**. Het is onduidelijk of het model is gevalideerd aan de hand van proeven op grote schaal. Bovendien vereist het model een breuktaaiheid die niet direct beschikbaar is.

Het overzichtsartikel van (St. Clair & Sinha, 2012) geeft een overzicht van modellen die sinds 2001 zijn ontwikkeld.

K Brosse breuk door gebreken aan het oppervlak van ABS kunststof buizen

Concepten van lineaire elastische breukmechanica om scheurgroei onder modus I-belasting voor ABS kunststof buizen te voorspellen zijn ook gemodelleerd (Lu, et al., 2003). Het effect van interne druk, doorbuigingsbelasting onder grond en restspanningen wordt gebruikt om een netto spanningsintensiteitsfactor te berekenen die van invloed is op de langzame scheurgroei en dus op de levensduur. Hiervoor zijn de **radiale diepte van een zwakke plek vanaf het oppervlak van de buisboring** en de **dikte van de buiswand nodig**. Om restspanningsprofielen te bepalen in een geëxtrudeerde buis die alleen aan externe oppervlaktekoeling wordt onderworpen, is bovendien de **afstand tot de buiswand aan de binnenzijde nodig**. Om via het Dugdale-model rekening te houden met puntplasticiteit zijn ook scheurlengtes nodig. De factoren die van invloed zijn op de levensduur van de buis (en dus op de toestand) zijn onder meer de **aanvankelijke grootte van de breuk**, de **breuktaaiheid**, de **weerstand tegen langzame scheurgroei** en de **plaats van de breuk** nodig. De auteurs hebben deze modellen later ook toegepast op PVC.

Het meer recente overzichtsartikel van (Wilson, Filion, & Moore, 2015) geeft ook een samenvatting van fysische modellen die sinds 2001 zijn ontwikkeld.

L Falen van PVC-buizen door inherente gebreken

Een fysisch probabilistisch model voor PVC waarin wordt aangenomen dat onder bedrijfsomstandigheden scheurinitiatie optreedt vanuit inherente defecten in de buiswand bestaat ook (Davis, et al., 2007). De lineaire elastische-breukmechanicatheorie wordt gebruikt om de tijd tot brosse breuk te voorspellen voor buizen met inwendige defecten die worden blootgesteld aan een gecombineerde inwendige druk en bodemvervorming in combinatie met restspanning door de wand. Naast de invoerparameters die als voorbeelden van belasting worden beschouwd, zijn aanvullende parameters nodig. In totaal zijn de **parameters voor de breuktaaiheid op korte termijn**, de **parameters voor langzame scheurgroei**, de **verdeling van de inherente defectgrootte**, de **elasticiteitsmodulus op korte termijn**, de **visco-elastische parameter voor vermindering van de elasticiteitsmodulus**, de **vloeigrens op korte termijn**, de **visco-elastische parameter voor vermindering van de vloeigrens** en de **straal aan de buitenzijde/binnenzijde** nodig. Afhankelijk van de relatie tussen deze twee invoerparameters is de wijze van bezwijken bros of ductiel. Dit model vereist dat men uitgaat van een verdeling van de defecten en moet worden gevalideerd. Het model houdt alleen rekening met breuken door hoepelspanningen. Het in aanmerking nemen van axiale spanningen zou nuttig zijn bij toepassing op buizen met een kleine diameter.

M Falen van asbestcementbuizen door aantasting

Er bestaat ook een fysisch probabilistisch model voor asbestcementbuizen die onderhevig zijn aan aantasting onder gecombineerde interne druk en externe belasting (Davis, et al., 2008). De degradatie van het interne oppervlak vindt plaats via zacht water en de degradatie van het externe oppervlak door contact met een bodemomgeving met een hoog sulfaatgehalte. De belastingen in dit model hebben een fysisch karakter; de opname van de aantasting van de buis heeft een meer statistisch karakter. Daartoe worden verdelingen van de degradatiesnelheid afgeleid op basis van verdelingen van **resterende treksterkte**, **oorspronkelijke treksterkte** en **leeftijd**. Voor dit model moet de **degradatiesnelheid** voor elke buis worden bepaald, wat een invasieve en dure procedure is.

3 Dataintegratie en -kwaliteit

3.1 Dataintegratie

3.1.1 Gewenste data

Het projectteam heeft een matrix met gewenste data geïdentificeerd, op basis van de expertise van de projectpartners. De volgende aspecten zijn meegewogen:

- Mechanisme klassen: categorieën waarin mechanismen geschaard zijn die van invloed kunnen zijn op het al dan niet functioneren van de leiding gedurende zijn levensduur. Voorbeelden zijn verbingsproblemen, AC wanddikteafname en statische/dynamische belastingen op de leidingen.
- Mechanismen: specifieke mechanismen zoals hoekverdraaiing, uitloging, interne waterdruk, verkeersbelasting, etc.
- Modelparameters (doelen): onderliggende oorzaken die de mate van het optredende mechanisme bepalen zoals mechanische en geografische eigenschappen van de leiding, zetting van de bodem, stijfheid van de koppeling, gemiddelde operationele druk, bodemsamenstelling, ligging nabij een weg, etc.
- Verklarende variabelen: specifieke gegevens uit modellen en databronnen die indicatief zijn voor een specifieke parameter of het overkoepelende mechanisme. Daarbij is waar mogelijk ook al een bronsuggestie gegeven.

Voor de bovengenoemde mechanismen, modelparameters en verklarende variabelen is aangegeven of deze noodzakelijk zijn voor de analyse of niet en of deze noodzakelijk zijn als trainingsdata. Zie Bijlage II voor de matrices en bijbehorende legenda. Zie bijlage III voor een lijst met partijen die data kunnen aanleveren.

3.1.2 Beschikbaarheid van data

Gewenste (historische) data benodigd van de eindgebruikers (Bijlage 1) bleek tot in 2022 grotendeels niet beschikbaar te zijn. Dit betreft voornamelijk hydraulische data (spatkrachten, druk en flow) en de chemische samenstelling van het medium welke respectievelijk op de leiding gewerkt heeft of door de leiding gestroomd is. Vaak zijn deze (gemodelleerde) data nog wel voorhanden voor de huidige situatie, maar niet nauwkeurig genoeg terug te rekenen naar maatgevende momenten in het verleden. Veel leidingdata als fabrikant, series, coating, kwaliteit van leidingaanleg, wijze van opvullen en opvulmateriaal, etc. ontbreekt.

Van de grondwaterkwaliteit is geen bron gevonden die data bevat met een afstand tot de leidingen < 10 km. gevonden, voor de analyse een redelijke afstand tot onderzochte leidingen.

Acquaint heeft inspectiedata aangeleverd van zestien drinkwater- en afvalwaterleidingen (tabel 'pipelines') van de eindgebruikers. In de tabel 'joints' zijn observaties als afstand tussen leidingdelen en hoekverdraaiingen vastgelegd. De tabel 'pipe_parts' bevat meetresultaten als ovaliteit, diameter, wanddikte, kromming en uitschuiven in meerdere dimensies. Daarnaast zijn tabellen aangeleverd waarmee genoemde tabellen onderling gekoppeld zijn. Tevens heeft Acquaint de dataset 'grondsoortenkaart' aangeleverd.

HDM heeft open databronnen aangeleverd zoals het nationale wegenbestand, Natura2000-gebieden, stadspolygonen, bodemgebruik en gevaarlijke pijpleidingen.

Spatial Insight heeft deze ook aangeleverd naast een dataset met daarin grondwaterniveaus en de AHN voor de maaiveldhoogtes.

3.1.3 Methoden

De verkregen data is door elke partij in een PostGIS-database geplaatst. Elke tabel is gemarkeerd zodat altijd de herkomst van de data getraceerd kan worden. Vervolgens is de beschikbare data geïntegreerd op het niveau van 'pipe_parts_refined', een bewerking van de tabel 'pipe_parts'.

Alle open data is met een ruimtelijke koppeling, dus op basis van geometrie-velden, met elkaar geïntegreerd. De inspectiedata is grotendeels op basis van id's gekoppeld, de 'joints' zijn ruimtelijk gekoppeld aan de 'pipe_parts_refined'.

Daarbij zijn de volgende buffers gehanteerd rondom het middelpunt van elk individuele pipe_part:

- Joints: binnen 5m
- NWB_wegvlakken: binnen 10m
- Gevaarlijke transportleiding: binnen 10m
- Wegdelen: binnen 10m
- Grondwaterdynamiek (glg/ghg): binnen 200m. Geïnterpoleerd met Inverse Direct Weighting tot de tweede macht (IDW2).
- AHN: binnen 10m. Geïnterpoleerd met IDW.

Zie bijlage IV voor een totaal overzicht van alle koppelingen en de daarvoor gekozen instellingen.

3.2 Analyse van de kwaliteit van en onzekerheden in databronnen

Datakwaliteitsbeoordeling is het proces van het wetenschappelijk en statistisch evalueren van data om te bepalen of ze voldoen aan de vereiste kwaliteit en van het juiste type en de juiste hoeveelheid zijn om het beoogde gebruik te kunnen ondersteunen. Het is een reeks richtlijnen en technieken om data te beschrijven, gegeven een toepassingscontext, en een reeks processen om de kwaliteit van data te beoordelen en te verbeteren.

Datakwaliteitsbeoordeling legt bijvoorbeeld problemen met geografische data bloot waardoor datawetenschappers en GIS-experts plannen kunnen maken voor het opschonen en verrijken van data. Dit wordt meestal gedaan om de integriteit van systemen, kwaliteitsborgingsnormen en nalevingsproblemen te behouden. Over het algemeen zijn technische kwaliteitsproblemen zoals inconsistente structuur- en standaardproblemen, ontbrekende of onbekende standaardgegevens en fouten in de gegevensvelden gemakkelijk te herkennen en te corrigeren, maar complexere problemen moeten worden benaderd met meer gedefinieerde processen.

De datakwaliteitsbeoordeling is uitgevoerd in 3 fasen.

- Planning en voorbereiding
- Implementatie en analyse
- Rapportage

3.2.1 Planning en voorbereiding

Na de data-integratie waren de volgende datacategorieën beschikbaar voor dit project:

1. Operatordata
2. Inline inspectiedata
3. Geografische data

Operatorgegevens zijn data die informatie geeft over de operationele aspecten van de pijpleiding, zoals druk en stroomsnelheden, ontwerpaspecten en materiaalkenmerken. De operatorgegevens in de beschikbare staat zullen worden gebruikt voor de analyse van de datacorrelatie en verdere beoordelingen.

De kwaliteit van de Inspectiedata is al bekend. De variaties zijn gekwantificeerd en behoeven geen verdere analyse.

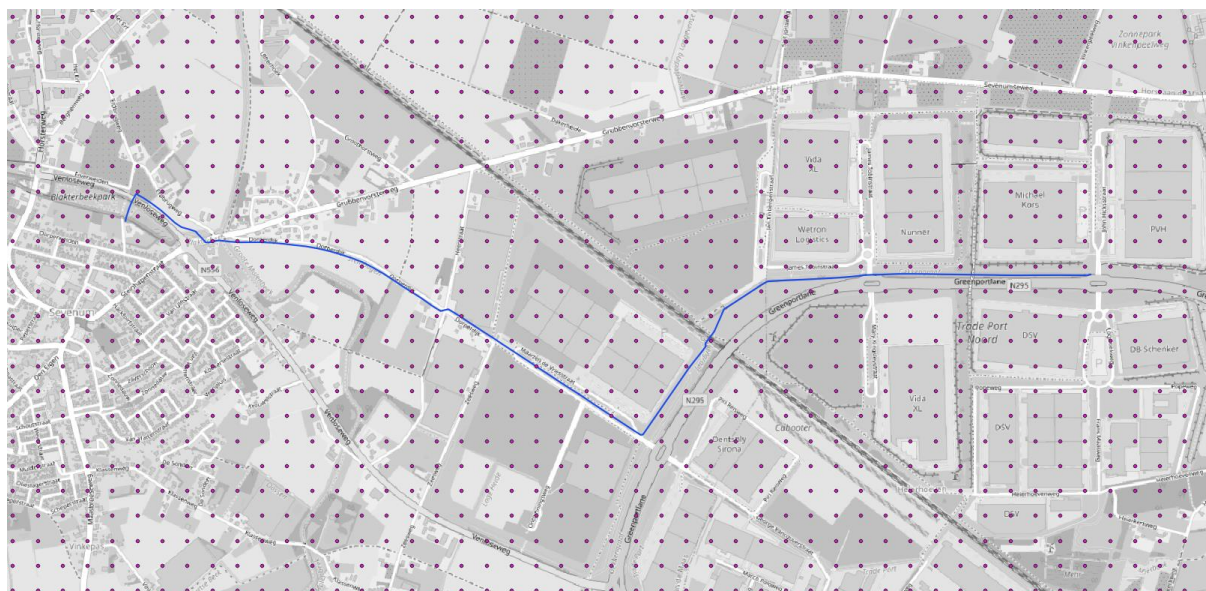
Geografische data zijn verkregen uit verschillende open bronnen en variaties in aspecten als wegen, bruggen, bodemparameters, waterstanden etc. zijn eerder waargenomen. De hoeveelheid data is enorm en een gedisciplineerde aanpak is gevolgd om deze variaties te bestuderen en te kwantificeren. De betrouwbaarheid van deze bronnen is bepaald voordat verdere analyse van de pijpleidingen is uitgevoerd.

3.2.2 Analyse kwaliteit databronnen

Om de onzekerheidspropagatie te modelleren is kennis nodig van onzekerheid (spreiding van observaties) in de gebruikte databronnen. Om deze onzekerheid in te schatten zijn twee analyse methoden geïdentificeerd:

- A. Het onderling vergelijken van verschillende databronnen
- B. Het valideren van databronnen middels validatiemetingen

Methode A is uitgevoerd door rondom leidingen waar onderzoeksgegevens van Acquaint bekend zijn een raster van punten te projecteren (Figuur 2). Aan deze punten is de waarde uit de verschillende vergelijkbare bronnen toegekend.



Figuur 2 gecreëerd raster (roze punten) rondom leiding (blauwe lijn) met inspectiedata

Doormiddel van methode B kan de spreiding in de aanwezige bronnen worden vergeleken met observaties die zijn gedaan volgens gevalideerde observatiemethoden. Als voorbeeld wordt de maaiveldhoogte zoals vastgelegd in de diverse bronnen te toetsen aan gecorrigeerde gps-metingen, een grote nauwkeurigheid. Zo kan bepaald worden in welke mate de ruimtelijke data de maaiveldhoogte beschrijft. Vanwege de kosten worden validatiemetingen slechts uitgevoerd wanneer duidelijke correlatie bestaat tussen specifieke bronnen en de afhankelijke variabele. Op dit moment is nog geen duidelijke correlatie gevonden.

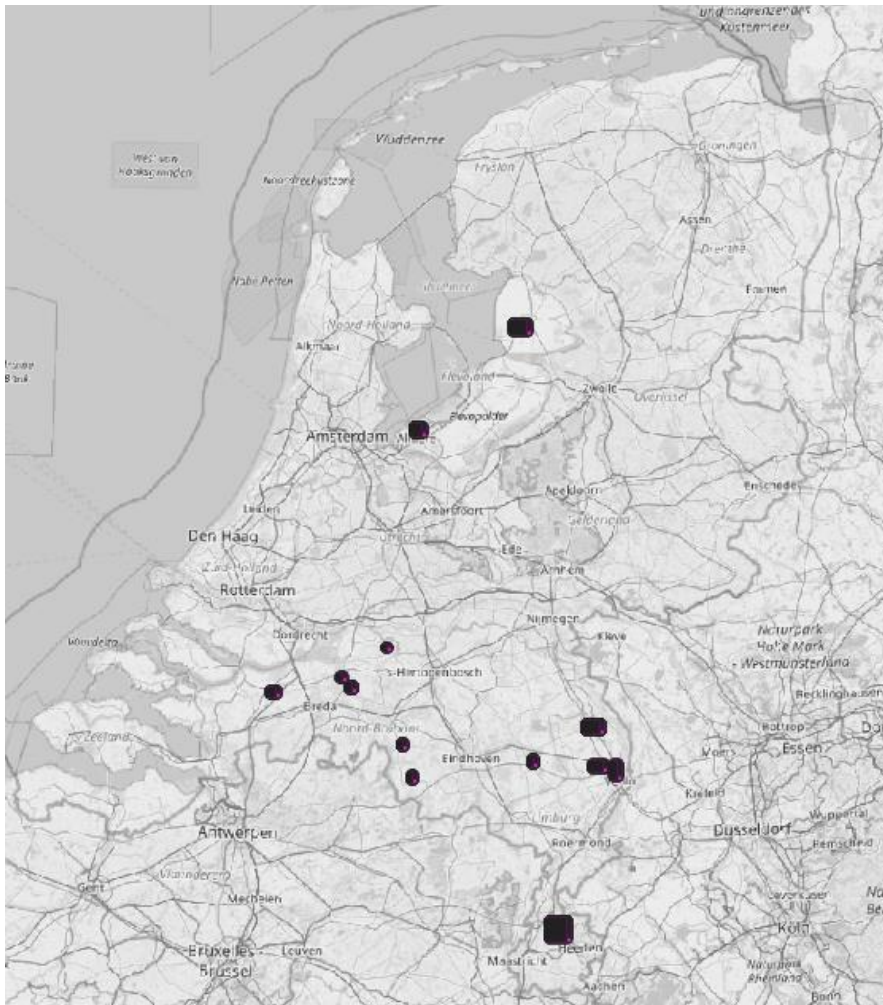
3.2.3 Implementatie en analyse

De datakwaliteitsanalyse is uitgevoerd in drie stappen:

1. Keuze van regio's voor analyse

De gekozen regio's maken deel uit van de secties waar persleidingen zijn aangelegd. Bijkomende overwegingen waren de ligging van de leidingen, de criticiteit van de leidingen, aanwezigheid van inspectiedata van leidingtoestand en het gebruiksgemak.

Figuur 3 geeft de locaties weer voor de analyse van de datakwaliteit. Zo is een dataset gecreëerd met 15108 punten in de ruimte waaraan observaties uit verschillende bronnen zijn gekoppeld. Vervolgens is per punt berekend welke verschillen bestaan tussen de diverse bronnen.



Figuur 3 Overzichtskaart van Nederland met geselecteerde locaties voor datakwaliteit controle

2. Vergelijking tussen gegevensbronnen

Alle categorieën databronnen zijn met elkaar vergeleken en de verschillen zijn gekwantificeerd. Voor de continue zijn de verschillen uitgedrukt in meters. Voor de discrete variabelen is vastgelegd per punt of de verschillende bronnen gelijke waarden hebben. Dit is vastgelegd als Boolean.

3. Kwantificering van variaties

De verschillen zijn weergegeven in histogrammen die zijn opgenomen (Bijlage V). Voor de meeste parameters blijkt een relatief kleine maar niet te verwaarlozen fractie (orde 10%) van de punten een afwijking te vertonen tussen de datasets; in sommige gevallen zijn de verschillen veel groter. Een numeriek overzicht wordt gegeven in Tabel 3 voor discrete variabelen en in Tabel 4 voor continue variabelen. Voor de meeste parameters zijn 2 verschillende bronnen beschikbaar. Let op dat in Tabel 4 80-99% van de gegevens niet beschikbaar zijn, en dat de vergelijking waar mogelijk dus is uitgevoerd op basis van de resterende fractie.

4 Correlatieanalyse

4.1 Inleiding

Bij het inschatten, modelleren of voorspellen van de conditie van een leiding komt het regelmatig voor dat benodigde gegevens, hetzij uit metingen, hetzij uit open bronnen, ontbreken. Deze ontbrekende gegevens kunnen de nauwkeurigheid verminderen of leiden tot valse schattingen van de buisconditie. In dit onderzoek is ons hoofddoel het ontwikkelen van een methode om de ontbrekende gegevens te benaderen met behulp van de bekende gegevens en de bijbehorende correlatie. De eerste stap is het vinden van de meest gecorreleerde informatie; daarna stellen we een efficiënte techniek voor om de ontbrekende waarde in de datasets aan te vullen. Ons algoritme selecteert automatisch de sterk gecorreleerde attributen. Na de selectie van sterk gecorreleerde kenmerkenparen wordt een clustering- of regressiemodel gebouwd om uiteindelijk ontbrekende datasets te genereren.

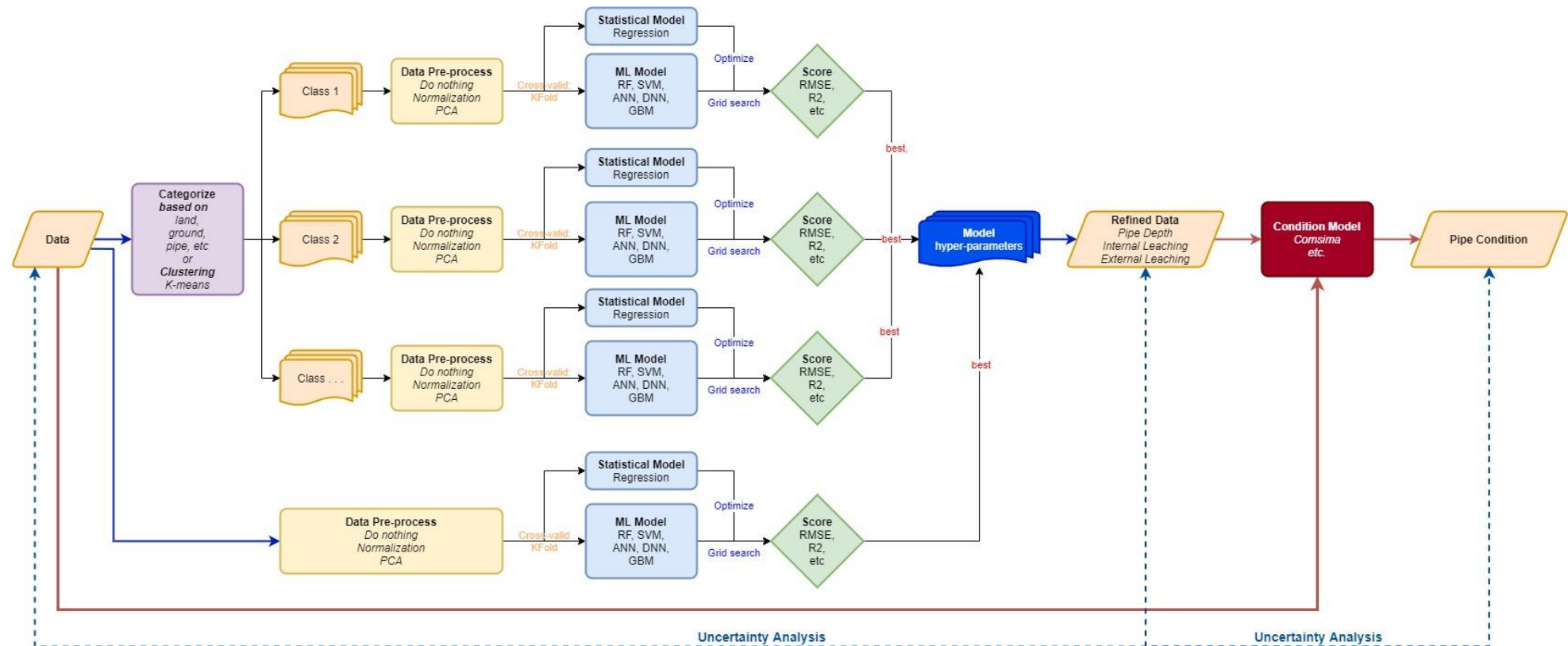
4.2 Methodologie

4.2.1 Stroomschema en stappen

Het onderstaande stroomschema (Figuur 4) toont de door ons voorgestelde stappen om de gezochte correlaties vast te stellen. In het geval er in de dataset geen waarden ontbreken, kan deze direct in het conditiemodel worden toegepast (rode pijl). Dit hoofdstuk richt zich specifiek op de situatie waarin dit niet het geval is. Het stroomschema bevat hiertoe twee parallelle oplossingen (donkerblauwe pijlen). De eenvoudige oplossing is het gebruik van alle beschikbare datasets om de correlaties te vinden, door het ML-model rechtstreeks te voeden met alle datasets.

De ingewikkeldere oplossing bestaat uit twee stappen. Om te beginnen worden datasets in subcategorieën opgesplitst op basis van expertkennis en datatype. Vervolgens wordt een gridsearch toegepast om het beste ML-model en de best passende hyperparameters¹ te selecteren. Dit proces bestaat uit drie stappen: gegevens voorbereiden, raster zoeken en prestatiescore. In de voorbereidingsstap wordt de ruwe dataset gefilterd op de ontbrekende gegevensrecords. Na het opschonen van de ruwe gegevens wordt normalisatie of dimensionaliteitsreductie toegepast om de gegevenskwaliteit voor modellen voor machinaal leren te verbeteren. De tweede stap is het gebruik van gridsearch. Om de meeste algoritmen voor machinaal leren en de bijbehorende hyperparameters op te nemen. Wij ontwierpen een nieuwe gegevensstructuur om al deze informatie op te slaan. De laatste stap is het controleren van de prestatiescores, en het selecteren van de beste om op te slaan. Voor elke subcategorie wordt een correlatiematrix berekend. Deze correlatiematrix leidt tot een beter begrip van de correlatie onder bepaalde omstandigheden.

¹ [https://en.wikipedia.org/wiki/Hyperparameter_\(machine_learning\)](https://en.wikipedia.org/wiki/Hyperparameter_(machine_learning))



Figuur 4: Stroomdiagram van het proces. De twee blokken aan de uiterste rechterkant vallen buiten het in dit hoofdstuk beschreven proces.

De door ons ontwikkelde methodologie is gebaseerd op domeinkennis over de toestand van de buis, ervaring met datascience en ideeën uit publicaties. Figuur 5 illustreert de primaire resultaten van elke stap. De eerste stap van de methodologie is de analyse van de volledige datasets. Om de correlatiematrix te genereren, passen wij de rangcorrelatiecoëfficiënt van Spearman toe. Deze correlatiecoëfficiënt wordt hoofdzakelijk gebruikt voor datasets met continue en discrete gegevens. De hoogste stijgende correlatie wordt weergegeven door 1, -1 staat voor de hoogste dalende trend.

Uit de Spearman-matrix links blijkt dat de correlatie niet sterk genoeg is om een goed model voor machinaal leren te bouwen. Daarom gebruikten wij de methode van datasplitsing om de correlatie van gesubcategoriseerde datasets te controleren. Uit de kleine matrices blijkt dat sommige kenmerken een hogere correlatie hebben. Na de correlatiematrix onderzochten wij het ruimtelijke distributiepatroon van datasets met subcategorieën. Het doel van deze stap is de gegevens te visualiseren, en deskundig oordeel toe te passen om na te gaan waar er patronen of trends zijn die kunnen worden gebruikt om een representatiemodel te bouwen om de correlatie te beschrijven.

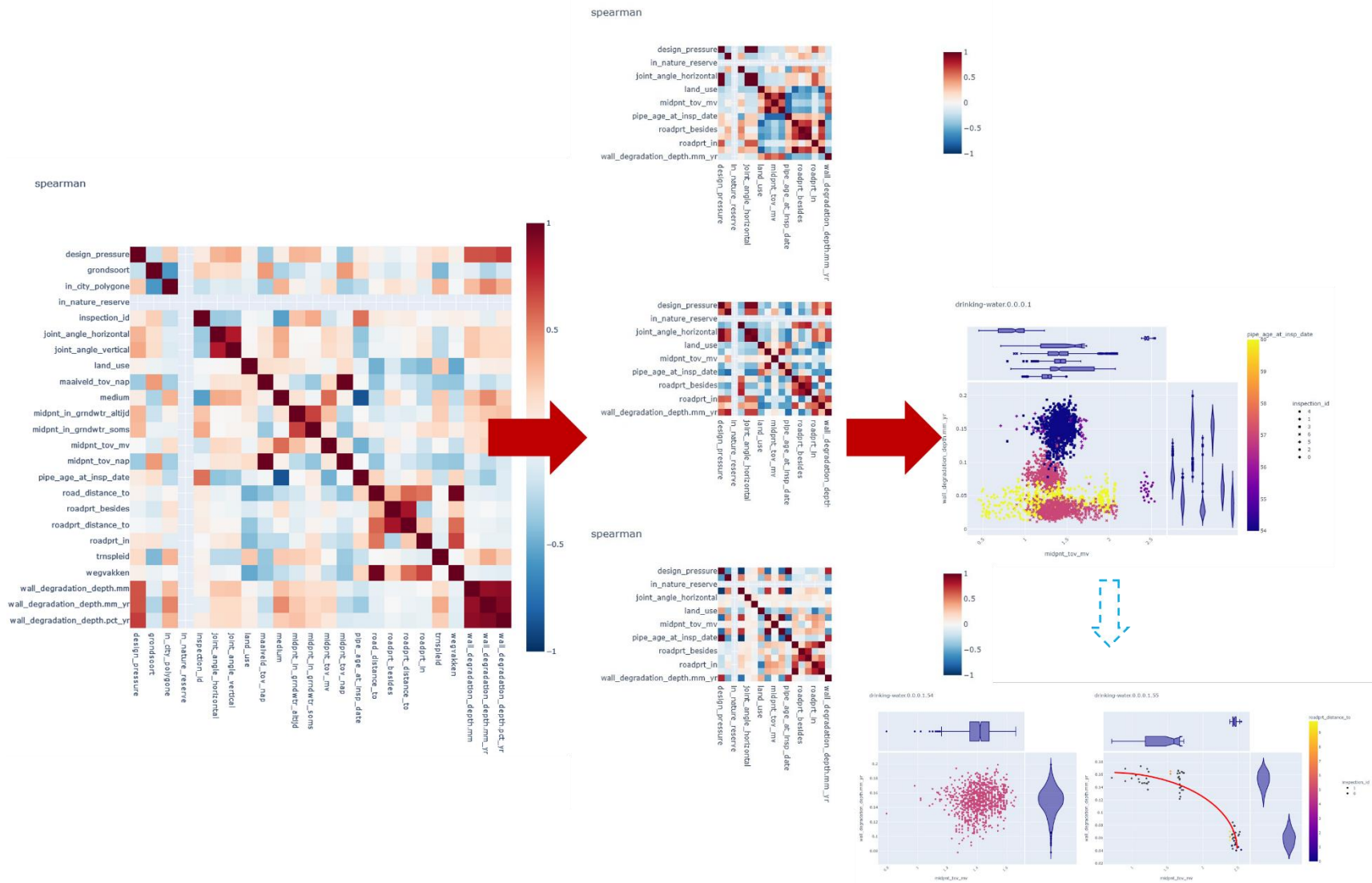
Samengevat zijn de drie achtereenvolgende stappen:

1. Correlatieanalyse:
 - a. Spearman's rangcorrelatiecoëfficiënt algemene beschrijving.
 - b. Structuur van de analyse van de volledige gegevensverzameling.
 - c. Uitleg over de reden en de methode om dieper in te gaan op categorisering.
2. Modellering (op weg naar verfijning met regressie of clustering):
 - a. Specifiek model a, regressie.
 - b. Specifiek model b, clustering.
3. Selectie van modellen:
 - a. Visualiseren van relaties tussen kenmerken.
 - b. Visie op het automatisch selecteren van de modelbenadering voor een gegeven set kenmerken en het toepassen ervan .

In het Midas-project is in praktische zin met name aan stap 1 gewerkt. De resultaten die hieronder worden gerapporteerd zijn uit deze stap verkregen.

4.2.2 Gebruikte programmeeromgeving

In dit onderzoek gebruiken we Python 3.8 als belangrijkste computertaal, een interactieve visualisatieplotbibliotheek om alle cijfers te genereren en scikit-learn. De laatstgenoemde bibliotheek is de meest gebruikte bibliotheek voor data-analyse en data-wetenschappelijke taken. Zij bevat functies voor gegevensvoorbewerking, modelselectie en prestatiematrix.



Figuur 5: Demonstratie van de correlatieanalyse en de verfijningsmodellijn

4.3 Resultaten

Deze paragraaf geeft een kort overzicht van de uit de correlatieanalyse en verfijning daarvan behaalde resultaten. Voor meer gedetailleerde resultaten wordt de lezer verwezen naar Bijlage VII.

De resultaten laten zich als volgt samenvatten:

- Er is slechts een enkel kenmerkenpaar met een sterke correlatie gevonden voor de volledige dataset, namelijk *maaiveld t.o.v. NAP* tegenover *buisdiepte t.o.v. NAP*. Dit is logisch is bij een min of meer constante buisdiepte en twee ordegroottes (grotere variatie in maaiveldhoogte).
- Voor subcategorieën zijn correlaties tussen diverse parameterparen gevonden. Deze variëren sterk van subcategorie tot subcategorie en het is niet uit te sluiten dat deze correlaties “toevallig” zijn en bij grotere waarnemingsgetallen verdwijnen. Vooralsnog zijn zij daarom als niet bruikbaar beoordeeld voor het gebruik van surrogaatparameters voor ontbrekende gegevens.
- De verfijnde analyse levert als aanvullend inzicht dat correlaties vaak voort lijken te komen uit een beperkt aantal clusters van metingen (die segmenten van dezelfde leiding representeren) waardoor eenvoudig een rechte lijn te trekken is. Het vergroten van de basis aan inspectiegegevens is nodig om tot betrouwbare en daarmee bruikbare resultaten te komen in deze analyse.

4.4 Conclusies en aanbevelingen

De conclusies van de analyses uit dit hoofdstuk zijn dat er geen sterke correlatie is met de aantasting van leidingen en geen bruikbare correlaties tussen omgevingsparameters. Daarmee zijn geen surrogaatparameters geïdentificeerd voor het aanvullen van ontbrekende data; dit is in de rest van dit rapport dus ook niet gedaan.

Wij bevelen ten eerste aan dat er meer inspecties, operationele gegevens en nieuwe datasets met betrekking tot de interne aantasting van de buis komen. Deze nieuwe informatie zou nuttig kunnen zijn om de werkelijke correlaties te vinden en de verfijningsmodellen verder te verbeteren.

5 Kwantificering en propagatie van de onzekerheid

Een van de opbrengsten van dit project is een visueel dashboard dat de asset manager helpt bij het nemen van passende maatregelen - vooral met betrekking tot het verzamelen van gegevens. In Na een algemene introductie van onzekerheid in paragraaf 5.2, worden in paragraaf 5.3 en 5.4 respectievelijk de aspecten van onzekerheidsvoortplanting en onzekerheidsanalyse toegelicht. Dit dient als opstap naar paragraaf 5.5 waarin de functionaliteiten en het gebruik van het dashboard worden toegelicht aan de hand van illustratieve voorbeelden.

5.1 Hoe de eindgebruikers onder operationele omstandigheden gebruik kunnen maken van het Midas dashboard

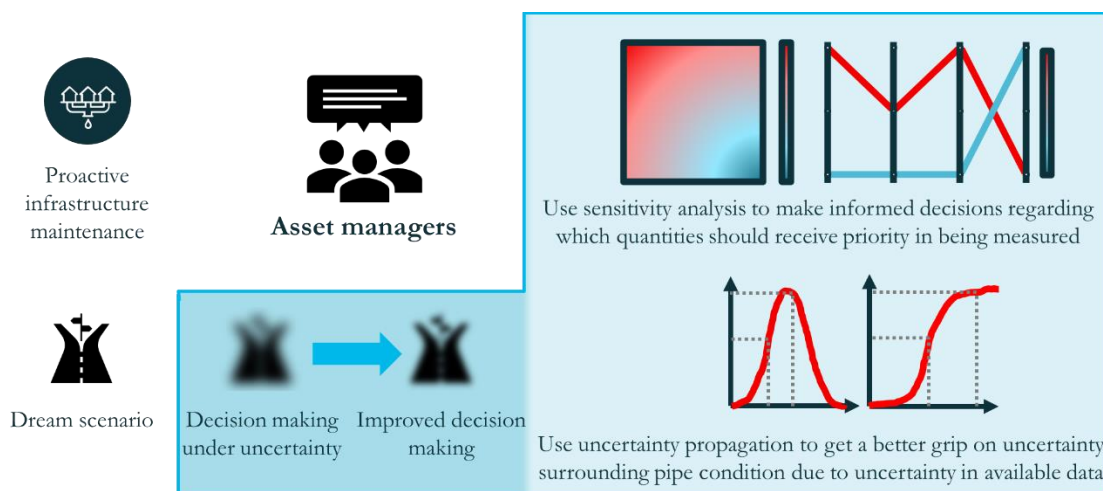
Een belangrijk resultaat van dit project is het Midas-dashboard dat gebruik maakt van gegevens die door de eindgebruikers zijn verstrekt. Asset managers moeten belangrijke beslissingen nemen onder verschillende onzekerheden rond de assets. De kosten van de assets zijn hoog en de besluitvorming rond deze assets speelt een centrale rol in de financiën van het bedrijf. Een van de belangrijkste taken van een asset manager is het proactief onderhouden van de infrastructuur waarvoor hij verantwoordelijk is. Er bestaan instrumenten zoals COMSIMA om de asset manager te helpen bij het nemen van beter gefundeerde beslissingen. Het zou ideaal zijn als een asset manager beslissingen zou kunnen nemen zonder enige ambiguïteit. In werkelijkheid zijn er echter verschillende onzekerheden die besluitvorming uitdagend maken. Het Midas-dashboard is ontwikkeld om de assetbeheerder te helpen beter geïnformeerde beslissingen te nemen in onzekere omstandigheden.

De resultaten op het dashboard zijn op COMSIMA simulaties gebaseerd. De uitkomst van de simulaties is een faalkans of conditiegetal. Een lagere waarde hiervan betekent betere conditie. Het dashboard bestaat uit twee componenten. De eerste component is die van de gevoeligheidsanalyse. De gevoeligheidsanalyse kan de beheerder helpen de algemene toestand van het leidingnet te beoordelen en beter geïnformeerde beslissingen te nemen over welke parameters bij voorkeur moeten worden gemeten. De tweede component van het dashboard is die van de voortplanting van onzekerheid. De onzekerheidsvoortplanting kan de asset manager helpen een beter beeld te krijgen van de onzekerheden in de door COMSIMA geschatte conditie als gevolg van de onzekerheden in de inputs. De begeleidende instructievideo over het dashboard bevat meer informatie over hoe de eindgebruiker het dashboard kan gebruiken om tot beslissingen te komen. Het dashboard dient als sjabloon. De eindgebruikers kunnen hun eigen COMSIMA-simulaties uitvoeren om soortgelijke plots te genereren en zo hun besluitvorming te verbeteren. Oftewel het beeld scherper krijgen (zie Figuur 6).

5.2 Inleiding in onzekerheid

Onzekerheid is een fenomeen dat de besluitvorming in tal van aspecten van ons leven - persoonlijk of professioneel - beïnvloedt. Als beheerder van drinkwaterinfrastructuur moeten beslissingen worden genomen ondanks een gebrek aan informatie. Soms is nauwkeurige informatie beschikbaar, maar vaker nog moet een assetmanager vertrouwen op ervaring en gefundeerde gissingen.

Volgens de NASA Technical Standard on Models and Simulations is "onzekerheid een brede en algemene term die wordt gebruikt om een onvolmaakte stand van kennis of een variabiliteit te beschrijven die het gevolg is van een verscheidenheid van factoren, waaronder, maar niet beperkt tot, gebrek aan kennis, toepasbaarheid van informatie, fysieke variatie, willekeur of stochastisch gedrag, onbepaaldheid, beoordeling en benadering".



Figuur 6: De toegevoegde waarde van het Midas dashboard aan de eindgebruikers – ondersteunen bij beslissingsvorming onder operationele omstandigheden.

De oorsprong van onzekerheid kan grofweg worden ingedeeld in de volgende twee categorieën:

- **Aleatorische onzekerheid:** Geassocieerd met de natuurlijke variabiliteit in de waarnemingen. Deze is typisch inherent aan het systeem en kan niet worden verminderd (“variatie” in normaal spraakgebruik).

Laten wij eens kijken naar het voorspellingsmodel voor leidingfalen, COMSIMA. Momenteel kan het model netwerken bestuderen die een statische belasting ondervinden. Momenteel wordt de interne druk in de leidingen tijdens de simulaties constant gehouden. In werkelijkheid zal de interne druk echter in ruimte en tijd fluctueren als gevolg van de dynamische en complexe aard van de waterleidingnetwerken en de vraag naar water. Dit geldt als een aleatorische onzekerheid bij het evalueren van simulaties in COMSIMA.

- **Epistemische onzekerheid:** Wordt geassocieerd met een gebrek aan kennis en kan worden verminderd door meer of betere gegevens te verkrijgen. Een dergelijke onzekerheid houdt meestal verband met meetfouten, gebrekkige schattingen of gewoon een gebrek aan kennis (“betrouwbaarheid” in normaal spraakgebruik).

Een voorbeeld hiervan is hoe de wanddikte in de verschillende versies van COMSIMA is verwerkt. In de eerste versie van COMSIMA werd uitgegaan van een constante waarde van de wanddikte op basis van geregistreeerde gegevens. Het is echter bekend dat de wanden van asbestcementbuizen doorgaans in de loop der tijd door uitloging degraderen. Dit aspect is later in de tweede versie van COMSIMA aangepakt met een verouderingsmodel. Aangezien er echter een gebrek aan kennis is over de werkelijke degradatieniveaus, moeten voor het verouderingsmodel bepaalde veronderstellingen worden gemaakt. Veldmetingen met pulse-echo kunnen dienen als een manier om deze epistemische onzekerheid te verminderen bij het verstrekken van invoergegevens voor COMSIMA.

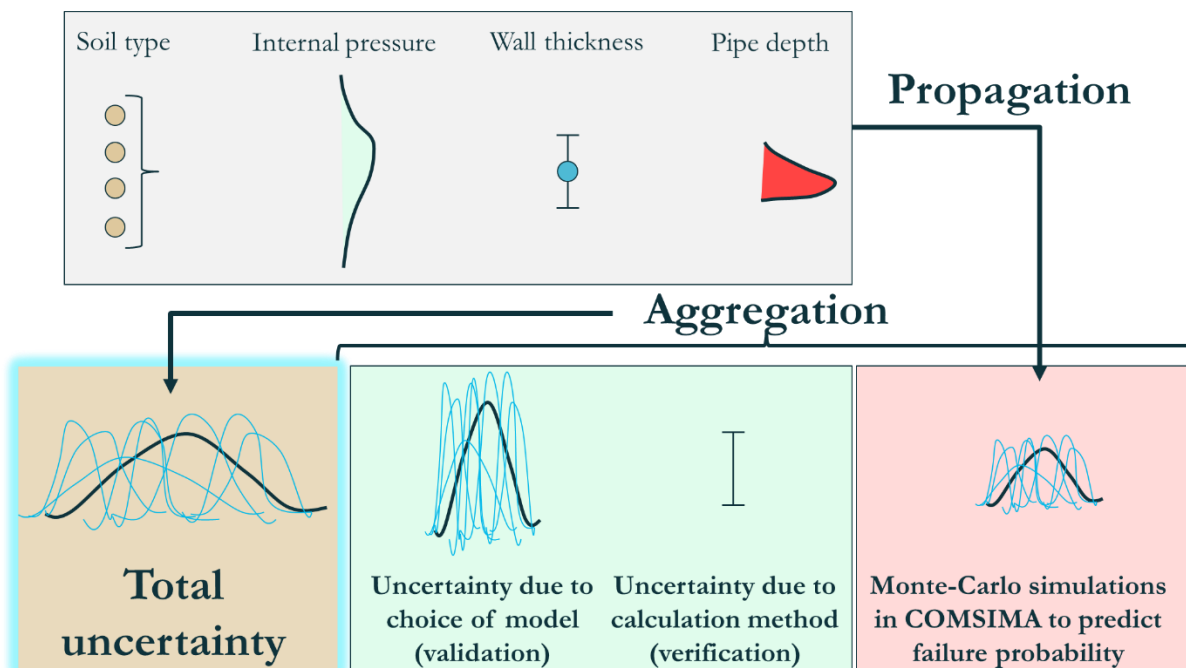
5.3 Onzekerheidspropagatie

5.3.1 Algemeen overzicht van onzekerheidspropagatie

Onzekerheidskwantificering is het proces van het karakteriseren van alle significante onzekerheden in een model, simulatie of experiment en van het kwantificeren van hun effect op de berekende of experimentele resultaten. Met behulp van Figuur 7 illustreren we hoe het kader voor onzekerheidskwantificering in de context van het onderhavige project zou werken.

Laten wij aannemen dat het door COMSIMA voorspelde falen van buizen uitsluitend afhankelijk is van het bodemtype, de inwendige druk, de wanddikte en de buisdiepte. Wanneer we het bodemtype als categorisch beschouwen, zal de waarde discreet van aard zijn (bijvoorbeeld "zand", "veen", "klei"). Voor variabelen als interne druk en buisdiepte kunnen de onzekerheden echter

de vorm aannemen van kansverdelingsfuncties (gebaseerd op sensormetingen van dagelijkse fluctuaties of geregistreeerde gegevens). Voor een grootheid als wanddikte kunnen dergelijke verdelingsfuncties niet triviaal worden gegenereerd vanwege een gebrek aan metingen, en moet een bereik worden aangenomen. Deze vier voorbeelden worden geïllustreerd in het grijze vak in Figuur 7.



Figuur 7: Kader voor de propagatie van onzekerheid. Onzekerheden in de inputparameters worden doorgegeven aan de uitkomsten van COMSIMA. Onzekerheden in verband met de berekeningstechnieken en de keuze van fysisch-mechanische modellen worden samen met de onzekerheden die naar de uitkomsten worden overgedragen, samengevoegd om de totale onzekerheid te kwantificeren. Zie de tekst voor een voorbeeld.

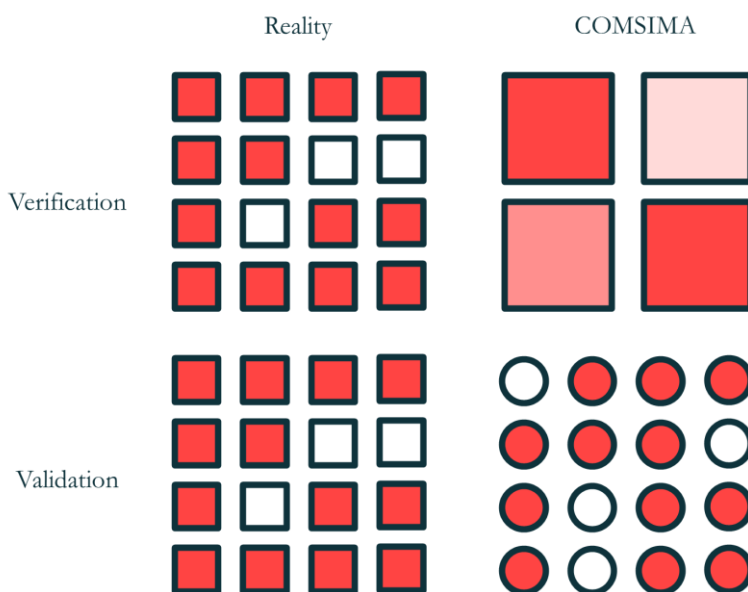
Elke combinatie van invoerparameters zal leiden tot een andere uitvoer. Zo kan COMSIMA worden uitgevoerd voor een bodemtype "zand", een wanddikte van "8 mm" en een buisdiepte van "1,2 m", terwijl de interne druk varieert tussen "2 en 6 bar" (zeg maar 100 simulaties). De variatie van de inwendige druk moet overeenkomen met de waarschijnlijkheidsverdelingsfunctie ervan. Als bijvoorbeeld het gemiddelde en de standaardafwijking van de interne druk respectievelijk 3 en 0,6 bar is, moet dit ook gelden voor de waarden die in de 100 simulaties worden gebruikt (het gemiddelde en de standaardafwijking van de interne druk die voor de 100 simulaties worden gebruikt, moeten respectievelijk ~3 en ~0,6 bar zijn).

Deze 100 simulaties met verschillende inwendige drukwaarden zullen leiden tot een variatie in de faalkans van een buis. Zo leidt een interne druk van 4 bar tot een bezwijkpercentage van 0,2, terwijl een interne druk van 5 bar tot een bezwijkpercentage van 0,5 leidt. Met de uitkomsten van de 100 simulaties kan een kansverdelingsfunctie worden gegenereerd (blauwe lijn in het rode kader). Dit is onzekerheidsvoortplanting. Soortgelijke simulaties waarbij bijvoorbeeld de wanddikte wordt gevarieerd, eerder via onzekerheidspropagatie tot een andere kansverdelingsfunctie van de uitkomsten (opnieuw een blauwe lijn in het rode kader). Op basis van meerdere simulaties kan een globale onzekerheidsverdelingsfunctie voor de uitkomst worden gegenereerd (zwarte lijn in het rode kader).

Hiernaast kunnen twee andere aspecten een rol spelen - onzekerheden in verband met de keuze van het model en de methode om het model toe te passen. De eerste onzekerheid houdt verband met de verificatie - worden de in COMSIMA opgenomen vergelijkingen nauwkeurig opgelost? De tweede onzekerheid heeft te maken met validatie - zijn de juiste vergelijkingen gekozen om de bijbehorende fysische mechanismen te onderzoeken? Dit wordt geïllustreerd in het groene kader van Figuur 7.

Een poging om dit op analoge wijze uit te leggen wordt geïllustreerd in Figuur 8. Laten we aannemen dat de werkelijkheid wordt uitgedrukt met zestien vierkanten met een mengeling van rood en wit. Wegens beperkingen in de middelen (bijvoorbeeld om de

omtrek te beperken) kunnen echter slechts vier vierkanten realistisch in COMSIMA worden geïmplementeerd. Dit resulteert in een verdunning van de kleuren van de vierkanten en leidt tot een onzekerheid over het werkelijke beeld. Deze onzekerheid hangt samen met de verificatie. Als vierkanten echter te ingewikkeld zijn om te implementeren (bijvoorbeeld vanwege hun grotere oppervlakte), kan ervoor worden gekozen de werkelijkheid in COMSIMA te bestuderen met zestien cirkels - wat een (iets) ander kader betekent. Dit hangt typisch samen met de aannames en beperkingen van de gekozen modellen. In deze specifieke illustratie is de kleur van de cirkels anders dan de corresponderende vierkanten vanwege de onzekerheden die voortvloeien uit het gekozen model. Deze onzekerheid hangt samen met validatie - het vergelijken van de modelresultaten met experimenten in een gecontroleerde omgeving. Voor de evaluatie van het buigmoment als gevolg van het verkeer moet bijvoorbeeld een vergelijking worden geïntegreerd. Hiervoor moet een vergelijking op een rooster worden opgelost. De kwestie van verificatie beantwoordt de vraag of het raster fijn genoeg is om de door het verkeer veroorzaakte spanningen nauwkeurig weer te geven. Terwijl het verkeer dynamisch van aard is, wordt het in COMSIMA op een statische manier geïntegreerd op basis van vermoeiingsmodellen. Bij de validatie moet de vraag worden beantwoord of de keuze van een dergelijk model al dan niet aanvaardbaar is.



Figuur 8: Een poging om verificatie en validatie uit te leggen aan de hand van analogieën. Zie tekst voor uitleg.

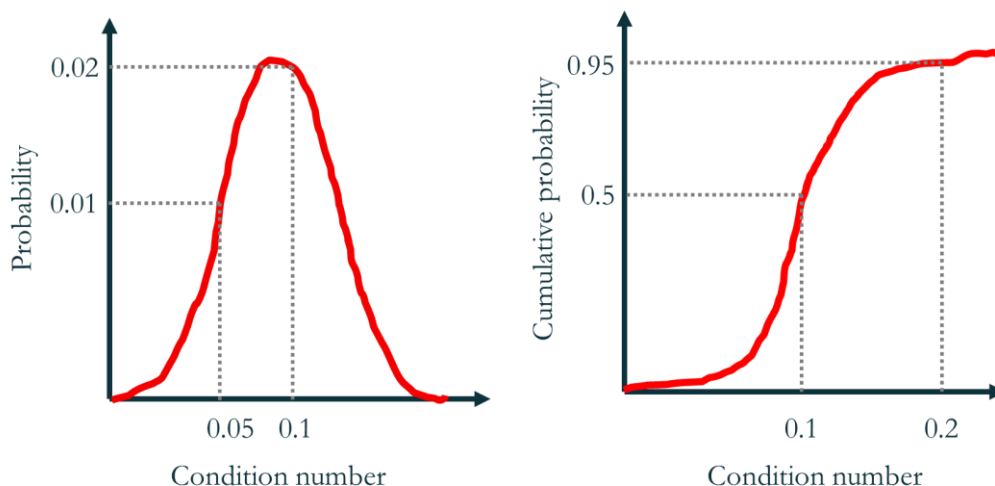
De onzekerheid in de output die voortkomt uit de input, kan samen met de onzekerheden in verband met de verificatie en validatie worden samengevoegd tot een resulterende onzekerheid (bruin vak in Figuur 7) - die van belang is in het besluitvormingsproces. Deze hele keten wordt voorwaartse onzekerheidspropagatie genoemd.

Een soortgelijk proces kan in omgekeerde richting worden uitgevoerd, wat doorgaans inverse onzekerheidspropagatie wordt genoemd. Dit proces beantwoordt de vraag: gegeven een output, wat had de inputparameter kunnen zijn? Dit wordt hier niet verder besproken omdat het buiten het bestek van dit project valt.

De onzekerheid in de inputparameters plant zich voort in de output. Om de verdeling van onzekerheden (in de invoer of de uitvoer) weer te geven, worden typisch twee soorten afspiegelingen van het onderliggende stochastische gedrag gebruikt - de kansverdelingsfunctie (pdf) en de cumulatieve verdelingsfunctie (cdf, zie Figuur 9). Een kansverdelingsfunctie kan worden gezien als een verfijning van het histogram, door de tellingen te vervangen door een waarschijnlijkheid. De cumulatieve verdelingsfunctie is afgeleid van de kansverdelingsfunctie als een voortschrijdende som van alle kansen.

De kansverdelingsfunctie vertelt "wat de kans is om een bepaalde situatie waar te nemen". In het voorbeeld, "de kans om een conditiegetal van 0,05 te hebben is 0,01 (of 1%)". Evenzo "is de kans op een toestand van 0,1 0,02 (of 2%)". Een kansverdelingsfunctie geeft een beeld van welke situatie het meest waarschijnlijk is en wat het bijbehorende verspreiding zou kunnen zijn.

De absolute waarden van de kans in een kansverdelingsfunctie zullen gevoelig zijn voor hoe fijn de categorieën van het histogram zijn. Om deze gevoeligheid te vermijden, kan de cumulatieve verdelingsfunctie ook nuttig zijn. In het voorbeeld is te zien dat "de kans om conditiegetal $\leq 0,1$ te hebben 0,5 (of 50%) is". Evenzo is "de kans om een conditiegetal van $\leq 0,2$ te hebben 0,95 (of 95%)". Kortom, de cumulatieve verdelingsfunctie beantwoordt de vraag "wat is de kans dat de variabele een bepaalde waarde niet overschrijdt".



Figuur 9: Schematische weergave van de kansverdelingsfunctie (links) en de cumulatieve verdelingsfunctie (rechts)

5.4 Gevoeligheidsanalyse

De belangrijkste vraag die een gevoeligheidsanalyse tracht te beantwoorden is "Hoe kan men bepalen welke input de output het meest beïnvloeden?". Het kan ook worden opgevat als een "wat-als-analyse" (Kenton, 2022). In een paar blogs worden de begrippen onzekerheid in lekentaal uitgelegd (Groen, 2016; Guo, 2020).

5.4.1 Types gevoeligheidsanalyses

Lokale gevoeligheidsanalyse: Een lokale analyse is een analyse waarbij het effect van variaties van één inputparameter (terwijl alle andere inputparameters ongewijzigd blijven) op de output wordt onderzocht. Deze techniek kan worden toegepast wanneer de beschikbare gegevens een "puntwaarde" zijn (een enkele waarde die het gegevenspunt vertegenwoordigt).

Als voorbeeld nemen wij COMSIMA-simulaties waarin de bezwijkgraad van de leiding alleen gebaseerd is op grondsoort, diepte onder de grond, wanddikte en inwendige druk. Bij wijze van voorbeeld ziet een lokale gevoeligheidsanalyse er als volgt uit. Aangenomen wordt dat de volgende kennis met een hoge mate van betrouwbaarheid beschikbaar is: de grondsoort is zand, de diepte onder de grond is 1 m, de inwendige druk is 3 bar. Er bestaat echter onzekerheid over de werkelijke waarde van de wanddikte. Op basis van een onderbuikgevoel wordt verwacht dat deze 8 mm zal bedragen. Vervolgens worden verschillende COMSIMA-simulaties uitgevoerd waarbij de wanddikte systematisch wordt gevarieerd van 2 mm tot 8 mm in stappen van 0,25 mm. Zo kan de lokale onzekerheid ten gevolge van de wanddikte worden beoordeeld. Op basis van de resultaten hiervan kan worden besloten of het drinkwaterbedrijf enkele veldmetingen moet uitvoeren om het onderbuikgevoel te verifiëren.

Screeninganalyse: Een screeningsanalyse kwantificeert het effect op de output wanneer een inputparameter wordt gewijzigd volgens het onzekerheidsbereik van een inputparameter. Zij kan worden beschouwd als een uitbreiding van de lokale gevoeligheidsanalyse. Het verschil hier is dat enige informatie over de onzekerheid in de inputparameter beschikbaar is.

Als uitbreiding op het voorbeeld van de "lokale gevoeligheidsanalyse", veronderstellen wij dat een gefundeerde schatting kan worden gemaakt van de wanddikte (bijvoorbeeld op basis van een historische database van wanddiktes van buizen). De geschatte wanddikte zal naar verwachting tussen 6 en 7 mm liggen. Er kunnen COMSIMA-simulaties worden uitgevoerd waarbij de

wanddikte systematisch wordt gevarieerd van 6 tot 7 mm in stappen van 0,1 mm. Als de variaties in wanddikte niet veel variatie in faalkans opleveren (bijvoorbeeld <5%), kan een dergelijke mate van onzekerheid aanvaardbaar zijn en hoeft het drinkwaterbedrijf geen veldmetingen uit te voeren.

Globale gevoeligheidsanalyse: Een globale gevoeligheidsanalyse kan worden gezien als een uitbreiding van de onzekerheidsvoortplanting. In tegenstelling tot de lokale analyse worden alle parameters gelijktijdig gevarieerd. Het doel is de verschillende inputparameters te rangschikken naar hun bijdrage aan de totale onzekerheid in de output. Een algemene eis is echter dat de volledige kennis van de ingangsonzekerheden vooraf bekend moet zijn (bijvoorbeeld als verdelingsfuncties). De zogenaamde Sobol-indices kunnen samenvatten welke inputparameters het meest bijdragen aan de variatie in de output. Het vorige voorbeeld van COMSIMA-simulaties wordt opnieuw bekeken. Aangenomen wordt dat voor elk van de relevante invoerparameters (bodemtype, ligging van de buis onder de grond, inwendige druk, wanddikte) een kansverdelingsfunctie voor de onzekerheden (onafhankelijk van elkaar) beschikbaar is. Er worden verschillende COMSIMA-simulaties uitgevoerd waarin alle parameters tegelijkertijd worden gevarieerd. Op basis van alle simulaties wordt de totale onzekerheid bepaald en blijkt dat de wanddikte de hoogste Sobol-index heeft (dus de meest "schuldige" parameter). In dat geval zou een uitgebreide meetcampagne voorrang moeten krijgen om de wanddikte te kwantificeren boven de andere invoerparameters.

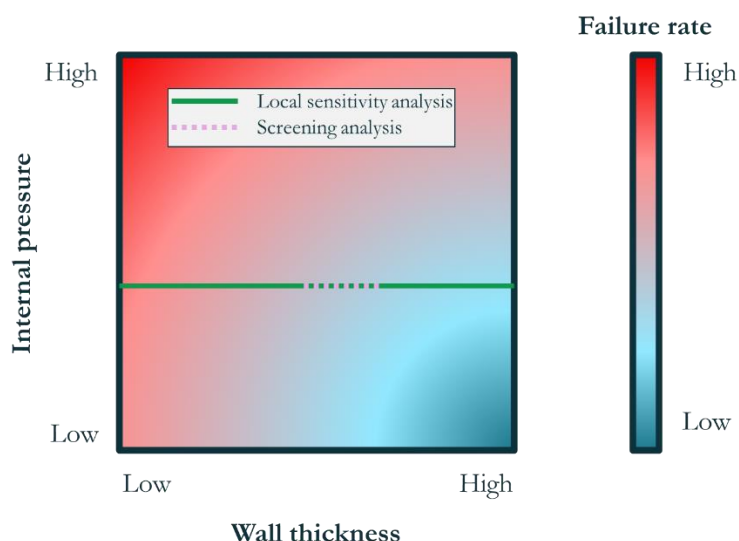
5.4.2 Visualisatie van de gevoeligheidsanalyse op het Midas-dashboard ²

In het als onderdeel van het Midas-project gemaakte dashboard (<https://kwrmidasapp.herokuapp.com/>) worden twee soorten plots gebruikt om de gevoeligheid van de output voor de keuze van de inputparameters te visualiseren - een heatmap en een plot met parallelle coördinaten. De heatmap is couranter en gemakkelijker te interpreteren, zelfs voor een ongehoofd oog. De plot met parallelle coördinaten is minder intuïtief en wordt vooral gebruikt wanneer er meerdere parameters in het spel zijn. Een heatmap is beperkt tot twee parameters.

Heatmap: Een heatmap is een plot met twee assen die de twee parameters bevat waarvan men het effect wil bestuderen. Voorwaarde daarbij is dus dat de overige parameters constant worden gehouden.

Het onderstaande schematische voorbeeld (Figuur 10) toont een heatmap tussen de inwendige druk en de wanddikte. Te zien is dat een hoge inwendige druk en een geringe wanddikte hogere kansen op buisbreuk oplevert en omgekeerd. Met een dergelijke heatmap kan een lokale gevoeligheids- of screeningsanalyse worden uitgevoerd. Zo is de ononderbroken groene lijn een voorbeeld van een lokale gevoeligheidsanalyse, terwijl de gestippelde magenta lijn een voorbeeld is van een screeninganalyse voor de wanddikte. Merk op dat deze twee lijnen horizontaal zijn, wat betekent dat de inwendige druk constant is. In dit specifieke voorbeeld suggereert de lokale gevoeligheids- of screeninganalyse dat voor de vaste waarden bodemtype (zeg zand), buisdiepte (zeg 0,8 m) en inwendige druk (zeg 4 bar) de bezwijkpercentages niet al te gevoelig zijn voor schommelingen in de wanddikte. Volgens de in paragraaf 5.4.1 besproken redenering betekent dit voor dit specifieke geval dat de exacte kennis van de wanddikte dus niet zo belangrijk is en een veldmeetcampagne wellicht niet veel toegevoegde waarde heeft. Met een dergelijke heatmap kan de globale gevoeligheidsanalyse echter niet worden gevisualiseerd, tenzij slechts twee parameters verantwoordelijk zijn voor onzekerheden in de output.

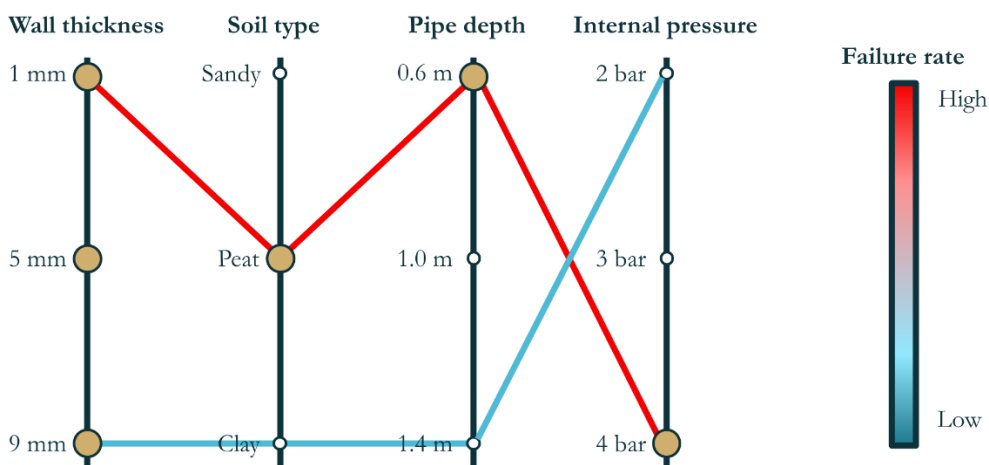
² Concrete voorbeelden van hoe de schematische voorbeelden in deze subparagraaf zich manifesteren in het Midas-dashboard zijn beschikbaar in bijlage VI, evenals een instructievideo



Figuur 10: Heatmap om de gevoeligheid van inputparameters op de output te visualiseren.

Parallele coördinatenplot: Naast de heatmap die beperkt tot twee parameters, wordt een plot met parallele coördinaten geïntroduceerd om het effect van meerdere parameters te visualiseren.

Het onderstaande schematische voorbeeld (Figuur 11) illustreert hoe een parallele coördinatenplot kan worden geïnterpreteerd. De vier coördinaten staan voor de vier inputparameters die van invloed zijn op de uitvalpercentages van leidingen. De kleur van de lijnen die over de coördinatenassen lopen, geeft de storingsgraad aan. In dit specifieke voorbeeld geeft het uitvoeren van een COMSIMA-simulatie met een wanddikte van 1 mm voor een leiding die in veen is gelegd, 0,6 m onder het oppervlak ligt en bij een inwendige druk van 4 bar werkt, aan dat de faalkans hoog is. Voor een leiding met een wanddikte van 9 mm, gelegd in klei, 1,4 m onder het oppervlak en werkend bij een inwendige druk van 2 bar, is de faalkans daarentegen laag. Een dergelijke plot maakt het mogelijk het effect van meerdere parameters tegelijk te visualiseren en biedt dus een globaal overzicht. Het kan echter nog steeds niet worden aangemerkt als een globale gevoeligheidsanalyse, aangezien de kansverdelingsfuncties van de inputparameters niet gemakkelijk zichtbaar zijn. Bovendien is het onduidelijk of het bekijken van deze plot een direct antwoord biedt op de vraag "onzekerheid in welke inputparameter de meeste onzekerheid in de output veroorzaakt". Wel biedt een dergelijke plot ook een mogelijkheid om een lokale gevoeligheids- of screeningsanalyse uit te voeren. Men kan eenvoudig het bereik van aanvaardbare invoerwaarden voor drie van de vier invoerparameters vaststellen (slechts één grote cirkel per bodemtype, diepteligging, interne druk in Figuur 11) en alleen het effect van de wanddikte afzonderlijk visualiseren (meerdere grote cirkels voor wanddikte in Figuur 11).



Figuur 11: Parallele coördinatenplot om het effect van meer dan twee inputparameters op de output te visualiseren. Gevulde grote cirkels laten zien hoe een dergelijke plot ook kan worden gebruikt voor lokale gevoeligheids- of screeningsanalyse.

5.4.3 Implementatie in Midas-dashboard

Zoals beschreven in paragraaf 5.1 is het bovenstaande geïmplementeerd in een dashboard. Op het dashboard kan het effect van acht parameters worden gevisualiseerd. Er zijn drie categorische en vijf numerieke parameters. De keuzes of bereiken van de acht parameters zijn:

- Buismateriaal: PVC of AC
- Grondwatertrap³: I of VIII
- Bodemtype: zand of klei
- Effectieve wanddikte: 10 - 80 mm
- Diepte (bovenkant buis): 0,5 - 1,2 m
- Diameter: 100 - 300 mm
- Druk: 100 -800 kPa
- Maximale grondzettingssnelheid: 0 - 1 m/j

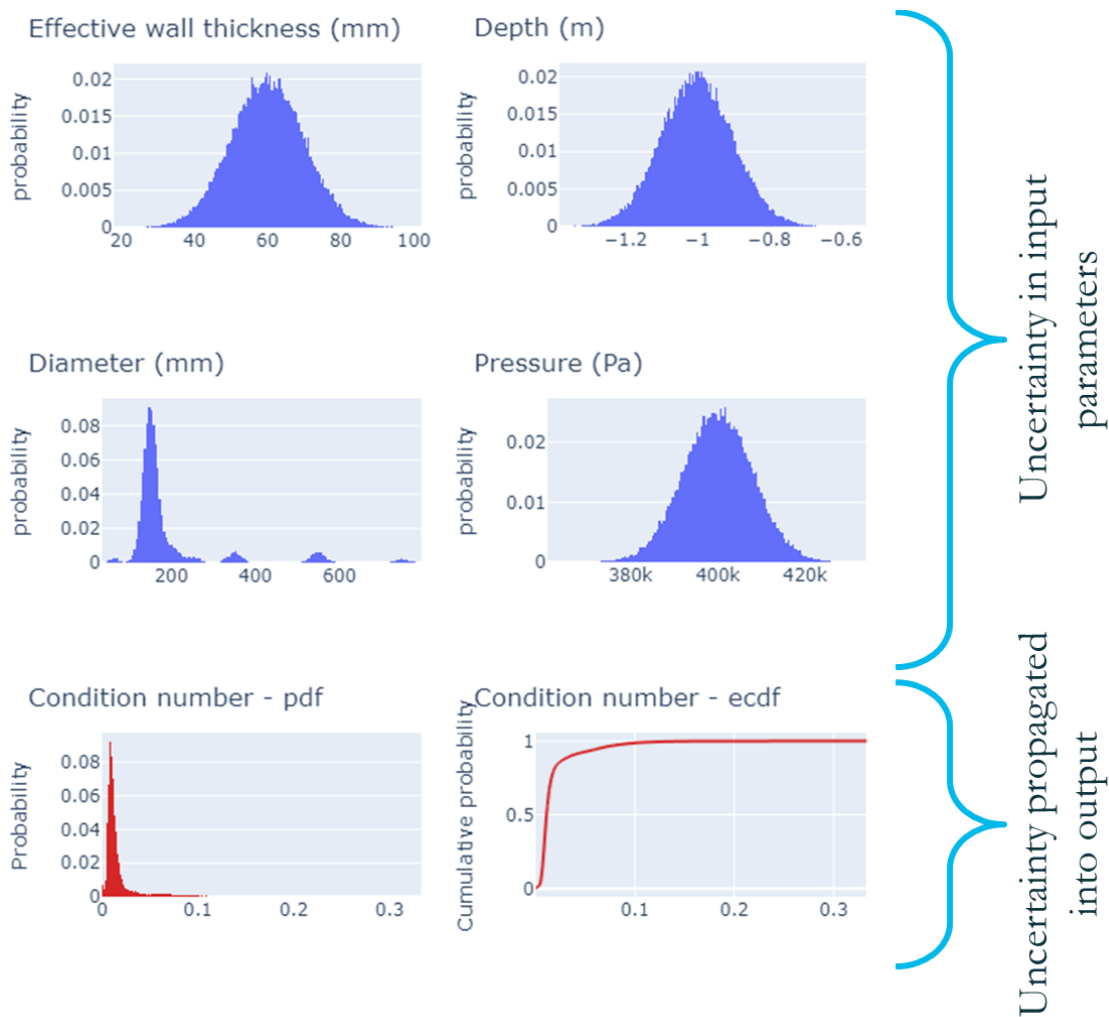
Door variatie van deze parameters en/of toepassing op relevante praktijkcasussen kan inzicht in de faalkans en meest relevante te meten parameters worden verkregen. Dit wordt nader geïllustreerd aan de hand van voorbeelden in Bijlage V.

5.5 Onzekerheidspropagatie

De onzekerheidspropagatiemodule op het dashboard visualiseert hoe de onzekerheid in de invoerparameters doorwerkt in de uitvoer. Een voorbeeld wordt getoond in Figuur 12. De vier panelen bovenaan geven de vier invoerparameters weer: effectieve wanddikte, diepte, diameter en interne druk. Voor elk van deze vier parameters is er een bijbehorende kansverdelingsfunctie voor de waarden die de invoerparameter kan aannemen. Behalve de diameter hebben alle andere variabelen een normale verdeling. De verdeling voor diameters is gebaseerd op een verdeling die is gegenereerd op basis van de huidige distributie in de praktijk. De twee panelen onderaan geven de output van het conditiegetal weer.

Door de verschillende mogelijkheden in de invoerparameters te combineren, zijn talrijke COMSIMA-simulaties uitgevoerd, die allemaal een overeenkomstige output voor het conditiegetal opleveren. De variabiliteit in de output wordt gevisualiseerd met een kansverdelingsfunctie en een cumulatieve verdelingsfunctie. De kansverdelingsfunctie voor het conditiegetal suggereert dat in de meeste COMSIMA-simulaties het conditiegetal 0,1 of lager was. Dit wordt bevestigd door de cumulatieve verdelingsfunctie, waarbij in meer dan 97% van de simulaties een conditiegetal van 0,1 of lager werd bereikt (een laag conditiegetal betekent een goede leidingconditie) .

³ indicatie van de absolute grondwaterstand en fluctuatie hiervan, zie <https://nl.wikipedia.org/wiki/Grondwatertrap>



Figuur 12: Impressie van de onzekerheidvoortplantingsmodule van het dashboard. Gebruikers kunnen ook de details zien door met de cursor op de verschillende plots te gaan staan.

6 Samenvatting en aanbevelingen voor eindgebruikers

6.1 Samenvatting

Aan het begin van het onderzoek is een literatuurstudie uitgevoerd. Hierbij is een inventarisatie gemaakt van parameters die de conditie van een leiding beïnvloeden. Hierbij hebben wij ons vooral gericht op (invoer van) fysische/mechanisme modellen. Er is een breed kwalitatief overzicht gepresenteerd met factoren die van invloed zijn op het falen van een leiding. Deze factoren zijn in drie categorieën ingedeeld: intrinsieke buis, omgeving en operationeel. Voor kwantitatief beschreven modellen van verschillende belastingen op ondergrondse buisleidingen zijn de invoerparameters geïdentificeerd (zoals inwendige druk, uitwendige diameter en wanddikte). Een volledig overzicht is weergegeven in Bijlage I.

6.1.1 Databeschikbaarheid en -kwaliteit

Er is een matrix opgesteld die faalmechanismen, modelparameters, en verklarende variabelen van de faalmechanismen samenbrengt en er is uitgezocht waar en in welke mate de betreffende parameters beschikbaar of verkrijgbaar zijn.

De beschikbaarheid van de verklarende variabelen bleek voor interne parameters (bijvoorbeeld: hydraulische en chemische data) beperkt. Indien beschikbaar, was dit uitsluitend het geval voor de huidige situatie maar niet voor de historische data. Verder bleken er een aantal geografische databronnen niet nauwkeurig genoeg voor het project.

Er is inspectiedata aangeleverd door Acquaint en openbare geografische data door HDM en Spatial Insight. Deze datasets zijn samengevoegd op basis van van geografische locatie. Vervolgens is op de dataset een volledige analyse uitgevoerd op de kwaliteit van de data. Deze analyse is uitgevoerd op de openbare geografische data. Hiervoor zijn twee methodes gebruikt. Onderling vergelijken van verschillende databronnen en vergelijking met validatiemetingen. Voor de eerst methode zijn verschillende databronnen met dezelfde waarden vergeleken. Hieruit bleek dat er voor veel parameters matige (orde 10%) tot grote (orde 80%) verschillen tussen de waarden voor categorische variabelen worden geconstateerd. Voor de overige variabelen worden kleine verschillen geconstateerd; wel blijken voor veel parameters geen gegevens in de beschikbare datasets aanwezig. Samenvattend moet worden geconcludeerd dat zowel de beschikbaarheid als de kwaliteit van de beschikbare data niet consequent voldoende is voor toepassing in conditiemodellen.

6.1.2 Correlatieanalyse

Om voor de ontbrekende gegevens een surrogaat – een manier om deze aan te vullen – te vinden, is onderzocht of er voldoende sterke correlaties met andere, beter bekende parameters bestaan. De resultaten van de correlatieanalyse en verfijning leveren vooralsnog geen bruikbare correlaties op.

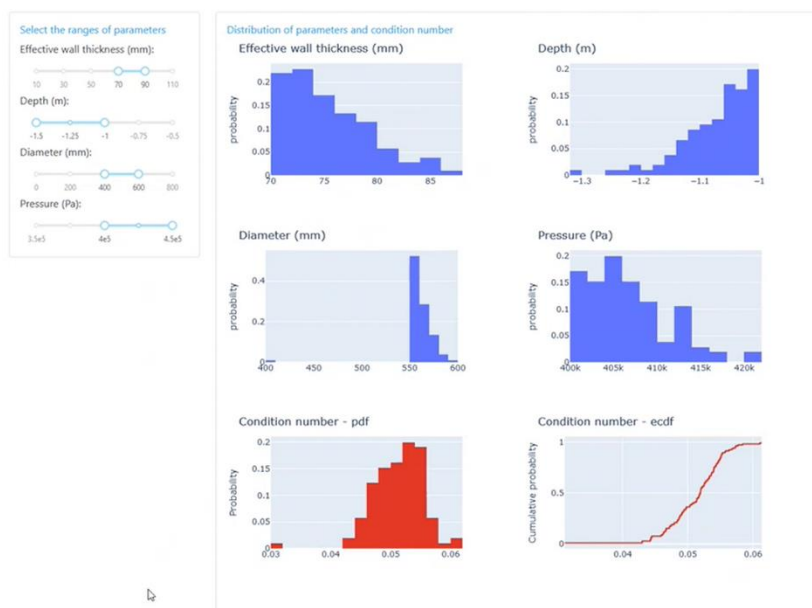
6.1.3 Gevoeligheidsanalyse en onzekerheidspropagatie

Een belangrijk resultaat van dit project is het Midas-dashboard (Figuur 13; let op dat hoewel een dashboard in de regel actuele gegevens weergeeft, het hier vooral bedoeld is voor het ontsluiten en analyseren van een grote dataset) dat gebruik maakt van gegevens die door de eindgebruikers aan de projectgroep zijn verstrekt. Het Midas-dashboard is ontwikkeld om de assetbeheerder te helpen beter geïnformeerde beslissingen te nemen (met actuele gegevens) in onzekere omstandigheden. Het dashboard bestaat uit twee componenten, die assetmanagers op de volgende wijze kunnen helpen:

- (1) Beoordeling van leidingnetconditie voor een leidingnetwerk: de analyse van onzekerheidspropagatie biedt een generiek kader voor het beoordelen van de kansverdeling van leidingcondities onder meerdere omstandigheden gegeven onzekerheid in relevante parameters. Eindgebruikers kunnen het diagram aanpassen op basis van de

beschikbare informatie over de belangrijkste parameters die ze al hebben. Vervolgens kunnen ze vanuit het dashboard zien of er aanvullende informatie moet worden toegevoegd en welke parameters beter moeten worden bepaald of gemeten.

- (2) Bepaling voor welke parameters een grove schatting voldoende is: de gevoeligheidsanalyse illustreert hoe de kansverdeling voor leidingfalen eruit zou zien op basis van veronderstelde of bekende kansverdelingen van de belangrijkste parameters die falen beïnvloeden. Dit is handig voor de gevallen waarin ervaren assetmanagers een redelijk bereik van een bepaalde invloedsfactor kunnen inschatten, zonder extra metingen.



Figuur 13: Dashboard voor het inzichtelijk maken van de gevoeligheden voor onzekere parameters en voortplanting van onzekerheid

6.2 Aanbevelingen voor praktische toepassing

- De ontwikkelde onzekerheidspropagatie en gevoeligheidsanalyse, geïmplementeerd in het dashboard, tonen het potentieel van een probabilistische benadering in besluitvorming omtrent assets aan. In de huidige versie kan het dashboard worden gebruikt om het begrip van de kansverdeling van leidingfalen en de invloed van de verschillende parameters (en onze kennis daarvan) op deze kansverdeling aan te scherpen. Aanbevolen wordt om het dashboard voor deze aanscherping van het begrip te gebruiken, evenals het gebruik van de methodiek om vast te stellen welke parameters het beste gemeten/bepaalde kunnen worden, omdat zij het meest bijdragen aan de reductie van onzekerheid van conditiemodellen.
- Tevens wordt aanbevolen om de probabilistische benadering van leidingfalen zoals behandeld in dit rapport in te bedden in het vervolgbesluitvormingsproces.

6.3 Aanbevelingen voor onderzoek

- Door een gebrek aan sterke correlaties in de huidige dataset is de mogelijkheid voor praktische toepassing van het voorspellen van ontbrekende gegevens (bijvoorbeeld degradatie) beperkt. Om dit te verbeteren bevelen wij aan dat er meer inspecties, operationele gegevens en nieuwe datasets met betrekking tot de interne aantasting van de buis worden verzameld. Op basis van deze nieuwe informatie kan opnieuw gezocht worden naar werkelijke correlaties en kunnen de verfijningsmodellen worden verbeterd.

- Er blijven onbeantwoorde vragen. Hoe kan de bandbreedte op faalkans worden verkleind? Op welk moment moeten oude leidingen vervangen worden? De beantwoording van dergelijke vragen kan worden vormgegeven in het kader van het bouwen van een besluitvormingstool (in een eventueel toekomstig project), die grotendeels steunt op de analyses die in dit rapport zijn gepresenteerd.

7 Referenties

Barton, N. A., Farewell, T. S., Hallett, S. H. & Acland, T. F., 2019. Improving pipe failure predictions: Factors affecting pipe failure in drinking water networks. *Water Research*, Volume 164, p. 114926.

Davis, P., Burn, S., Moglia, M. & Gould, S., 2007. A physical probabilistic model to predict failure rates in buried PVC pipelines. *Reliability Engineering & System Safety*, 92(9), pp. 1258-1266.

Davis, P. et al., 2008. Failure prediction and optimal scheduling of replacements in asbestos cement water pipes. *Journal of Water Supply: Research and Technology—AQUA*, 57(4), pp. 239-252.

Groen, E., sd *Introduction: uncertainty and sensitivity analysis in environmental modeling*. [Online]
Available at: <https://evelynegroen.github.io/introduction/Introduction.html>

Guo, S., 2020. *Uncertainty quantification explained*. [Online]
Available at: <https://towardsdatascience.com/managing-uncertainty-in-computational-science-and-engineering-5e532085512b>

Hetenyi, M., 1946. *Beams on elastic foundation*. Ann Arbor, Michigan: University of Michigan Press.

Kenton, W., 2022. *Sensitivity analysis definition*. [Online]
Available at: <https://www.investopedia.com/terms/s/sensitivityanalysis.asp>

Kiefner, J. F. & Vieth, P. H., 1989. *A modified criterion for evaluating the remaining strength of corroded pipe*, sl: Battelle Columbus Div., OH (USA).

Lu, J. P., Davis, P. & Burn, L. S., 2003. Lifetime prediction for ABS pipes subjected to combined pressure and deflection loading. *Polymer Engineering & Science*, 43(2), pp. 444-462.

Moerman, A. J. van Vossen en R.H.S> Beuken (2016) UKNOW: Zich op leidingdegradatie door samenhang in informatiesystemen. BTO 2016.031

Moser, A. & Folkman, S., 2008. *Buried pipe design*. sl:McGraw-Hill Education.

Rajani, B. B., Makar, J. M. & McDonald, S. E., 2001. Mechanical properties of grey cast iron water mains. *ASCE, Journal of materials in Civil Engineering*, pp. 1-45.

Rajani, B. & Kleiner, Y., 2001. Comprehensive review of structural deterioration of water mains: physically based models. *Urban Water*, 3(3), pp. 151-164.

Rajani, B., Zhan, C. & Kuraoka, S., 1996. Pipe soil interaction analysis of jointed water mains. *Canadian Geotechnical Journal*, 33(3), pp. 393-404.

Salehi, S., Jalili Ghazizadeh, M. & Tabesh, M., 2018. A comprehensive criteria-based multi-attribute decision-making model for rehabilitation of water distribution systems. *Structure and Infrastructure Engineering*, 14(6), pp. 743-765.

St. Clair, A. M. & Sinha, S., 2012. State-of-the-technology review on water pipe condition, deterioration and failure rate prediction models!. *Urban Water Journal*, 9(2), pp. 85-112.

van Laarhoven, K. A., 2020. *BTO 2020.001 Een model voor scheurgroei in PVCU buizen, geïmplementeerd in Comsima*, Nieuwegein: KWR.

Wilson, D., Filion, Y. & Moore, I., 2015. State-of-the-art review of water pipe failure prediction models and applicability to large-diameter mains. *Urban Water Journal*, 14(2), pp. 173-184.

Wols, B. A. & Moerman, A., 2017. *BTO 2017.074 Verouderingsmodule voor spanningsberekeningen in leidingen met COMSIMA*, Nieuwegein: KWR Watercycle Research Institute.

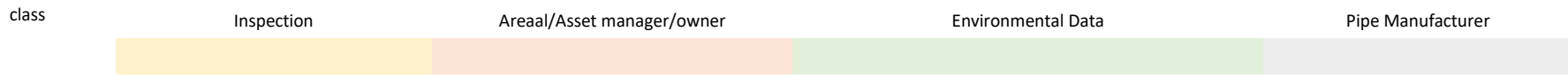
Wols, B. A., Moerman, A. & Vertommen, I., 2015. *BTO 2015.082 Comsima: model voor spanningen op ondergrondse leidingen*, sl: KWR.

Mechanisme →	Parameter ↓	COMSIMA	Hydrostatische druk	Drukpleken	Bodemdruk	Belasting in de lengterichting	Wielbelasting (dynamisch)	Bodemdaling	Belastingen door temperatuurstijging	Vorstbelasting	Analyse buis-bodeminteractie	Resterende structurele weerstand	Brosse breuk door gebreken aan het oppervlak van ABS kunststof buizen	Falen van PVC-buizen door inherente gebreken	Breuk van asbestcementbuizen door aantasting	Gegevens beschikbaar DWB/MS	ideale situatie directe gegevensbron
	Lengte buissegment	✓															
	Tijdsafhankelijkheid van de PVC-treksterkte/vloeispanning	✓															
	Statische treksterkte buismateriaal	✓									✓						productspecificaties
	PVC leeftijdgecorrigeerde (huidige) treksterkte	✓															
	Diepte grondwater	✓															
	Veerconsante bodem	✓									✓						Tekstboek
	Dichtheid droge bodem	✓			✓						✓						Tekstboek
	Dichtheid natte bodem	✓			✓						✓						Tekstboek
	Statische differentiële zetting	✓															
	Omvang van defect in materiaal			✓								✓	✓	✓			[--]
	Fluctuaties in belasting			✓													[--]
	Materiaalspecifieke scheurgroeiparameters			✓							✓	✓	✓	✓			Wetenschappelijke literatuur
	Elasticiteitsmodulus (vloeistof)			✓													Tekstboek
	Dichtheid (vloeistof)			✓													Tekstboek
	Hoeveelheid lucht			✓													[-]
	Leidingtracélengte			✓													LIS
	Effectieve sluitingstijd van de afsluiter			✓													[-]
	Grootte van snelheidsverandering			✓													Hydraulisch netwerkmodel
	Sleufwijdte				✓						✓						[-]

Mechanisme →	ideale situatie directe gegevensbron															
Parameter ↓	Gegevens beschikbaar DWB/MS	Breuk van asbestcementbuizen door aantasting	Falen van PVC-buizen door inherente gebreken	Brosse breuk door gebreken aan het oppervlak van ABS kunststof buizen	Resterende structurele weerstand	Analyse buis-bodeminteractie	Vorstbelasting	Belastingen door temperatuurstijging	Bodemdaling	Wielbelasting (dynamisch)	Belasting in de lengterichting	Bodemdruk	Drukplekken	Hydrostatische druk	COMSIMA	
Poreusheid van de bodem							✓									Tekstboek
Segregatiepotentieel							✓									[-]
Onbevoren watergehalte en thermische gradiënt aan het vriesfront							✓									[-]
Vorstdiepte							✓									[-]
Temperatuurverschil						✓										[-]
Reactiemodulus buis-bodem						✓										Tekstboek
Maximale diepte van het defect					✓											[-]
Doorsnede van het metaal dat in het aangetaste gebied verloren is gegaan, geprojecteerd op de lengteas van de buis					✓											[-]
Oorspronkelijke doorsnede van het aangetaste gebied					✓											[-]
Foliasfactor					✓											
Voorlopige of kortetermijn-breuktaaiheid					✓											productspecificaties
Afstand tot de buiswand aan de binnenzijde																[-]
Plaats van de breuk																[-]
Verdelingen van resterende treksterkte														✓		[-]

II Datavereisten en -beschikbaarheid

II.I Model-based



Class	Data required					Models classification				
	Description	Specifications	Website	Direct/indirect parameter	Notes KWR	Physical (Analysis of the stresses and the strengths of pipes)	Statistical (Model of the relationships between input variables and pipe breakage)		Simplistic (Ranks likelihood of pipe break using a single parameter)	
							Determ.	Probab.	ML	
Climate	Rainfall		https://www.knmi.nl/home							
	Temperature									
Ground temperature			-		The mean temperature increases with the depth. The amplification depends on parameters such as location and vegetation (ground cover)					

Soil properties (Backfill materia and Native soil)	Surface temperature	-							
	Resistivity			It requires extensive amounts of data that are historically unavailable and costly to acquire					
	Angle of internal friction								
	Settlements (mm/year) (differential settlement)	https://www.spaceoffice.nl/en/							
	Length of differential settlement								
	Axial pipe-soil reaction modulus			It represents the stiffness of the embedment soil placed on the sides of buried flexible pipe. E' is used to estimate the deflection of the pipe due to dead and live loads over the pipe. E' varies with the type of embedment soil and with the degree of compaction					
	pH								
	Soil suction	BRO							

	Soil moisture			Difficult to quantify directly because of the spacial variations in soil types and precipitation rates to which WDS are subject. It affects the bulk and shear moduli and linear soil expansion coefficient					
	Density	soilgrids.org							
	Soil type	http://maps.bodemdata.nl/bodemdata.nl/index.jsp https://bodemdalingkaart.nl/							
	Ondergrondgegevens	https://www.pdok.nl/introductie/-/article/basisregistratie-ondergrond-bro-							
	Volumetric swelling and shrinkage of clays								
	Corrosivity								
Pipe attributes	Material		Elasticity, yield stress						
	Diameter						X		
	Type								
	Length						X		
	Age (Installation year)								X

	Pipe vintage		Pipes made at a particular time and place may be more vulnerable to failure.						
	Internal average pressure								
	Internal and external coating								
	Original wall thickness		Internal and external corrosion will reduce it		X				
	Current Wall thickness								
	Status (operating/not operating)								
	Date of latest change of status								
	Latitude and longitude of a pipe's centroid								
	Current Depth								
	Depth of burial								
	Density of the material								
	Initial Ultimate tensile strength								
	Current tensile strength								
	Initial Yield Strength								
	Current yield strength								
	Elasticiteit modulus								
	Pipe deformation								

	Bedrijfdruk/Ontwerpdruk								
	Pipe laying condition								
	Transient pressure								
	Pipe resistance		It is based on pipe material, diameter and wall thickness						
	Technical status								
	Number of valves on the pipe								
Water temperature				In warm and cold seasons					
Water mineral content									
Angular displacement of joints									
Joint gap									
Type of joints									
Number of house connection in the same strict section									
Number of hydrants in The same street section									
Number of valves in the same street section									
External loadings	Ground water levels		-		-				
	Wind		-		-				

	manufacturer									
	water quality pH									
	Flow rate									
	Waterdeel		https://www.pdok.nl/introductie/-/article/basisregistratie-topografie-brt-topnl							
	Actueel Hoogtebestand Nederland		https://www.pdok.nl/introductie/-/article/actueel-hoogtebestand-nederland-ahn3-							
	Basisregistratie Kadaster		https://www.pdok.nl/introductie/-/article/basisregistratie-kadaster-brk-							
	Bebouwing (Building)		https://www.pdok.nl/introductie/-/article/basisregistratie-adressen-en-gebouwen-ba-1							
	Buisleiding gevaarlijke inhoud / ondergrondse infra (Pipeline hazardous content / underground infrastructure)		https://www.risicokaart.nl/welke-risicos-zijn-er/risico-transport-gevaarlijke-stoffen/buisleidingen							
	Waterweg/vaarweg		https://www.pdok.nl/introductie/-/article/nationaal-wegen-bestand-nwb-							

operating)													
Date of latest change of status	3	Applicable?	/										
Latitude and longitude of a pipe's centroid	3		YES	line geometry, no centroid please	3	Location of the pipe	Description of the route and depth (X, Y, Z); this information is available and reliable.		X				X
Current Depth	3	Combination with maaiveld			3	Depth of the pipes							
Depth of burial	3	If it is known	YES										
Density of the material					3								
Initial Ultimate tensile strength					3								
Current tensile strength					3								
Initial Yield Strength					3								
Current yield strength					3								
Elasticiteit modulus					3								
Pipe deformation					3								
Bedrijfdruk/ Ontwerpdruk					3		Pressure and pipe class.	X				X partially	X partially
Pipe laying condition	3		/										
Transient pressure	3		YES										
Pipe resistance	2	metallic pipes?	/										
Technical status							current technical condition and remaining technical lifespan						

	Number of valves on the pipe	2	Ontluchters(air vents)?	/									
	Water temperature	3		/									
	Water mineral content	3											
	Angular displacement of joints	3				3	Joint angles/Joint displacement						
	Joint gap	3											
	Type of joints	3		YES									
	Number of house connection in the same strict section	2	Focus on transport pipes	YES									
	Number of hydrants in The same street section	2		/									
	Number of valves in the same street section	2		/									
	External loadings	Ground water levels	3	See soil data	YES								
		Wind	1		YES								
		Ground		Soil type/density?	YES								
		Traffic	3		YES								
		Seismic load	3	Maybe only in Limburg of influence, for the clients in the project	YES								
		Frost depth	3	Ground temperature?	/								
Pipe failure hystorical data	Years of historical break records	2	Focus on condition instead of failures	/				maintenance history; analysis of historical failure events	X				X partially
	Number of pipes included in the dataset	2	Focus on condition instead of failures	/				maintenance history; analysis of historical failure events	X				X partially

	Natuur-gebieden					3		land registry, property, third party underground infrastructure. use of ground level/locations in relation to other infrastructure (gas pipeline, railways, barriers, dykes); this information is available and reliable.		X				X
Location of roads and roads classes		3	Correlation with traffic	/		3	Nationaal Wegen Bestand							
Spoorvak (Railway section)						3								
Stedelijk (Urban)						3								
3rd party activities				YES										
thrust in corners/bends				YES										
manufacturer				YES					X					X partially
water quality pH						3								
Flow rate						3		hydraulic functioning						X
Waterdeel						3								
Actueel Hoogtebestand Nederland						3								
Basisregistratie Kadaster						3								
Bebouwing (Building)						3								
Buisleiding gevaarlijke inhoud / ondergrondse infra (Pipeline hazardous content / underground infrastructure)						3		land registry, property, third party underground infrastructure						
Waterweg/vaarweg						3								

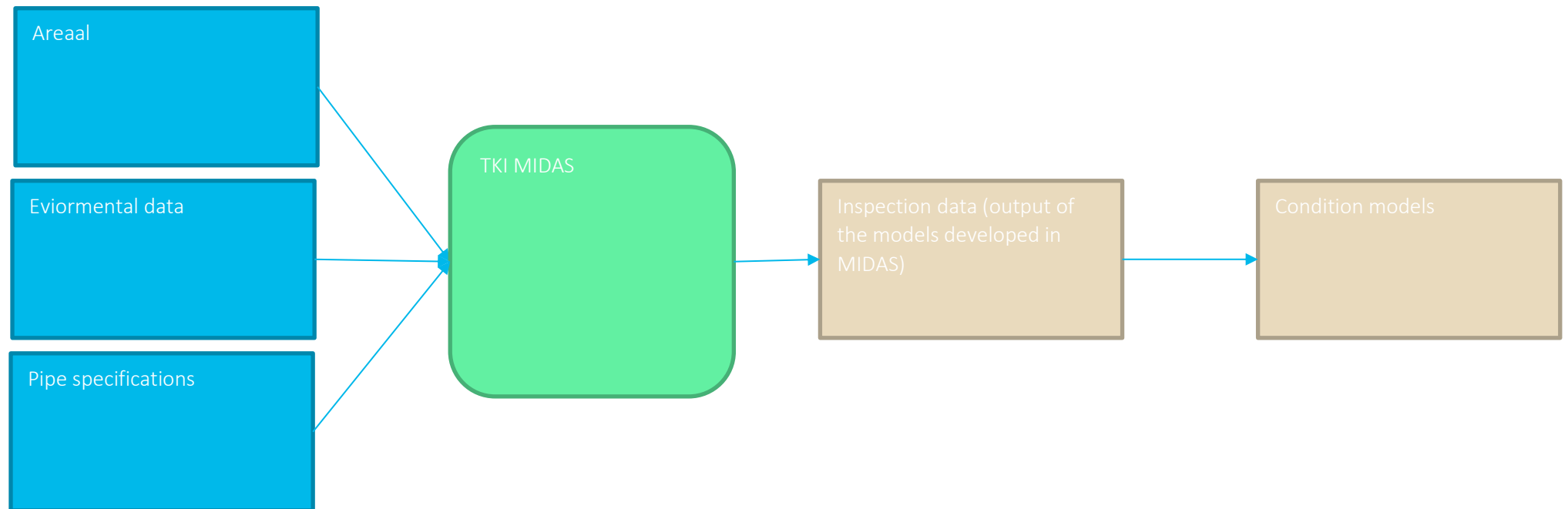
Special components (such as de-aerators and aerators, water hammer devices and valves)								repair pieces, vents, aerators, blocking devices, drilling saddles, valves, protective structures, jacket pipes, cathodic protection, water hammering devices. This information is available and reliable .	X					X
Type verbinding (type of junctions)									X					
sterkteberekeningen (strength calculation)								little is known					X partially	X partially
Environment								geografische informatie		X				X
Risk Inventory & Evaluation									X					
Connections									X	X			X partially	X partially
Hydraulic profile/cross profile according to design								working pressure, negative pressure (if applicable), air accumulation. this info is available and reliable	X					X partially
Allowable pressure, energy consumption								hydraulic functioning						X partially
Operating instructions (including venting/water precipitation)								partly available						X partially
pressure line, design pressure height									X					X
composition wastewater/ground water								only when separately authorised						X partially

Soil characteristics								Few to unknown						
----------------------	--	--	--	--	--	--	--	----------------	--	--	--	--	--	--

II.II Gemeten door Acquaint

Description	Specifications	Sensor	Direct/Indirect	Accuracy
Wall thickness	Wall thickness of measured material. In most cases we measure the residual healthy wall thickness	UT	Indirect: speed of sound required	~0.45mm
Leaching	Leaching in AC pipes	UT	Indirect: speed of sound , and most cases original wall thickness	~0.45mm
H2SOS4 degradation	Sulfide attack in cement based wastewater pipelines	UT	Indirect: speed of sound	~0.45mm
Corrosion in cast iron	Corrosion of cast iron pipes	UT	Indirect: speed of sound	~0.45mm
Inner radius	Inner radius on 16 clock positions		indirect: speed of sound medium	~1mm
Ovality	Tangential deformation of the pipe	UT	indirect: speed of sound medium	~1mm
Gap widths	Gap widths of joints on 16 positions	UT + Odometrie	Direct	minimum gap width of 24mm and ~1mm accuracy
Vertical displacement	Joint displacements in degrees	IMU	Direct	0.25 degrees
Horizontal displacement	Joint displacements in degrees	IMU	Direct	0.25 degrees
X coordinate	Coordinate in any requested reference system (mostly RD)	IMU+ODO+Markers	Indirect: GPS markers needed	1 meter
Y coordinate	Coordinate in any requested reference system (mostly RD)	IMU+ODO+Markers	Indirect: GPS markers needed	1 meter
Z coordinate	Coordinate in any requested reference system (mostly RD)	IMU+ODO+Markers	Indirect: GPS markers needed	1 meter
Axial deformation horizontal	Deviation from straight line	IMU+ODO	Direct	0.25 meter
Axial deformation vertical	Deviation from straight line	IMU+ODO	Direct	0.25 meter

Wire breaks in PCCP	If there is a wire break yes/no	MFW	direct	yes/no
Leakages	If there is a leakage	Hydrophone	Direct	yes/no
Anomalies	Large range of anomalies detected in one of the sensors	All	Direct	?



III Databeschikbaarheid bij de verschillende betrokken partijen

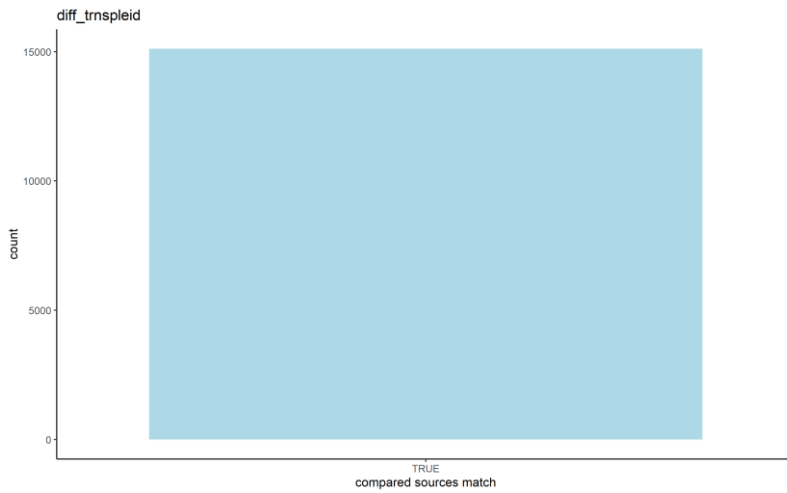
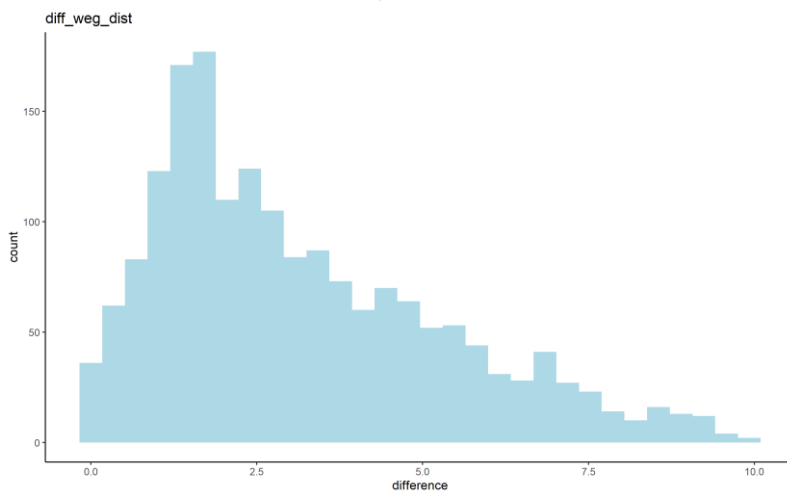
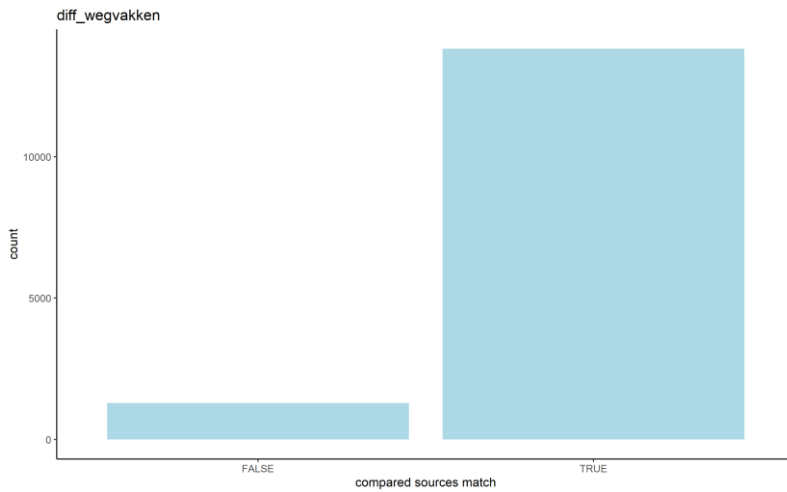
Parameters		Spatial Insight	HDM	ACQ	KWR
Soil chemistry (pH, lime, components)	pH external	x	x		
	Lime content				x
	Components soil	x	x		
Product type (series, manufacturer, vintage)	Series	x	x		
	Manufacturerer	x	x		
	Age	x	x	x	
Ground water depth and quality (SI, pH, components) (BRO, ...)	Saturation index	pH			x
	Ground water level	ja	possible, not checked		
	Components water	ja	possible, not checked		
External coatings soil type (BRO, ...) hazardous pipelines, (?) land use (?) (basisregistratie, Nature 2000, stadspolygonen, nationaal wegenbestand...)	External coatings	x	x	x	x
	Soil type	x	ja		
	Hazardous pipelines	ja	ja		
	Land use	x	ja		
	basis registratie	x	ja		
	Nature 2000	ja	ja		
	stadspolygonen	ja	ja		
	nationale wegenbestand	ja	ja		
	Internal water quality	x	x		
	Saturation index	x	x		
Water quality (SI, pH, components)	pH internal water	x	x		
	components	x	x		
	bron water (waar het water is gewonnen)	x			
Product type (series, manufacturer, vintage)	Industrie afvalwater		x		
	Internal coatings	x	x		x
Internal coatings	Relining	x	x		x
	flow rate (?)	x	x	x	
flow rate (?)	over time	x	x		
	operating pressure (?)	x	x	x	
operating pressure (?)	over time	x	x		
	Design pressure	x	x	ja	
actueel hoogtebestand NL (?) backfill and pipe installation quality	AHN	x	x		
	Pipe installation quality	x	x	x	x
	NEN 3650 is it used at installation	x	x		
	Mantelbuis			x	
	fundering	x	x	x	x
land use (?) (basisregistratie, Nature 2000, stadspolygonen, nationaal wegenbestand...)	Material Properties		x		
	Pipe depth	tov NAP	x	x	
Externe data beschikbaar					
Externe data nog niet beschikbaar					
Data klant/eindgebruiker niet beschikbaar					
Data klant/eindgebruiker beschikbaar					

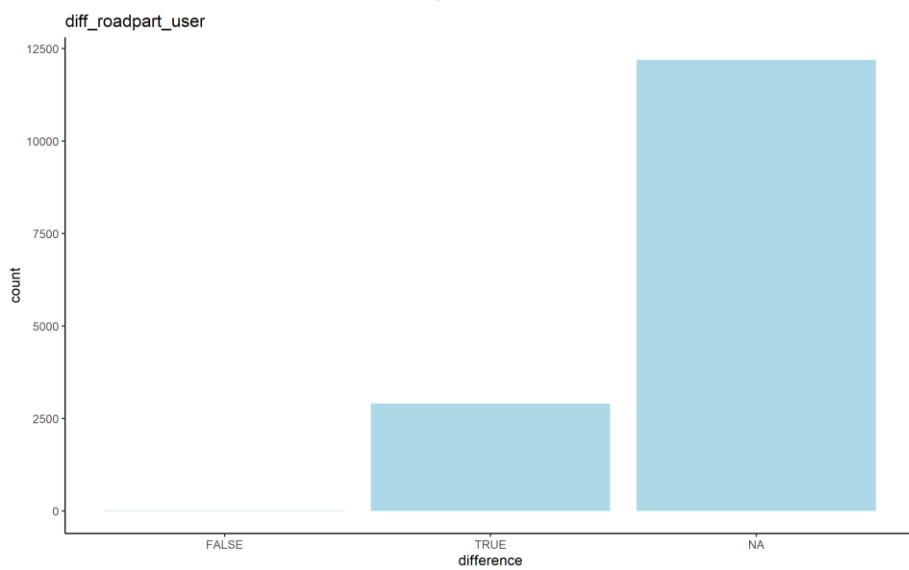
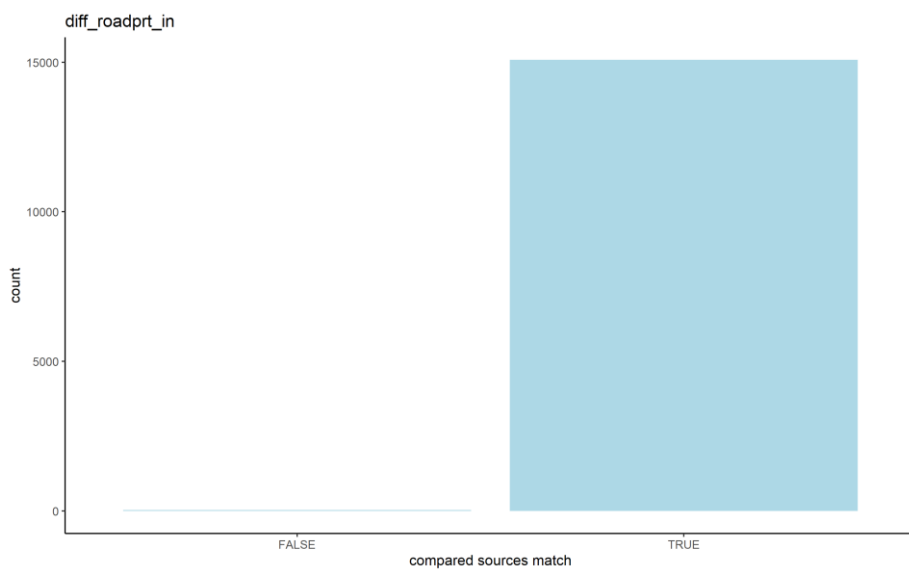
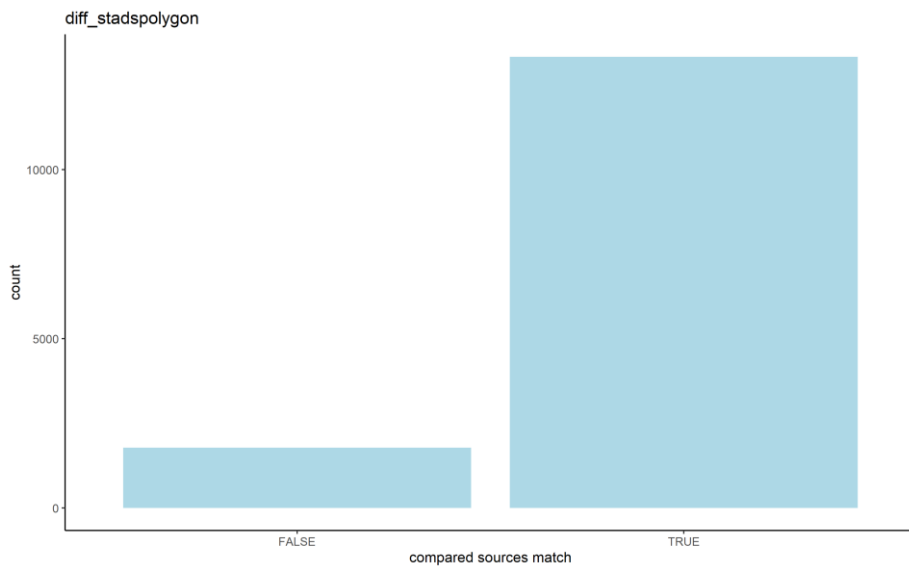
IV Koppelingen

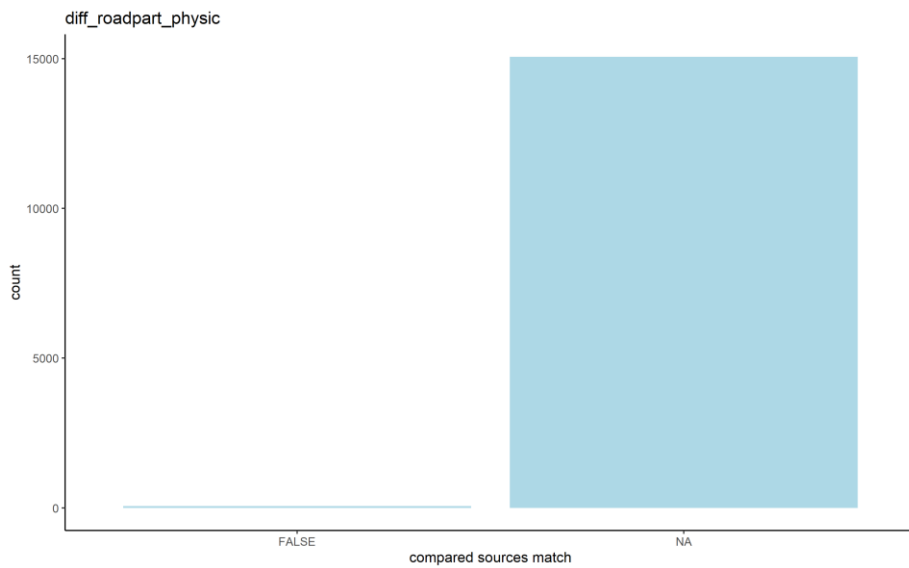
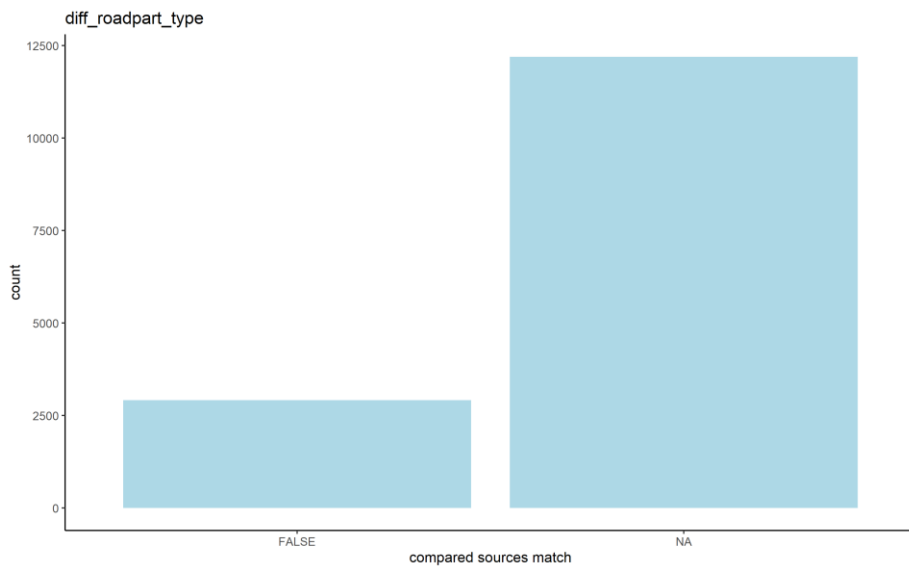
Main Table	Join Table	Join Type	Join Expression	comments	used fieldnames
aq_pipe_parts_refined	aq_joints	1:n	[Aq Pipe Parts Refined].midpoint.Buffer(5).Interacts([Aq Joints].geomaspoin)		joint_id, ..., joint_angle_vertical
aq_pipe_parts_refined	aq_pipe_parts	1:1	[Aq Pipe Parts Refined].pipe_part_id = [Aq Pipe Parts].id		pipe_part_id, pipe_part_type, created_at, ..., w_th_rem_std_min
aq_pipe_parts_refined	aq_inspections	1:1	[Aq Pipe Parts].inspection_id = [Aq Inspections].id		inspection_date -> pipe_age_at_insp_date
aq_pipe_parts_refined	aq_pipelines	1:1	[Aq Inspections].pipeline_id = [Aq Pipelines].id		medium, construction_year -> pipe_age_at_insp_date
aq_pipe_parts_refined	aq_grondsoortenkaart	1:n	[Aq Pipe Parts Refined].midpoint.Within(Grondsoortenkaart.geom)		grondsoort
aq_pipe_parts_refined	hdm_ahn1_stadspolygonen	1:1	[Aq Pipe Parts Refined].midpoint.Within([Hdm Ahn1 Stadspolygonen].geometry)		in_city_polygone
aq_pipe_parts_refined	hdm_bodemgebruik2015	1:1	[Aq Pipe Parts Refined].midpoint.Within([Hdm Bodemgebruik2015].geometry)		land_use
aq_pipe_parts_refined	hmd_bodemkaart_map_area		[Aq Pipe Parts Refined].midpoint.Within([Hdm Bodemkaart Map Area].geometry)		
aq_pipe_parts_refined	hdm_natura	1:1	[Aq Pipe Parts Refined].midpoint.Within([Hdm Natura].geometryaspolygon)		in_nature_reserve
aq_pipe_parts_refined	hdm_nwbwegenwegvlakken	1:n	[Aq Pipe Parts Refined].midpoint.Buffer(10).Interacts([Hdm Nwbwegenwegvlakken2].geometry)		wegvakken
aq_pipe_parts_refined	hdm_risicokaart	1:n	[Aq Pipe Parts Refined].geom.Buffer(10).Interacts([Hdm Risicokaart2].geometry)		trnspleid
aq_pipe_parts_refined	hdm_wegdeel_vlak	1:n	[Aq Pipe Parts Refined].midpoint.Buffer(10).Interacts([Hdm Wegdeel Vlak2].geometry)		wegdeel, ..., wegdeel_verhardingsbreedte klasse

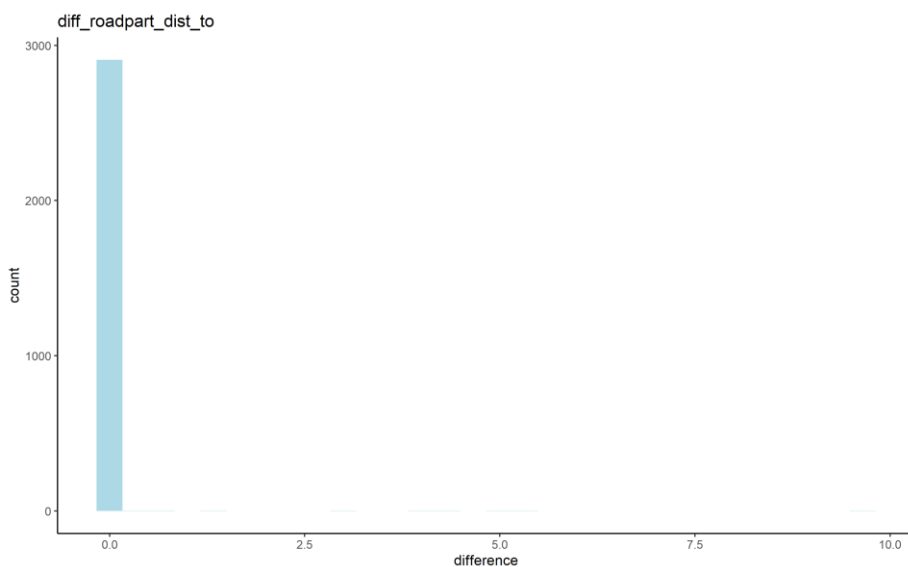
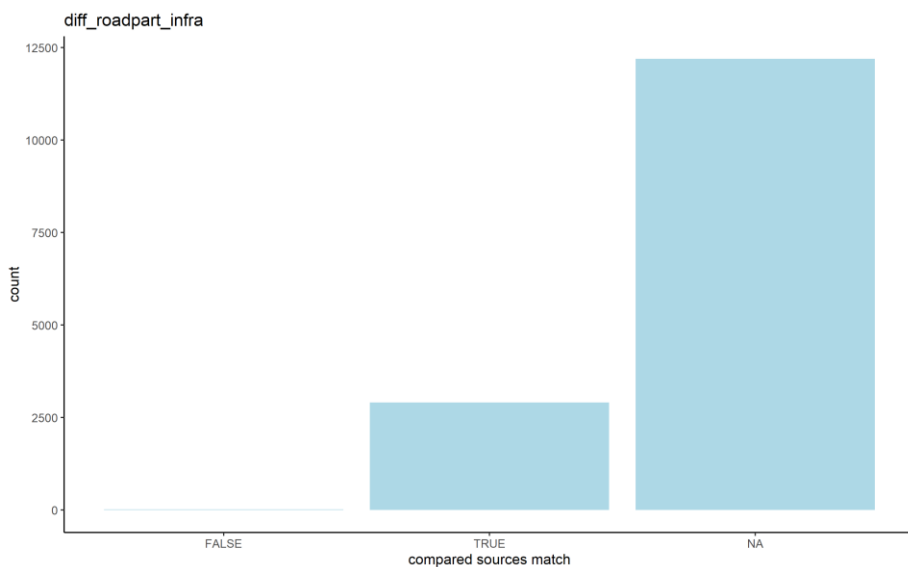
aq_pipe_parts_refined	si_grondwaterdynamiek_glg_ghg	1:n	[Aq Pipe Parts Refined].midpoint.Buffer(200).Interacts([Si Grondwaterdynamiek Glg Ghg].geometry)	IDW^2 as method	groundwater_glg, groundwater_ghg
aq_pipe_parts_refined	si_grondwaterkwaliteit	1:n	[Aq Pipe Parts Refined].midpoint.Buffer(15000).Interacts(Grondwaterkwaliteit.geom)	most recent einddatum, grootheidcode = "pH", filter_id = "001" (filter closest to maaiveld). IDW^2 as method	pH
aq_pipe_parts_refined	si_ahn	1:n	[Aq Pipe Parts Refined].midpoint.DistanceTo(AHN.geometry) < 10	IDW^2 as method	maaiveld_tov_nap, midpnt_tov_mv, midpnt_in_grndwtr_soms, midpnt_in_grndwtr_altijd

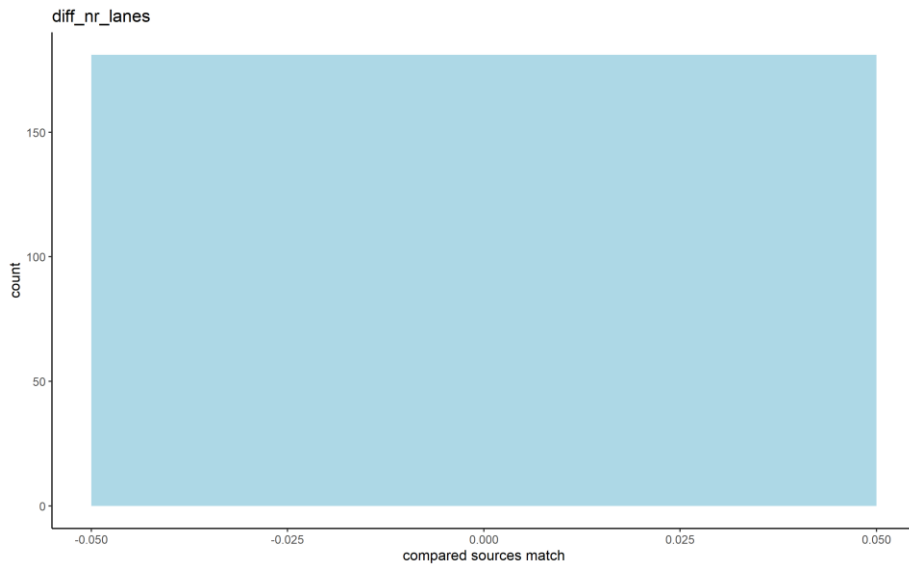
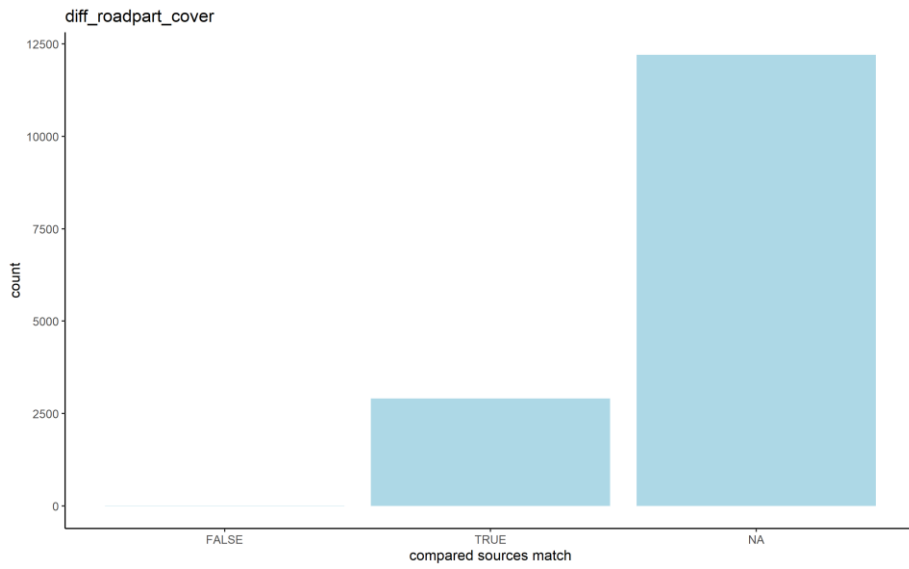
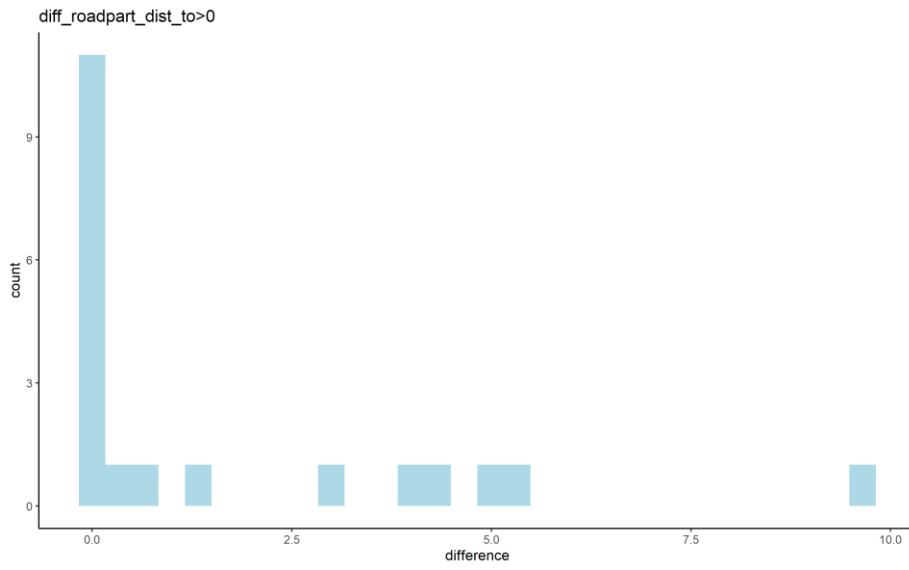
V Histogrammen datakwaliteitsbeoordeling

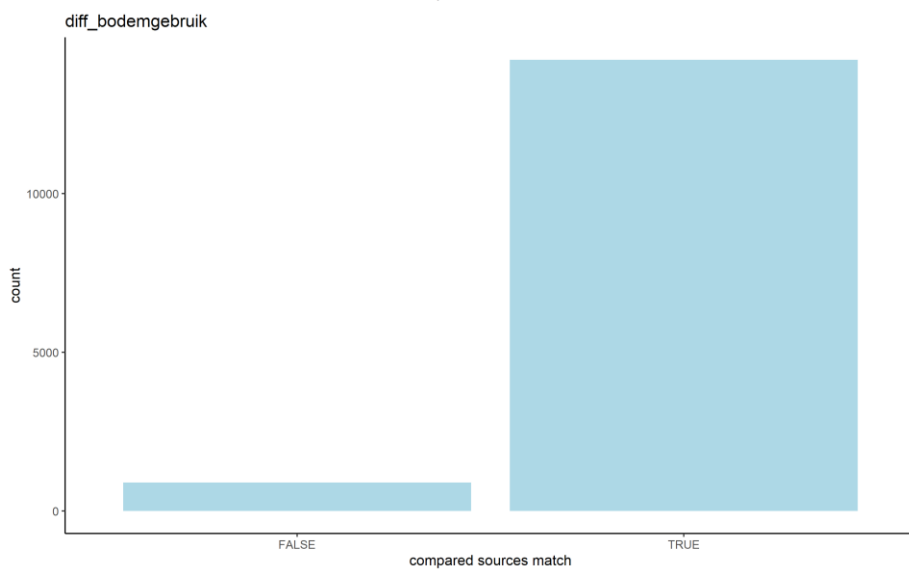
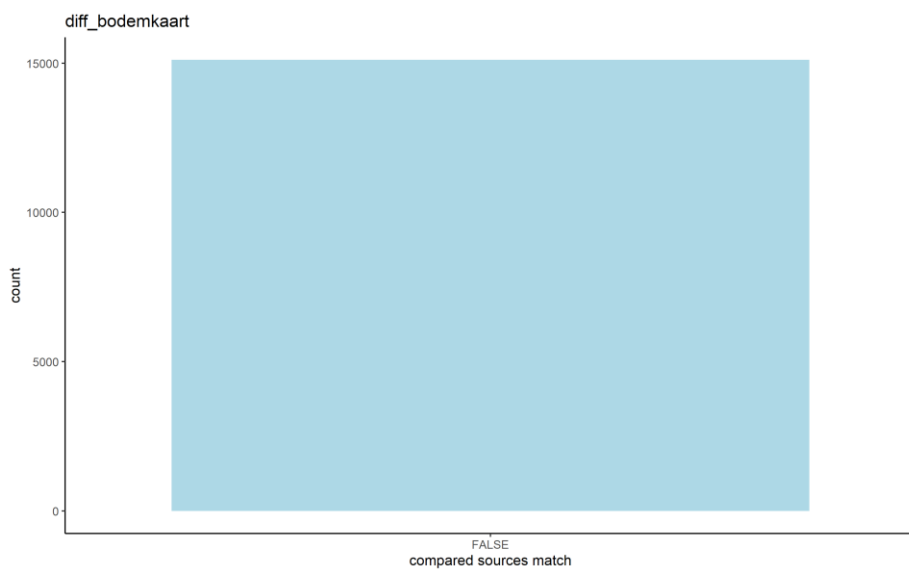
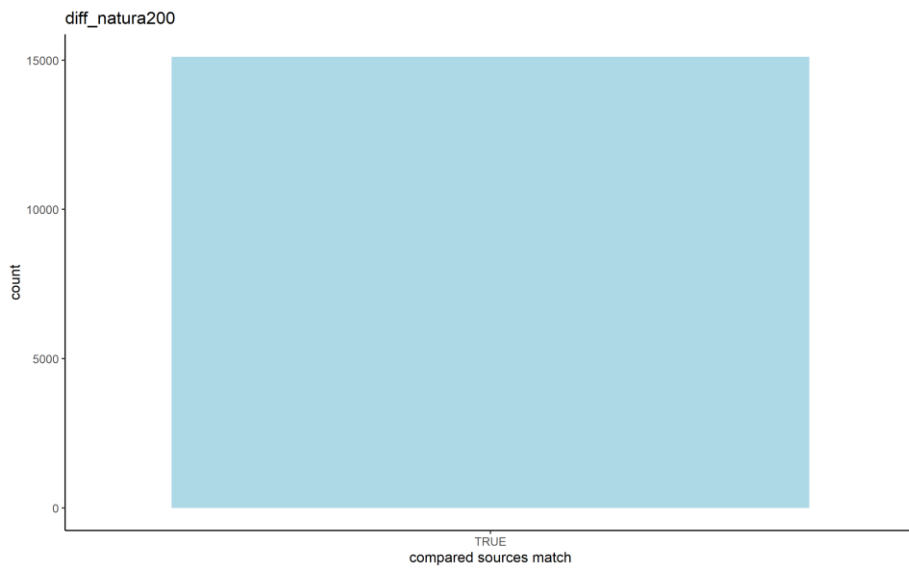












VI Gebruiksvoorbeelden Midas-dashboard

Heatmap

Een heatmap is een plot met twee assen die de twee parameters bevat waarvan men het effect wil bestuderen. Daarom is het een vereiste dat de overige parameters constant worden gehouden. De twee coördinatenassen stellen de te onderzoeken parameters voor, terwijl de kleuren de overeenkomstige conditiescores zijn.

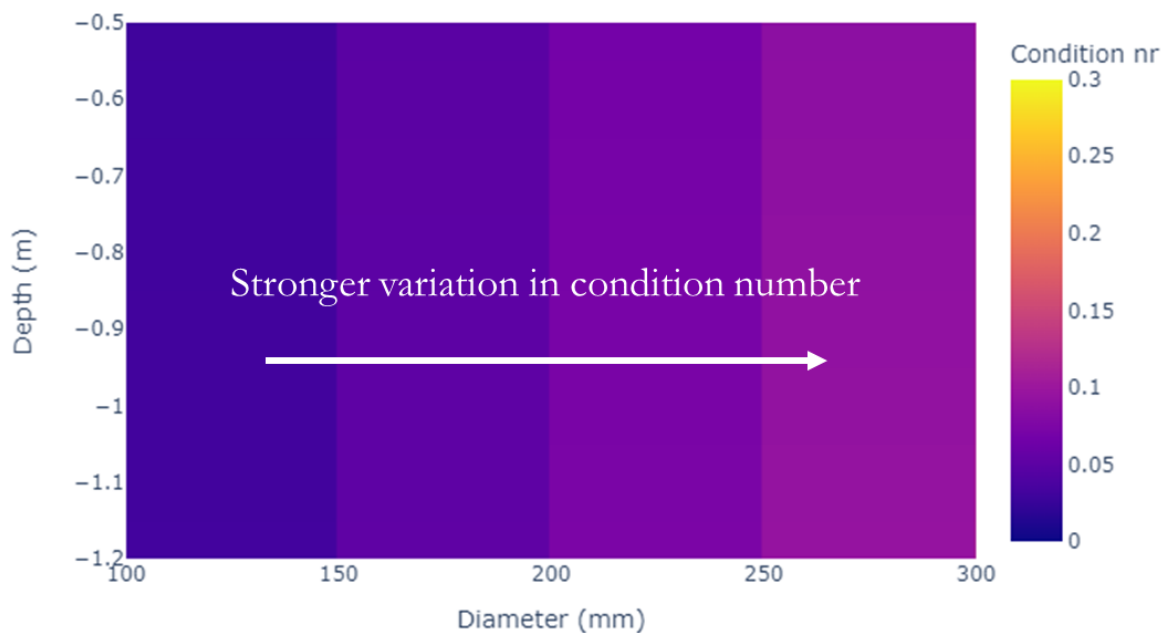
Aangezien een heatmap slechts het effect van twee parameters visualiseert, moeten de overige zes (drie categorische en drie numerieke) parameters worden vastgesteld. Het "leidingmateriaal" wordt vastgesteld als "PVC", het "grondwatertype" als "I", en het "bodemtype" als "zand". De waarde van de numerieke parameters is het gemiddelde van het overeenkomstige bereik. Het bereik van de effectieve wanddikte is 10-80 mm. Het gemiddelde hiervan is 45 mm. Evenzo zijn de gemiddelde waarden voor diepte, diameter, druk en maximale grondzettingssnelheid respectievelijk 0,85 m, 200 mm, 450 kPa en 0,5 m/j.

Voorbeeld 1:

Laten we een voorbeeld bekijken waarin "Diameter" wordt gekozen als eerste parameter (horizontale as) en "Diepte" als tweede (verticale as, Figuur 14). De andere numerieke parameters zijn vastgesteld: effectieve wanddikte op 45 mm, druk op 450 kPa, en maximale zettingssnelheid op 0,5 m/j. Er wordt een heatmap verkregen in een rasterstructuur waarbij de kleur de faalratio aangeeft. De heatmap laat zien dat de conditiescore toeneemt met toenemende "Diameter" (zichtbaar met het blote oog) en met toenemende "Depth" (alleen zichtbaar bij hoveren met de cursor; bij een diameter van 100-140 mm en een diepte van -0,6 m is de conditiescore 0,02999342; bij een diameter van 100-140 mm en een diepte van -1,1 m is de conditiescore 0,03124038).

Wat echter anders is, is dat de snelheid van toename van de conditiescore aanzienlijk verschilt voor deze twee parameters (of richtingen). Te zien is dat de heatmap die overeenkomt met het voorbeeld een horizontale gradiënt vertoont (de kleur verandert aanzienlijk in de horizontale richting, maar nauwelijks in de verticale richting). Dit suggereert dat de faalratio tamelijk ongevoelig is voor "diepte" (aangezien de kleuren in de verticale richting niet significant variëren). De faalratio is echter duidelijk gevoelig voor de "Diameter" (aangezien de kleuren in de horizontale richting aanzienlijk variëren).

Daarom is het in dit specifieke voorbeeld belangrijker om met grote zekerheid over de diameter te beschikken dan over de begraafdiepte. Dit komt door de grotere gevoeligheid van het faalpercentage van de buis voor de diameter. Merk op dat deze conclusie geldt voor de vaste combinatie van de overige zes invoerparameters.



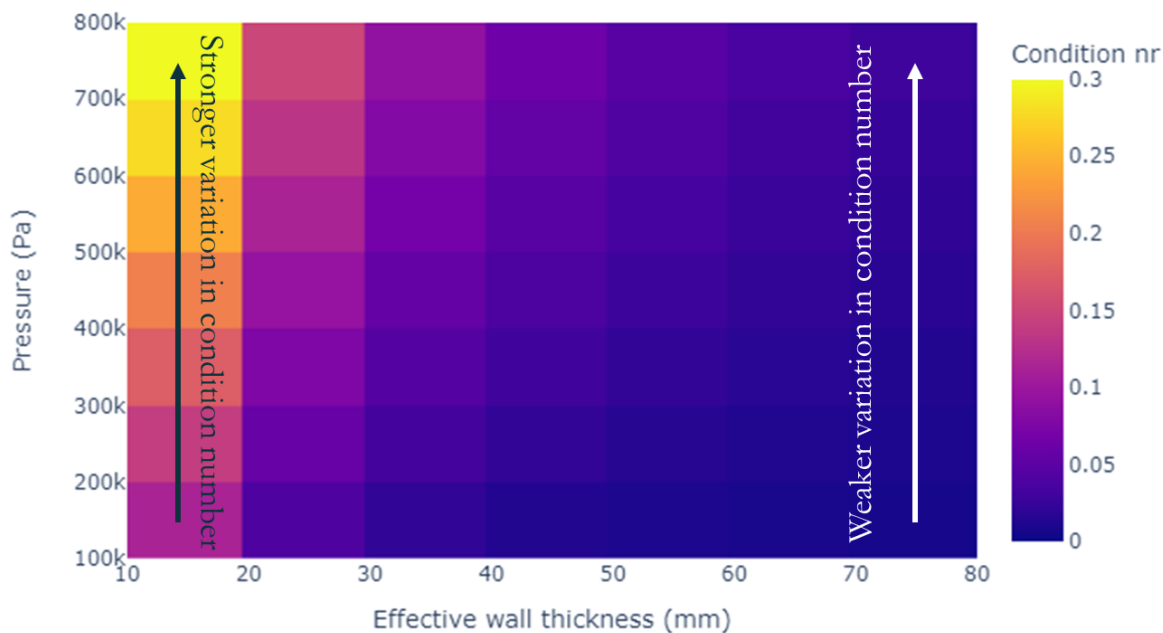
Figuur 14: Visualisatie van de heatmap voor voorbeeld 1.

Voorbeeld 2:

Laten we een voorbeeld nemen waarin "Effectieve wanddikte" wordt gekozen als eerste parameter (horizontale as) en "Druk" als tweede (verticale as, Figuur 15). De andere numerieke parameters zijn vastgesteld: diepte op 0,85 m, diameter op 200 mm, en maximale zettingssnelheid op 0,5 m/j. Er wordt een heatmap verkregen in een rasterstructuur waarbij de kleur de faalsnelheid aangeeft.

De heatmap in dit voorbeeld verschilt kwalitatief van het vorige voorbeeld. De gradiënten in de kleur zijn bijvoorbeeld niet-triviaal. Wanneer de effectieve wanddikte hoog is (> 70 mm), is er nauwelijks verandering in het conditiegetal langs de drukas. Dit betekent dat wanneer de effectieve wanddikte hoog is, het meten van de druk geen extra informatie oplevert over de conditie van de buis. Wanneer de effectieve wanddikte daarentegen laag wordt gemeten (10-20 mm), zal het meten van de druk veel informatie bieden. Dit komt omdat er een sterke variatie is in de kleuren of conditiescores bij variatie in druk (toenemende conditiescore bij toenemende druk).

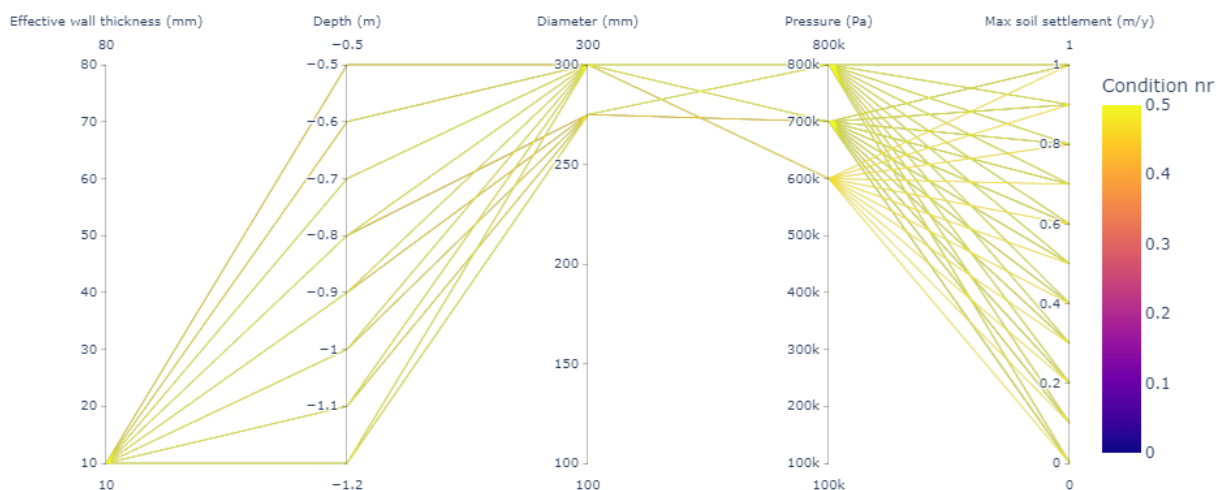
Daarom kan in dit specifieke voorbeeld op basis van de gemeten wanddikte worden besloten of het meten van de druk al dan niet een meerwaarde oplevert. Merk op dat deze conclusie geldt voor de vaste combinatie van de overige zes invoerparameters.



Figuur 15: Visualisatie van de heatmap voor voorbeeld 2

Parallele coördinatenplot

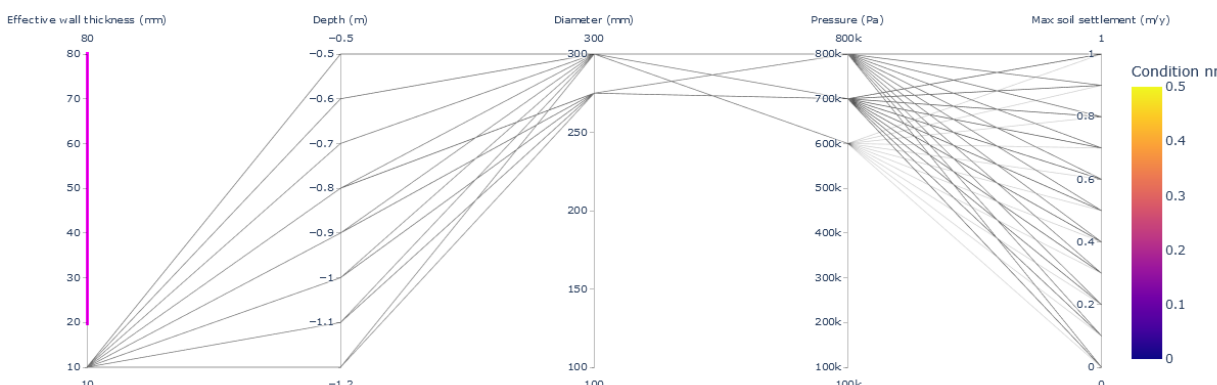
Om de invloed van meerdere parameters te visualiseren wordt een parallelle coördinatenplot (Figuur 16) ingevoerd, die een globaal overzicht biedt. Er zijn twee soorten invoerparameters: categorische (leidingmateriaal, grondwatertype, bodemtype) en numerieke (effectieve wanddikte, diepte, diameter, druk, maximale grondzetting). Door de gewenste categorische invoerparameters te selecteren, wordt de parallelle coördinatenplot voor de vijf numerieke invoerparameters gegenereerd. De kleur van de lijnen die over de coördinatenassen lopen, geven het conditiegetal aan. Door de drempel voor de conditiescore te verlagen, neemt het aantal lijnen toe.



Figuur 16: Impressie van de parallelle coördinatenplot met de drempelwaarde gekozen als 0,45, leidingmateriaal als AC, grondwatertype als I, bodemtype als zand.

Twee gevallen kunnen in aanmerking worden genomen. In beide gevallen wordt een (extreem hoge) drempelwaarde van "0,45" gekozen als conditiescore. Vanuit praktisch oogpunt zou deze waarde een drempel voor de conditiescore kunnen vormen die bij overschrijding een reden tot bezorgdheid zou zijn. Bovendien draagt de keuze van deze drempelwaarde bij tot een betere visualisatie van de plot met minder rommel. De aanwezigheid van een lijn wijst op een reden tot bezorgdheid (doordat de conditiescore de drempel overschrijdt) en omgekeerd. Verder beschouwen we het voorbeeld waarbij het leidingmateriaal "AC" is, het grondwatertype "I", en het bodemtype zand.

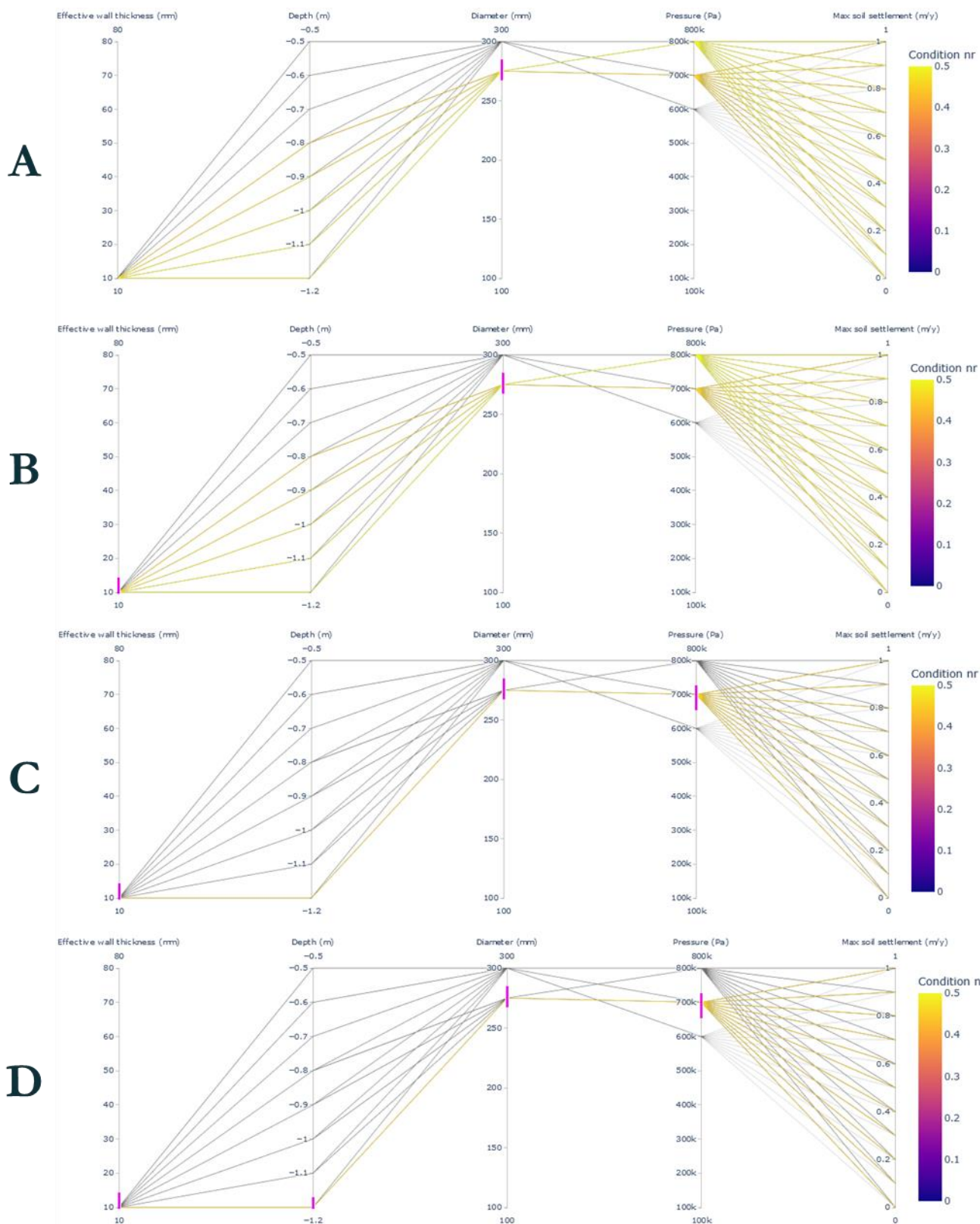
Geval I (Figuur 17): we nemen aan dat we weten dat de effectieve wanddikte 20 mm of meer bedraagt. In de plot met parallelle coördinaten kan men deze waarde/bereik opgeven. Door het bereik te selecteren, is te zien dat alle lijnen grijs zijn. De lijnen zijn grijs omdat er geen lijnen door deze selectie van effectieve wanddikte gaan. Het resultaat suggereert dat er geen verdere metingen nodig zijn voor de overige parameters die de conditiescore van de buis beïnvloeden (voor de huidige conditiedrempel van "0,45").



Figuur 17: Visualisatie van geval I voor de plot met parallelle coördinaten.

Geval II (Figuur 18): is bekend dat de buisdiameter 275 mm bedraagt.
 - Na selectie van de waarde op de plot zullen enkele lijnen grijs worden (met name die door diameters van 300 mm. De resterende niet-grijze lijnen kunnen helpen bij de beslissing welke parameters voorrang moeten krijgen om met grote zekerheid te worden gemeten. Dit is subplot A in Figuur 18.

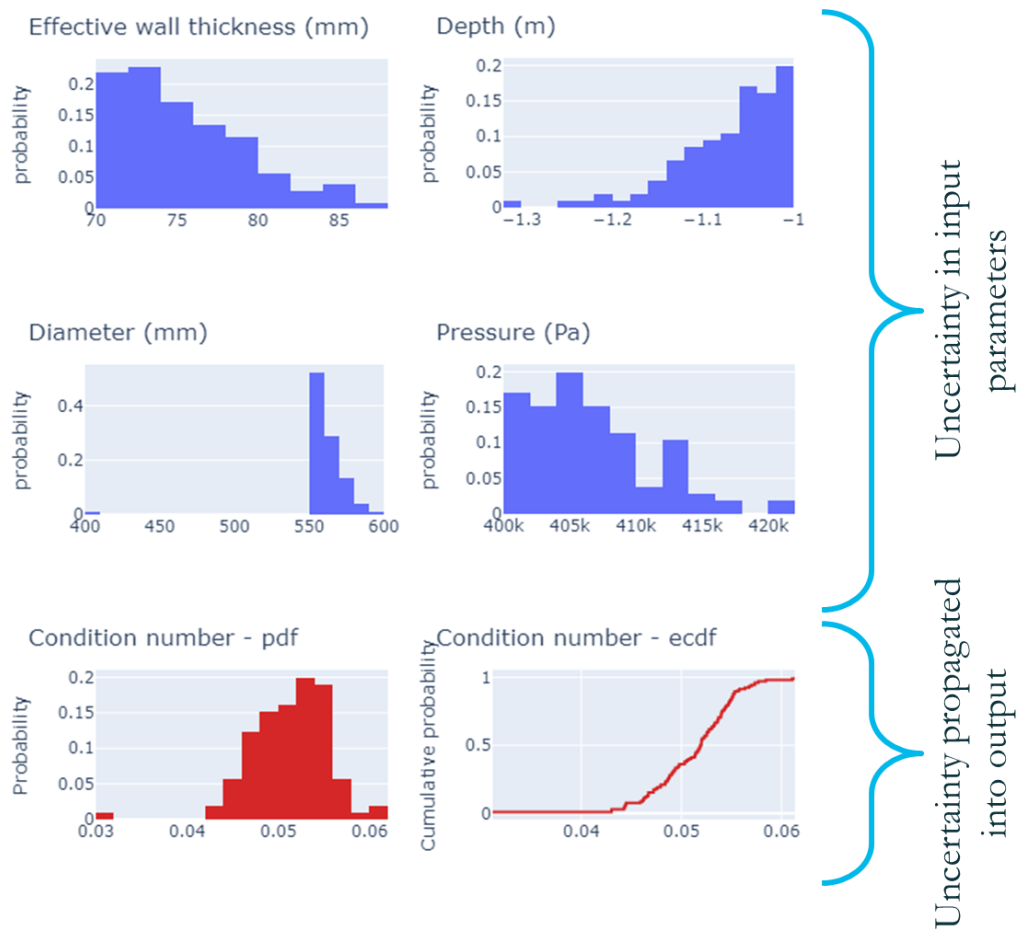
- Het beste is om eerst een parameter te kiezen met het smalste bereik waar (niet-grijze) lijnen doorheen lopen. Voor "Effectieve wanddikte" gaan de lijnen door een smal bereik (10 mm). Dit geldt ook voor de druk (600-800 kPa). Voor "Diepte" en "Maximale grondzetting" doorlopen de lijnen alle waarden, wat suggereert dat het eerst meten van deze waarden geen nuttige informatie oplevert. De "Effectieve wanddikte" wordt als eerste geselecteerd. Dit is deelplot B in Figuur 18.
- Te zien is dat de enige lijnen die door "Effectieve wanddikte" gaan, bij 10 mm liggen. Dit betekent dat eerst de effectieve wanddikte moet worden gemeten. Indien de effectieve wanddikte 20 mm of meer bedraagt, is het niet nodig metingen van andere parameters uit te voeren (aangezien er geen lijnen passeren).
- Is dit echter niet het geval (d.w.z. de effectieve wanddikte is 10 mm of lager), dan moet een andere parameter worden gemeten. Het beste is een parameter te kiezen die een duidelijke variatie vertoont. In dit geval zou dat "Druk" zijn. Indien de druk onder 700 kPa wordt gemeten, is het niet nodig andere parameters te meten (aangezien alle lijnen grijs of afwezig zijn). Dit is subplot C in Figuur 18.
- Als de druk 700 kPa bedraagt, blijven er nog minder niet-grijze lijnen over. Er blijven nog verscheidene niet-grijze lijnen over die door "Maximale bodemzetting" lopen. Daarentegen loopt er slechts één niet-grijze lijn door de as "Diepte" op 1,2 m. Het meten van "Diepte" biedt dus de meeste inzichten. Als de diepte wordt gemeten op minder dan 1,2 m, is het niet nodig verdere metingen te verrichten (aangezien er alleen grijze lijnen overblijven). Dit is subplot D in Figuur 18.
- Wordt echter een diepte van 1,2 m gemeten, dan ontstaat een moeilijke situatie. De enige overblijvende parameter is "Maximale zettingsgraad van de bodem". Er is duidelijk geen/nauwelijks variatie in de kleuren van de lijnen over het hele bereik (0-1 m/j). Dit betekent dat het meten van de maximale zettingsnelheid geen verdere inzichten zou opleveren en dat de metingen voor "effectieve wanddikte", "druk" en "diepte" tevergeefs waren.



Figuur 18: Visualisatie van geval II voor de plot met parallelle coördinaten

Voortplanting van onzekerheid

Door in het dashboard in te zoomen op de invoerparameters (via de schuifregelaars) kan het effect van het verkleinen van het bereik van de invoerparameters op de conditiescore worden gevisualiseerd. Dit wordt getoond in Figuur 19. De output (plots die overeenkomen met het conditiesgetal) wordt gewijzigd afhankelijk van het bereik van de inputparameters. Ook kan worden waargenomen dat de kansverdelingsfuncties hun fijnheid beginnen te verliezen en als histogrammen beginnen te verschijnen. De maximumwaarde in de kansverdelingsfunctie is 0,2, veel hoger dan 0,02 in Figuur 12). Dit geeft aan waarom het gebruik van een cumulatieve verdelingsfunctie wellicht beter is, omdat deze robuust is en ongevoelig voor de keuze van het invoerparameterbereik.



Figuur 19: Impressie van de voortplanting van de onzekerheid na verkleining van de invoerparameters.

VII Resultaten correlatieanalyse

VII.1 Correlatie voor alle datasets

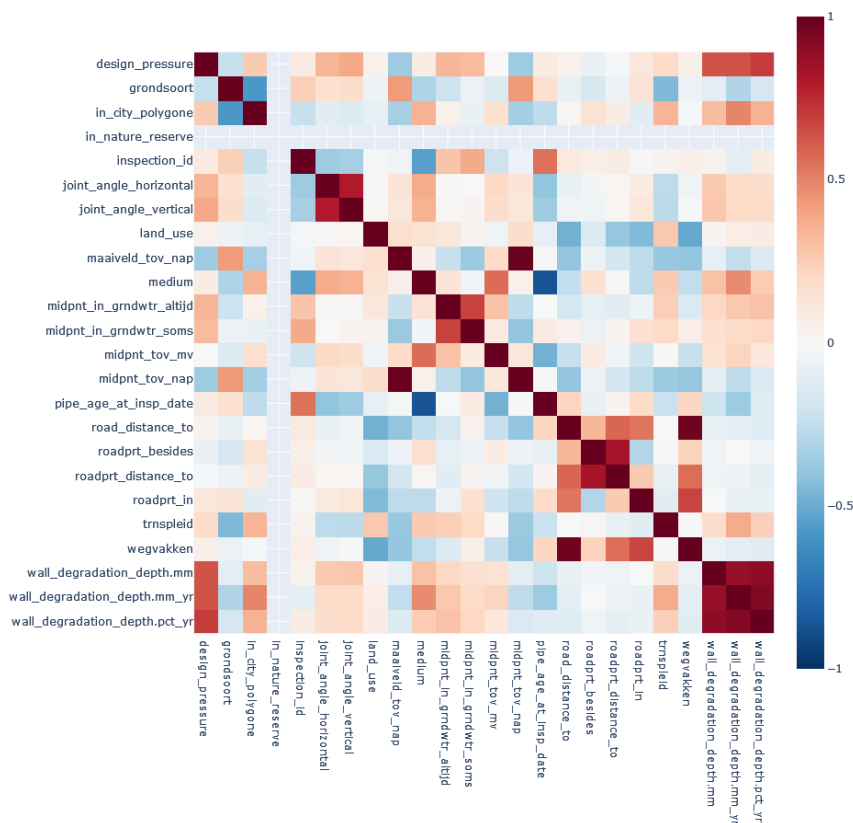
De volledige dataset bevat buisinspectiegegevens en open datasets. De inspectiegegevens hebben betrekking op de hoekverdraaiing van de verbinding en de degradatie van de wand. De open dataset omvat omgevingsfactoren, wegkenmerken, ondergrondse eigenschappen, grondgebruik, bodemtype en stadsgegevens.

De gegevens voor de correlatieanalyse worden opgeschoond door records met ontbrekende gegevens te verwijderen en de stringgegevens worden omgezet in gecategoriseerde gegevens. Om de degradatie van de buiswand goed weer te geven, berekenen we deze door gebruik te maken van de kolommen met de oorspronkelijke dikte van de wand en de resterende wanddikte in de loop van het jaar.

Uit de Spearman-afbeelding (Figuur 20) van alle datasets blijkt dat de correlatie tussen elk kenmerkenpaar in het algemeen niet sterk is. Er is echter één kenmerkenpaar, namelijk de buisdiepte t.o.v. NAP en het maaiveld t.o.v. NAP, dat een uitstekende correlatie vertoont, maar deze correlatie is vanzelfsprekend triviaal.

Een ander interessant resultaat is dat de degradatiegraad van de buiswand enigszins (orde 0.5) samenhangt met de ontwerpdruk. Hoe hoger de ontwerpdruk, hoe groter de degradatie. De degradatiesnelheid van de buiswand hangt echter niet samen met de leeftijd van de buis. De reden hiervoor is onderzocht en toegelicht in Van Laarhoven et al. (2019).

Na volledige analyse van de dataset bleek dat de tweede correlatie tussen de twee kenmerken die verband houden met de degradatiegraad van de wanddikte het medium is, of de transportleiding en of de leiding zich in de stad bevindt. Het grondwater vertoont ook een correlatie met de degradatie van de buiswand, maar die is iets lager dan de vorige kenmerken. Een logische vervolgvraag is dus hoe de correlatie eruit zal zien als de volledige dataset wordt opgesplitst in subcategorieën.



Figuur 20: Spearman's rangcorrelatiecoëfficiënt van alle datasets

VII.II Correlatie voor subcategorieën van de datasets

Het opsplitsen van de dataset in subcategorieën zou meer informatie kunnen opleveren. Na de bespreking met deskundigen hebben wij besloten de gegevens te splitsen volgens deze vier kenmerken: medium (mediumtype), trnspleid (of het transportleiding is), in_city_polygone (of het in de stad is), grndwtr_altijd (grondwaterstand).

Het totale aantal subcategorieën is 12, in plaats van 16 groepen. Dit komt doordat sommige leidingen niet onder de splitsingscriteria vallen, of doordat de records ontbrekende waarden bevatten die tijdens de voorbereiding van de gegevens zijn verwijderd.

De eindresultaten tonen correlatiescores tussen verschillende kenmerken. De leidingen in elke subcategorie hebben hun eigen kenmerken. De analyse van de subcategorieën toont het verschil tussen de volledige gegevensanalyse.

Tabel 6 toont twee klassen van kenmerkenparen, de eerste klasse bevat de hoogst gecorreleerde kenmerken, en de tweede klasse illustreert de zwak gecorreleerde kenmerken. Uit de matrix blijkt dat voor elke subcategorie de sterk gecorreleerde kenmerken variëren. De individuele correlatiematrices worden getoond in Figuur 21 - Figuur 31.

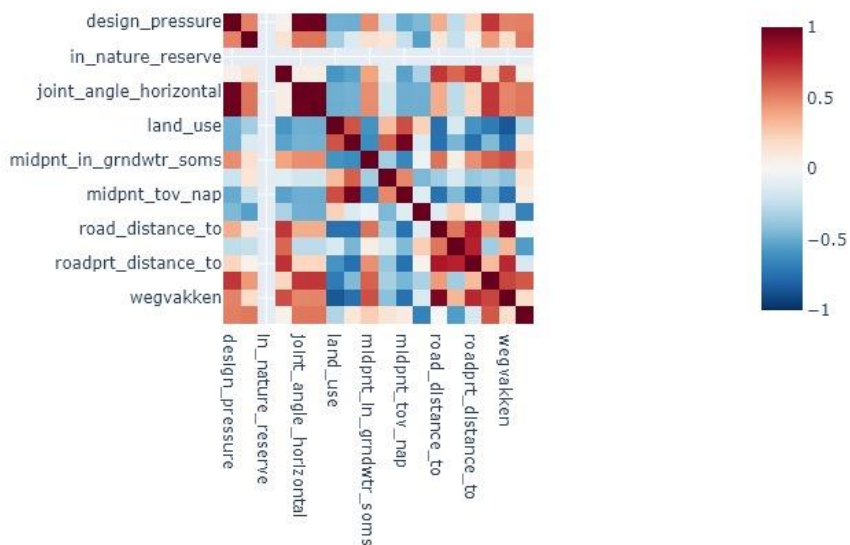
Tabel 6: Correlatie van sub-categorieën

Sub-categorie	Kenmerkpren van de eerste klasse	Kenmerkpren van de tweede klasse
drinking-water.0.0.0	<ul style="list-style-type: none"> - ontwerpdruk & verbinding, weg - verbinding & weg - grondgebruik & maaiveld, grondwater, weg - maaiveld & grondwater, buisdiepte, weg - leidingdiepte & weg - leeftijd leiding & degradatie - weg & degradatie 	<ul style="list-style-type: none"> - ontwerpdruk & bodemtype, grondwater, afbraak - grondsoort & verbinding, leeftijd van de buis, weg, degradatie - verbinding & grondgebruik, grondniveau, grondwater, buisdiepte, buisleeftijd, degradatie - grondgebruik & leeftijd leiding, afbraak - leeftijd leiding & weg
drinking-water.0.0.1	<ul style="list-style-type: none"> - ontwerpdruk & verbinding - verbinding & degradatie - landgebruik & weg - grondniveau & grondwater, leeftijd leiding, weg - leidingdiepte & weg - leidingleeftijd & weg, degradatie 	<ul style="list-style-type: none"> - ontwerpdruk & bodemtype, grondgebruik - bodemtype & grondniveau, buisdiepte, buisleeftijd - verbinding & grondgebruik - grondgebruik & leidingdiepte, afbraak - leidingdiepte & afbraak
drinking-water.0.1.1	<ul style="list-style-type: none"> - buisdiepte & weg 	<ul style="list-style-type: none"> - maaiveld en degradatie
drinking-water.1.0.0	<ul style="list-style-type: none"> - maaiveld & grondwater, buisdiepte, weg - grondwater & buisdiepte, weg - buisdiepte & weg 	<ul style="list-style-type: none"> - landgebruik & weg, degradatie - maaiveld & degradatie
waste-water.0.0.0	<ul style="list-style-type: none"> - ontwerpdruk & maaiveld, buisdiepte, buisleeftijd, degradatie - maaiveld & leidingdiepte, leeftijd leiding - leeftijd van de buis & degradatie 	<ul style="list-style-type: none"> - ontwerpdruk & grondgebruik - bodemtype & grondniveau, buisdiepte, buisleeftijd - grondgebruik & leeftijd leiding, weg, afbraak - maaiveld & weg, afbraak - grondwater & leidingdiepte

		- leidingdiepte & afbraak
waste-water.0.0.1	- ontwerpdruk & leidingleeftijd, degradatie - maaiveld & leidingdiepte, leidingleeftijd - buisdiepte & buisleeftijd - ouderdom van de buis & degradatie	- ontwerpdruk & maaiveld, buisdiepte - grondsoort & maaiveld, buisdiepte - maaiveld & weg - buisdiepte & weg
waste-water.0.1.0	- grondgebruik & maaiveld, leidingdiepte - maaiveld & leidingdiepte - grondwater & leeftijd leidingen	- bodemtype & landgebruik, maaiveld, buisdiepte, weg - leidingdiepte & weg
waste-water.0.1.1	geen	- grondgebruik & maaiveld, leidingdiepte - maaiveld & buisdiepte
waste-water.1.0.0	- bodemtype & landgebruik, grondwater, weg - verbinding & leeftijd leiding - grondgebruik & maaiveld, grondwater, buisdiepte, weg - maaiveld & buisdiepte - grondwater & weg	- bodemtype & leidingdiepte, leidingleeftijd - verbinding & landgebruik, grondwater, buisdiepte, weg - landgebruik & degradatie - maaiveld & grondwater, leeftijd buis, weg - grondwater & degradatie - leidingdiepte & weg - weg & degradatie
waste-water.1.0.1	- bodemtype & verbinding , grondniveau, buisdiepte, buisleeftijd, degradatie - verbinding & maaiveld, buisdiepte, buisleeftijd, degradatie - grondniveau & buisdiepte - buisdiepte & buisleeftijd	- grondsoort & weg - verbinding & weg - grondgebruik & weg - grondniveau & leeftijd leiding, weg, afbraak - leidingdiepte & weg, afbraak - leeftijd leiding & weg, afbraak
waste-water.1.1.0	- maaiveld & buisdiepte	geen
waste-water.1.1.1	- landgebruik & weg - maaiveld & buisdiepte	- landgebruik & leidingdiepte

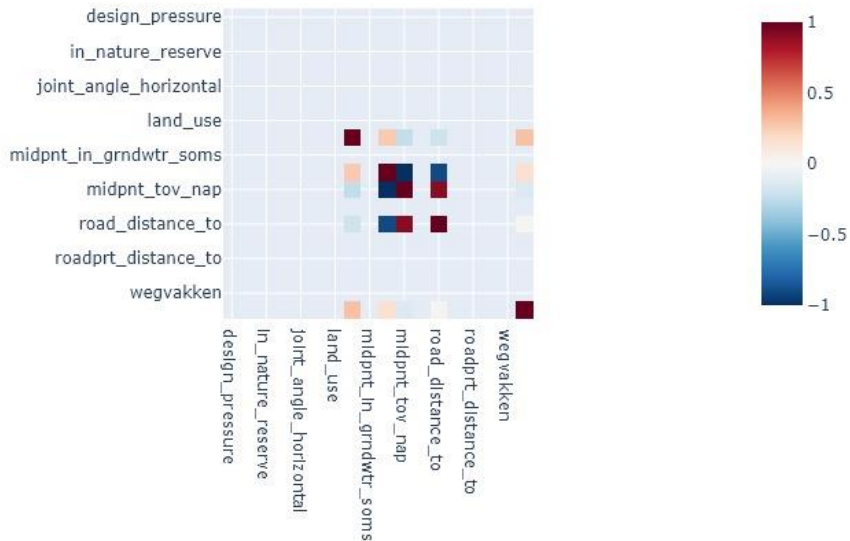
Naam subcategorie: medium, transportleiding, in stad, altijd in grondwater; 0: Nee, 1: Ja.

drinking-water.0.0.0



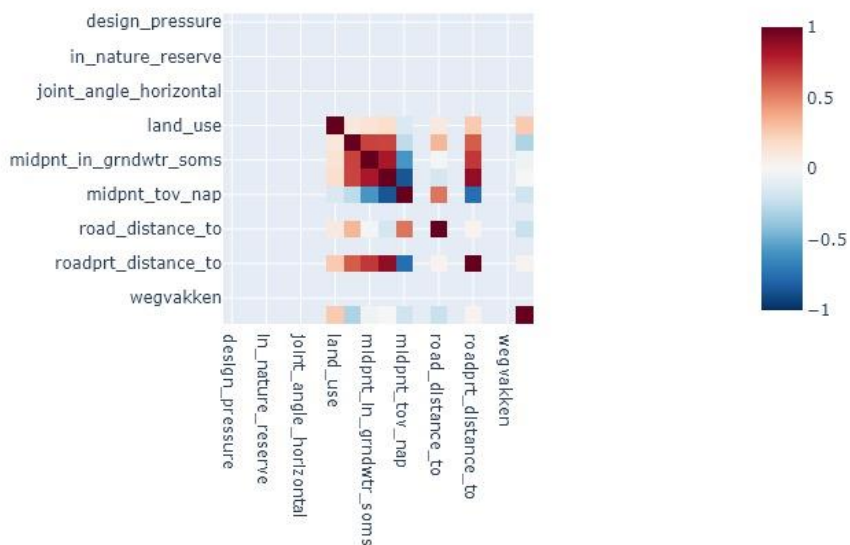
Figuur 21: Correlatiecoëfficiënt van categorie: drinkwater, geen transportleiding, niet in stad, niet altijd in grondwater

drinking-water.0.1.1



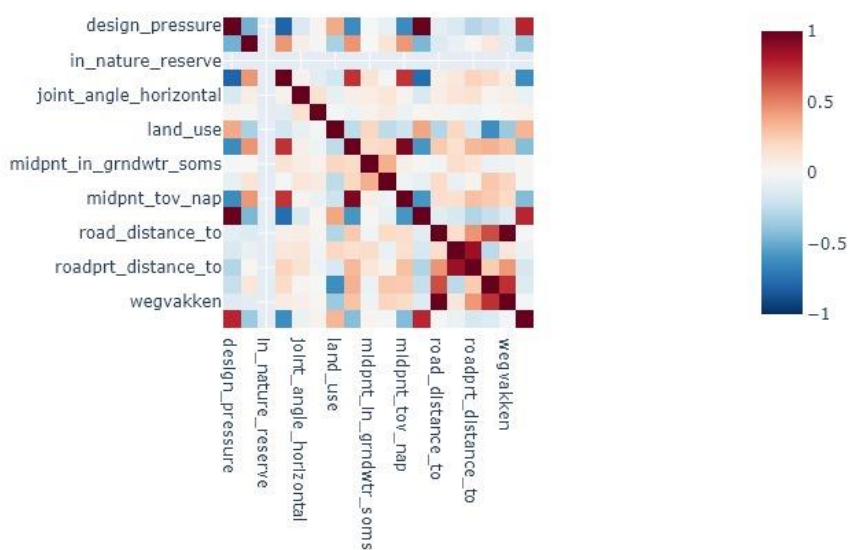
Figuur 22: Correlatiecoëfficiënt van categorie: drinkwater, geen transportleiding, is in stad, is altijd in grondwater

drinking-water:1.0.0



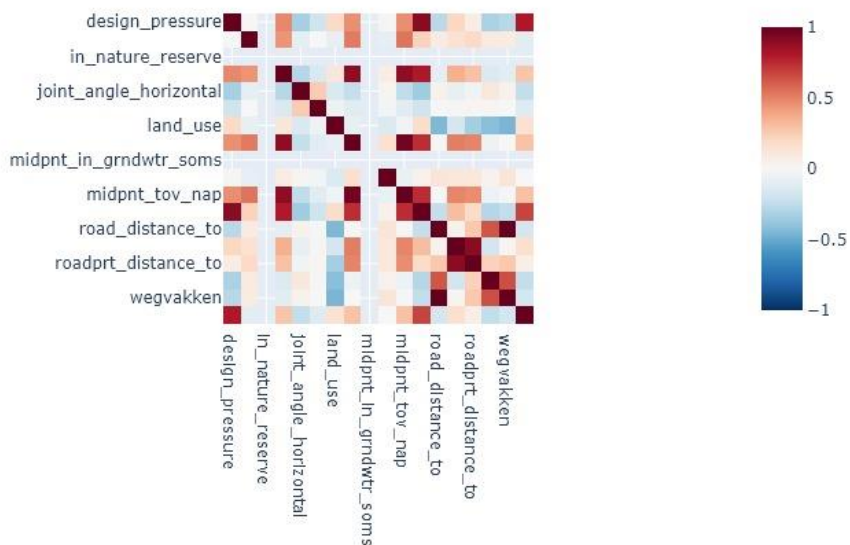
Figuur 23: Correlatiecoëfficiënt van categorie: drinkwater, is transportleiding, niet in stad, niet altijd in grondwater

waste-water:0.0.0



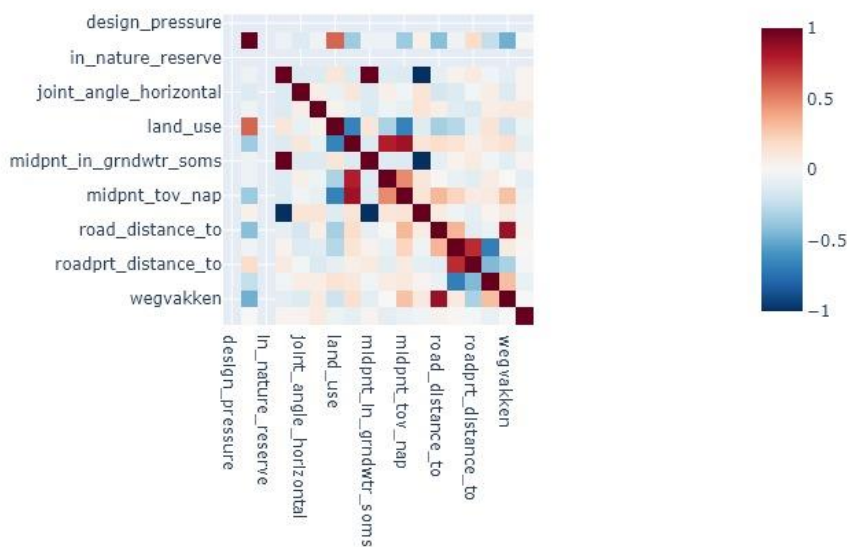
Figuur 24: Correlatiecoëfficiënt van categorie: afvalwater, geen transportleiding, niet in stad, niet altijd in grondwater

waste-water.0.0.1



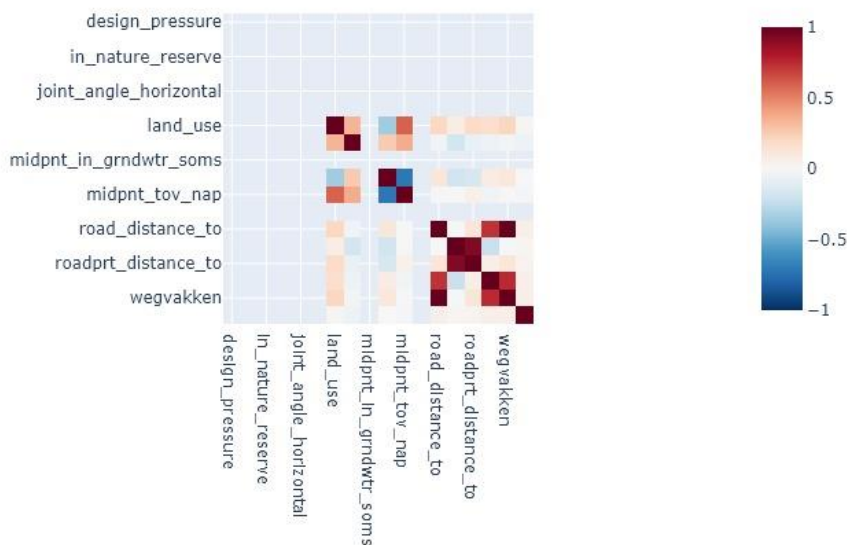
Figuur 25: Correlatiecoëfficiënt van categorie: afvalwater, geen transportleiding, niet in stad, zit altijd grondwater

waste-water.0.1.0



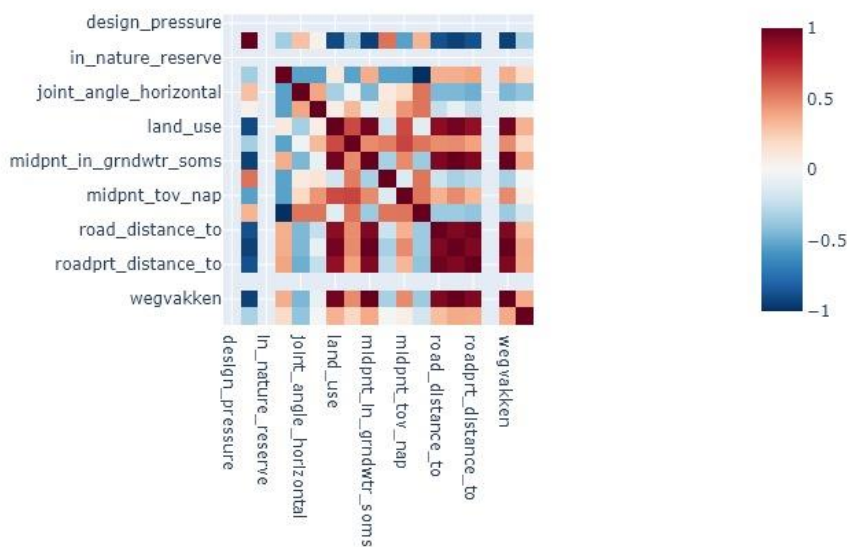
Figuur 26: Correlatiecoëfficiënt van categorie: afvalwater, geen transportleiding, is in stad, niet altijd in grondwater

waste-water.0.1.1



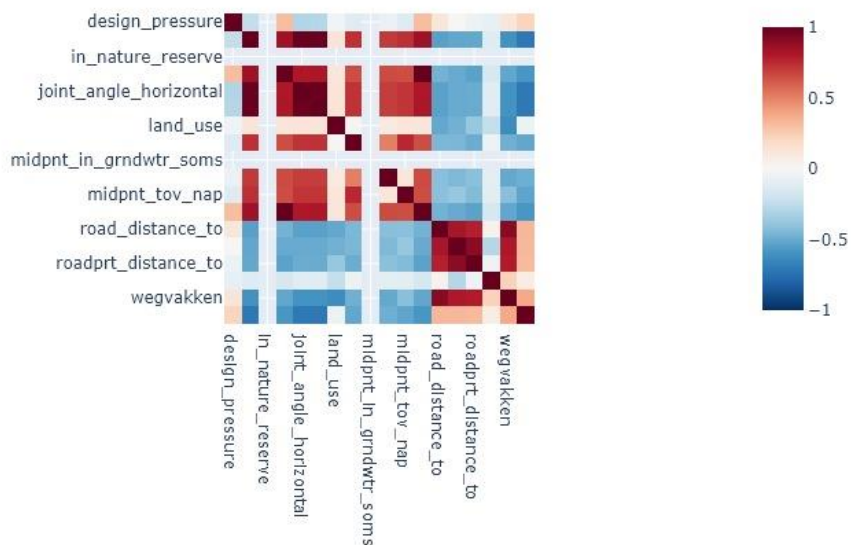
Figuur 27: Correlatiecoëfficiënt van categorie: afvalwater, geen transportleiding, bevindt zich in de stad, bevindt zich altijd in het grondwater

waste-water.1.0.0



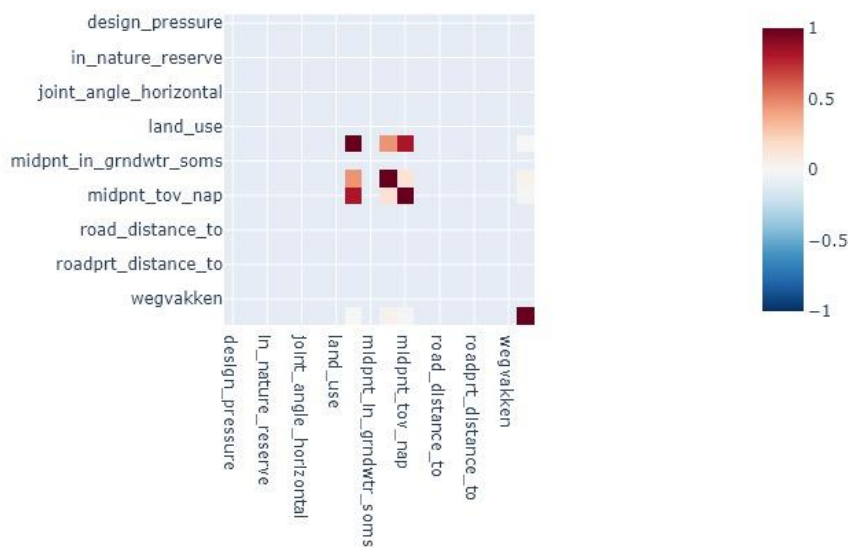
Figuur 28: Correlatiecoëfficiënt van categorie: afvalwater, is transportleiding, niet in stad, niet altijd in grondwater

waste-water.1.0.1



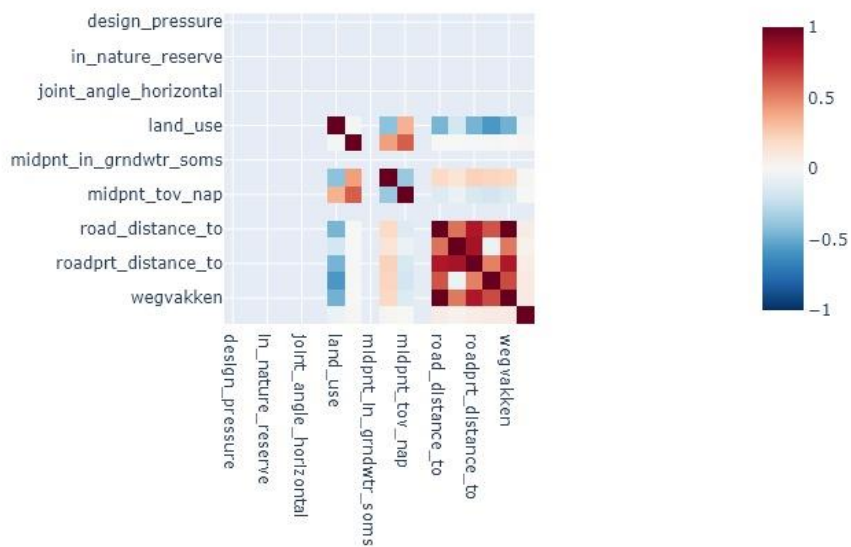
Figuur 29: Correlatiecoëfficiënt van categorie: afvalwater, is transportleiding, niet in stad, zit altijd in grondwater

waste-water.1.1.0



Figuur 30: Correlatiecoëfficiënt van categorie: afvalwater, is transportleiding, is in stad, niet altijd in grondwater

waste-water.1.1.1



Figuur 31: Correlatiecoëfficiënt van categorie: afvalwater, is transportleiding, is in stad, altijd in grondwater