



# ***Backtracing van verontreinigingen in het distributienet***

**Methoden en bronbepaling met  
stochastische watervraag**

**BTO 2013.030  
April 2013**

# ***Backtracing van verontreinigen in het distributienet***

## **Methoden en bronbepaling met stochastische watervraag**

**BTO 2013.030  
April 2013**

© 2013 KWR

Alle rechten voorbehouden.

Niets uit deze uitgave mag 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, zonder voorafgaande schriftelijke toestemming van de uitgever.

# Colofon

**Titel**

Backtracing van verontreinigingen in het distributienet - Methoden en bronbepaling met stochastische watervraag

**Opdrachtnummer**

B111791-001

**Rapportnummer**

BTO 2013.030

**Onderzoeksprogramma**

BTO (PBC Waterdistributie)

**Projectmanager**

Nellie Slaats

**Opdrachtgever**

CvO

**Kwaliteitsborger**

Mirjam Blokker

**Auteurs**

Peter van Thienen en Dirk Vries

**Verzonden aan**

Dit rapport is verspreid onder BTO-participanten en is openbaar

# Samenvatting

Waterkwaliteitssensoren geven een lokaal beeld van waterkwaliteitsparameters in het distributienet. Door de gegevens van deze sensoren in te voeren in hydraulische modellen is belangrijke informatie te krijgen over stroomopwaarts gelegen delen van het netwerk. Hydraulische modellen kunnen worden ingezet om met behulp van *backtracing*-algoritmes de effectiviteit van verschillende configuraties sensornetwerken te bepalen. Ook kunnen ze bij een verontreinigingscalamiteit worden ingezet om nauwkeuriger te bepalen waar een mogelijke bron zich bevindt. Dit is bijvoorbeeld uitgewerkt voor leidingnetgebieden in Eindhoven en Almere.

## **Belang: data waterkwaliteitssensoren optimaal benutten**

Als een verontreiniging in het leidingnet wordt ontdekt, is het zaak snel te handelen om de consequenties te minimaliseren. Wanneer een waterkwaliteitssensor een verontreiniging meet, is een van de hoofdvragen waar deze verontreiniging vandaan komt. Als dit is vastgesteld, kunnen maatregelen worden getroffen om de verdere verspreiding te beperken en om de oorzaak aan te pakken. Door vóór installatie het "blikveld" van mogelijke sensorconfiguraties te bepalen, kan bovendien een optimale configuratie worden ontworpen en kan worden voorspeld hoe goed dat cluster sensoren zal functioneren bij het opsporen van de oorzaak van vervuilingen.

## **Aanpak: *backtracing* in hydraulische modellen**

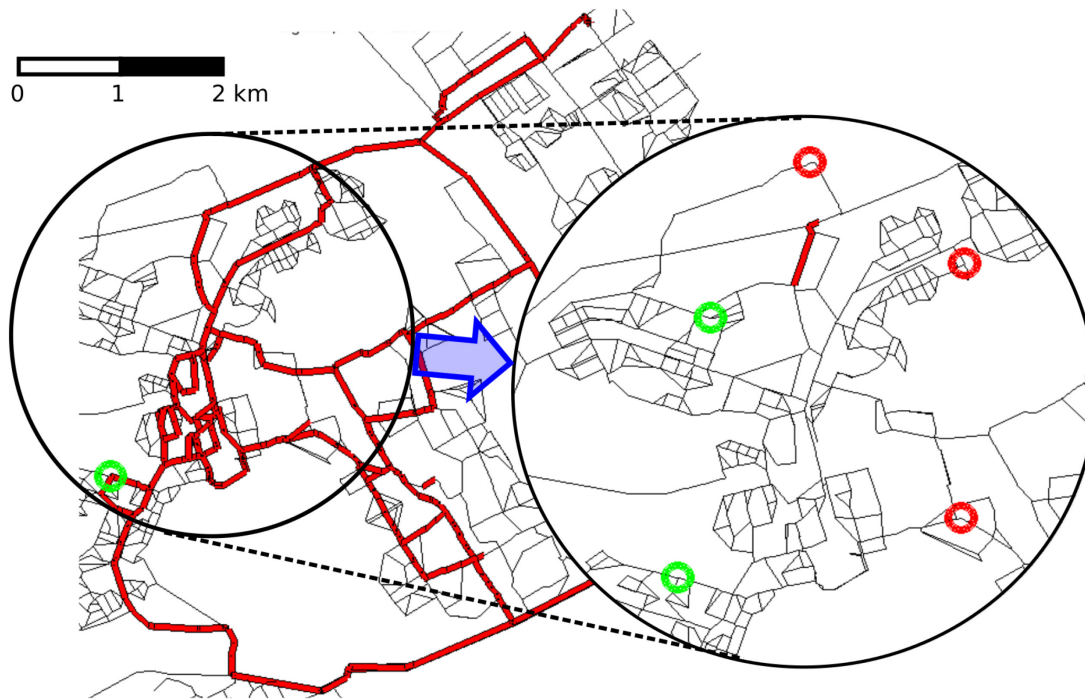
Hydraulische simulaties zijn een krachtig hulpmiddel om de verspreiding van een stof door een netwerk te berekenen. Na een inventarisatie van beschikbare benaderingen zijn twee benaderingen uitgewerkt voor *backtracing*, het terug in de tijd volgen van water (verontreinigingen) in een netwerkmodel. Na uitwerking en implementatie bleek een van de twee geschikt om direct toe te passen op echte netwerkmodellen. Daarnaast is een methodiek uitgewerkt voor het combineren van observaties op meerdere punten in een netwerk: als een verontreiniging op meerdere plaatsen in het netwerk wel of niet wordt geconstateerd, kan de combinatie van deze constatering leiden tot een snelle en eenvoudige inperking van het mogelijke brongebied. Deze combinatiemethodiek is toegepast op netwerkmodellen van Eindhoven en Almere.

## **Resultaten: dekkingskaarten en nauwkeurig lokaliseren calamiteit**

Met de ontwikkelde methodieken zijn dekkingskaarten gemaakt voor geselecteerde combinaties van meetlocaties in de proefgebieden. Deze dekkingskaarten geven duidelijk weer of een verontreiniging die in een bepaald deel van het netwerk optreedt door een of meerdere sensoren kan worden gedetecteerd bij de gekozen locaties voor de waterkwaliteitssensoren. Met een beperkt aantal bepalingen op verschillende locaties in het leidingnet blijkt het potentiële herkomstgebied van een verontreiniging terug te brengen is naar slechts enkele segmenten van een leidingnetmodel. De combinatie van sensoren met hydraulische modellen (inclusief *backtracing* en bronzoekmethoden) verkleint hiermee de onzekerheid over wat er in het netwerk gebeurt. Essentieel hierbij is actuele en accurate kennis van het distributiesysteem: de beschikbaarheid van een accuraat leidingnetmodel, actuele kennis van afsluiterstanden en bedrijfsvoering.

## **Implementatie: zinvolle plaatsing sensoren en opsporen vervuiliingsbronnen**

De toegevoegde waarde van waterkwaliteitssensoren in het distributienet in combinatie met hydraulische modellen is in de eerste plaats dat deze combinatie zinvolle plaatsing van waterkwaliteitssensoren in het distributienetwerk mogelijk maakt. Op basis van metingen van dergelijke slim geplaatste sensoren, kan bij calamiteiten bovendien worden bepaald op welk deel van het netwerk de aandacht moet worden geconcentreerd voor het opsporen van een vervuiliingsbron. Zo kan de bron van een vervuiling sneller worden gelokaliseerd en dus de verspreiding beter worden ingeperkt en de gevolgen daarmee geminimaliseerd. De ontwikkelde methodiek is opgezet rondom het modelleringsprogramma EPANET. Voor toepassing bij en door de waterbedrijven zelf is implementatie in de bij hen gebruikte hydraulische software noodzakelijk.



*Figuur S1: Bij een vervuiling in het leidingnet is het gebied waar de bron van de vervuiling zich mogelijk bevindt (rode leidingnetsegmenten) in te perken door de gegevens van waterkwaliteitssensoren in te voeren in een hydraulisch model. Met de gegevens van de waterkwaliteitssensoren in dit uitgebreide potentiële brongebied (groen cirkeltje = vervuiling aanwezig; rood cirkeltje = geen vervuiling) blijkt het meest waarschijnlijke brongebied veel kleiner te zijn (rode segmenten in de grote cirkel). Zo is de oorzaak van een calamiteit in het distributienetwerk gemakkelijker op te sporen.*

# Inhoud

<b>Samenvatting</b>	<b>1</b>
<b>Inhoud</b>	<b>3</b>
<b>1 Inleiding</b>	<b>5</b>
1.1 Aanleiding en doel	5
1.2 Aanpak	5
1.3 Leeswijzer	6
<b>2 Probleembeschrijving en bestaande methodieken</b>	<b>7</b>
2.1 Probleem	7
2.2 Gepubliceerde benaderingen in de wetenschappelijke literatuur	8
2.2.1 Raamwerken en mate van zekerheid	8
2.2.2 Lagrangiaanse en verwante benaderingen	8
2.2.3 Euleriaanse/optimalisatiebenaderingen	9
2.3 Uitdagingen	9
2.4 Implementaties in commerciële en vrij beschikbare software	10
<b>3 Ontwikkelde en toegepaste benaderingen</b>	<b>13</b>
3.1 Inleiding	13
3.2 Benadering 1: Lagrangiaanse backtracing op basis van deterministisch voorwaarts stoftransport	13
3.3 Benadering 2: Backtracing door probabilistische inversie	18
3.3.1 Herkomstkaarten en herkomstkaskaarten voor een enkele sensor	20
3.3.2 Combinaties van sensoren	20
3.3.3 Dekkingskaarten	22
3.3.4 Beschermingsgebieden	28
3.4 Gebruikte en ontwikkelde software	28
<b>4 Toepassing van de ontwikkelde methodieken</b>	<b>31</b>
4.1 Inleiding	31
4.2 Uitgangspunten en onderzoeksvragen	31
4.3 Testgebieden	31
4.4 Resultaten Eindhoven	34
4.4.1 Variabiliteit van het stromingsveld	34
4.4.2 Herkomst en dekking	35
4.4.3 Nadere beschouwing van stochastiek	40
4.4.4 Bronbepaling	41
4.5 Almere	44
4.5.1 Variabiliteit van het stromingsveld	44
4.5.2 Herkomst en dekking	44
4.5.3 Bronbepaling	48

4.6	Vergelijking van Eindhoven en Almere	48
4.7	Gevoeligheidsbeschouwing	50
4.7.1	Afsluitermanipulatie	50
4.7.2	Tijdstip van meting	50
4.8	Nut en noodzaak	53
<b>5</b>	<b>Conclusies en aanbevelingen</b>	<b>55</b>
5.1	Conclusies	55
5.1.1	Benaderingen	55
5.1.2	Toepassing van Lagrangiaanse methodiek op Eindhoven en Almere	55
5.2	Nut en noodzaak	56
5.3	Aanbevelingen	56
5.3.1	Voor toepassing	56
5.3.2	Voor vervolgonderzoek	56
<b>6</b>	<b>Referenties</b>	<b>59</b>
I	Verklarende woordenlijst	61
II	Opzet van backtracing met voorwaarts Lagrangiaanse benadering	63
III	Discreet mathematisch-fysisch model voor stoftransport met probabilistische benadering	65
IV	Dekkingskaarten	72
V	Gebruikte software	74
VI	Test van implementatie	76
VII	Vorbereiding van de hydraulische modellen van Eindhoven en Almere	82

# 1 Inleiding

## 1.1 Aanleiding en doel

Mede vanuit waterkwaliteitsbewakingsoogpunt bestaat de wens om continu een actueel beeld van diverse waterkwaliteitsparameters in het volledige distributienet te hebben. Gezien de omvang van distributie-infrastructuur is het echter niet te verwachten dat het distributienet op korte termijn volledig is voorzien van sensoren. Ook willen waterbedrijven op basis van nu nog een beperkt aantal sensormetingen na een alarm op een sensor snel kunnen vaststellen waar een eventuele verontreiniging in het net vandaan komt. Het doel van dit project is te onderzoeken hoe op basis van een beperkt aantal metingen het mogelijke brongebied van een geconstateerde verontreiniging is in te perken, door de stroming van het water tegen de tijd in te volgen in een numeriek model. In de Engelstalige literatuur wordt dit aangeduid met *backtracing*. Dit is een iets preciezere en beperktere omschrijving van het bij aanvang geformuleerde doel, namelijk hoe op basis van een beperkt aantal metingen een beeld te verkrijgen is van de toestand van het waterkwaliteitsparameters overal in een leidingnet.

Hiertoe worden twee benaderingen beschouwd: 1) een Lagrangiaanse benadering, waarin virtuele deeltjes worden gevolgd terug in de tijd, en 2) een inversie-benadering, waarin een krachtige inversiemethode wordt toegepast om voor het hele netwerk tegelijkertijd deze *backtrace* te bewerkstelligen. Met deze inversiemethode is binnen het BTO in een heel andere toepassing ervaring opgedaan (Van Thienen en Floris, 2012). Op deze manier wordt een beeld van de waterkwaliteit stroomopwaarts van de sensoren verkregen. In geval van een besmetting kan de bron hiervan snel worden gelocaliseerd, waardoor de gevolgen voor de consument en het imago van het waterbedrijf door snel ingrijpen beperkt kunnen worden.

Het stromingsveld in een drinkwaterdistributienet is zeer variabel van aard, zowel in de ruimte als in de tijd. Met name in de periferie van het net komen stagnatie en omkering van de stromingsrichting vaak voor. Deze variabiliteit wordt tot dusver nauwelijks in hydraulische modellen meegenomen, maar is voor de doelstelling van dit project van groot belang. De stroomsnelheden in het netwerk door de dag heen worden in dit project gebaseerd op een gedetailleerd hydraulisch model. Voor het genereren van de verbruikspatronen die het hydraulische model aandrijven wordt gebruik gemaakt van SIMDEUM® (Blokker, 2010), dat in een reeks van simulaties de stochastische aard (variantie) van de stroomsnelheden eveneens goed kan kwantificeren. Vanwege de stochastische aard van de watervraag kunnen meerdere stromingsmodellen voor ieder moment van de dag worden gemaakt, die ieder een bepaalde waarschijnlijkheid van realisatie hebben.

Het goed kunnen *backtracen* is tevens relevant voor de vraag waar men het beste waterkwaliteitssensoren in het distributienet kan plaatsen. Laird et al. (2005) en De Sanctis et al. (2010) wijzen er terecht op dat de optimalisatie-doelstellingen die in software als TEVA-SPOT (Berry et al., 2006, 2010) worden toegepast alle gerelateerd zijn aan de impact van een eventuele besmetting. Er wordt geen functionaliteit geboden om een optimale configuratie voor een zo nauwkeurig mogelijke bepaling van een besmettingsbron in de tijd en de ruimte (grootste pakkans) te bepalen. Het is daarom een secundaire doelstelling van dit rapport om handvatten te bieden voor de bepaling van optimale sensorlocaties met specifiek deze doelstelling.

In dit rapport wordt veel gesproken over *sensoren*. Dit woord moet in brede zin worden opgevat. Het omvat niet alleen fysieke sondes die in het distributienet zijn geïnstalleerd die waarden van chemische of fysische grootheden constateren en communiceren, maar ook monsternamen-analysetrajecten en zelfs meldingen van klanten.

## 1.2 Aanpak

Allereerst is in de wetenschappelijke literatuur gezocht naar methoden voor het *backtracen* van geconstateerde verontreinigingen van een constatering naar de bron. Hiernaast is een methodiek opgezet voor het combineren van waterpaden van individuele sensoren tot kaarten die iets zeggen over de



mogelijke herkomst van verontreinigingen. Vervolgens zijn de geselecteerde en ontwikkelde methoden geïmplementeerd in onderzoekstools, gebouwd rondom EPANET-MSX (Shang et al., 2008). Deze zijn tenslotte toegepast op netwerkmodellen van Eindhoven en Almere. Afgevaardigden van Brabant Water, Dunea, PWN en Vitens hebben deel uitgemaakt van de projectgroep.

### **1.3 Leeswijzer**

Hoofdstuk 2 biedt een overzicht en beschouwing van in de literatuur beschreven methoden voor *backtracing*, inclusief de uitdagingen waar men voor wordt gesteld. Ook worden de mogelijkheden op dit gebied van commercieel en anderszins beschikbare software op dit vlak op een rijtje gezet. Hoofdstuk 3 beschrijft de in dit project ontwikkelde en toegepaste methoden. Ten behoeve van de leesbaarheid en toegankelijkheid is ervoor gekozen om gedetailleerde uitwerkingen en tests onder te brengen in bijlagen. In Hoofdstuk 0 worden de beschreven technieken toegepast op netwerkmodellen van de distributienetten van Almere en Eindhoven. Hier worden de bruikbaarheid en beperkingen van de methodieken onder praktische omstandigheden onderzocht. Ten slotte biedt Hoofdstuk 5 een overzicht van conclusies en aanbevelingen die volgen uit dit werk. Dit rapport bevat het nodige jargon, dat toegelicht wordt in de verklarende woordenlijst van Bijlage I.

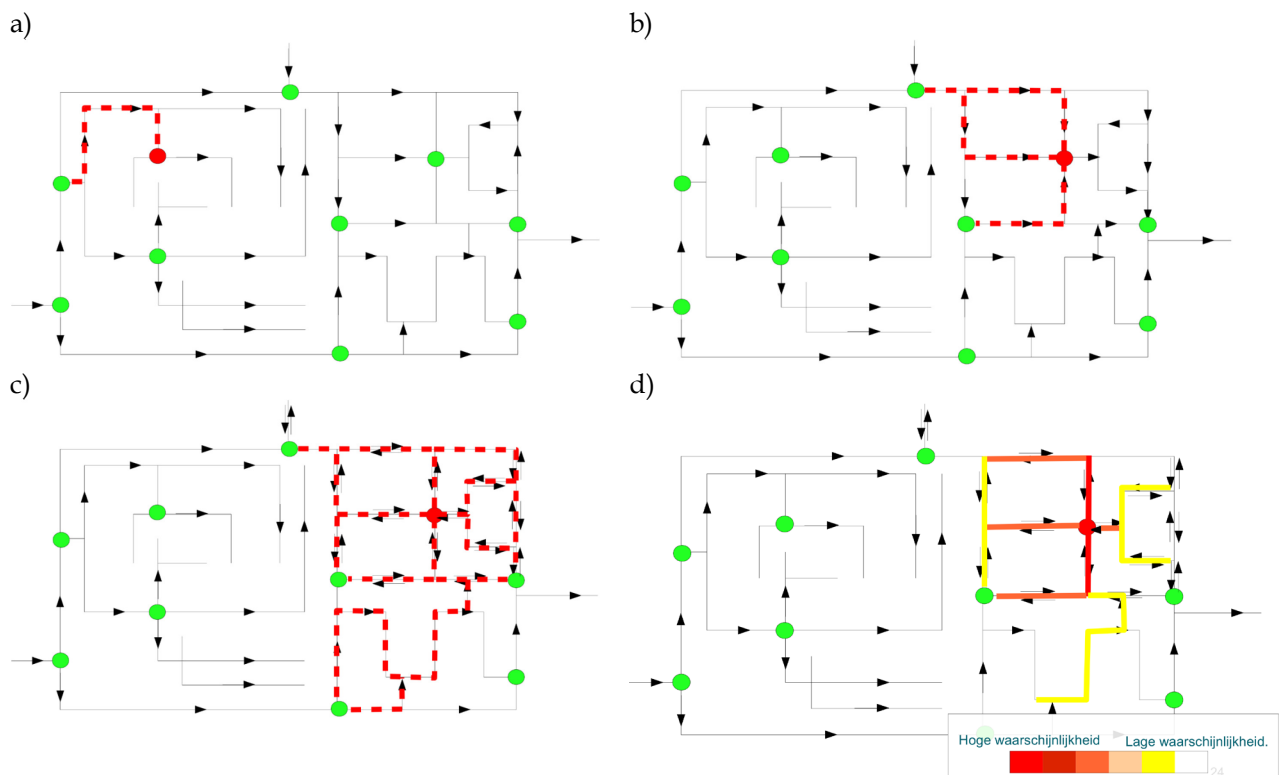
## 2 Probleembeschrijving en bestaande methodieken

### 2.1 Probleem

Wanneer een sensor een afwijkende concentratie voor een bepaalde stof meet en andere sensoren in de omgeving deze afwijking niet constateren, kan het in sommige gevallen gemakkelijk zijn om te achterhalen waar deze stof het netwerk binnen is gekomen, met name in vertakte netten (Figuur 1a). Men volgt simpelweg de stroming in tegenovergestelde richting totdat men een sensor tegenkomt die niet de betreffende verontreiniging meet. De afgelegde weg komt overeen met het mogelijke brongebied. In vermaasde netten is dit vaak echter al minder recht-toe-recht-aan (Figuur 1b). Het mogelijke brongebied omvat nu alle wegen naar de sensor toe. Wanneer we ons bewust worden van het feit dat stroomsnelheden en soms ook -richtingen gedurende de dag kunnen veranderen en van dag op dag ook verschillend kunnen zijn, wordt het mogelijke herkomstgebied van een geconstateerde verontreiniging mogelijk vele malen groter (Figuur 1c).

Variaties in stroomsnelheid- en richting kunnen een gevolg zijn van:

- reguliere variaties in watervraag door de dag heen;
- stochastische variaties in de watervraag;
- veranderingen in de bedrijfsvoering;
- afsluitermanipulaties.



Figuur 1: Het reconstrueren van een besmettingsbron op basis van afwijkende (rood) en niet-afwijkende (groen) sensormetingen in een vertakt net (a), een vermaasd net met constante stroomrichting (b), en een vermaasd net met variërende stroomrichting (c). In het laatste geval kan een herkomstkanskaart(d) helpen om zoekacties naar de besmettingsbron te prioriteren.

Met name in het tertiaire netwerk leiden stochastische variaties op basis van SIMDEUM® in modellen tot beduidend grotere en meer pendelzones en significante stagnatieperioden in vergelijking met modellen zonder een stochastische component in de watervraag (Blokker, 2010).

In de praktijk betekent dit dat er een grote onzekerheidsmarge zit op iedere schatting die wordt gemaakt van een instantaan snelheidsveld in een (tertiair) distributienet. Op basis van stochastische simulaties kan echter wel een kansverdeling van snelheidsvelden worden berekend. Deze kan worden gebruikt om in te schatten hoe waarschijnlijk het is dat een geconstateerde verontreiniging van een bepaalde plaats in het netwerk afkomstig is. Dergelijke inschattingen kunnen worden gecombineerd in een herkomstkanskaart, zoals schematisch weergegeven in Figuur 1d.

## 2.2 Gepubliceerde benaderingen in de wetenschappelijke literatuur

### 2.2.1 Raamwerken en mate van zekerheid

Er bestaan verschillende conceptuele en mathematische benaderingen voor het *backtrace*probleem. Ten eerste kan een mathematisch model hiervoor in twee verschillende raamwerken worden opgesteld. Het voor dergelijke toepassingen meestgebruikte is het Lagrangiaanse raamwerk. Hierin is de in het netwerk bewegende materie het referentiekader en wordt de beweging van deeltjes (waterpakketjes) door een vastliggend netwerk beschreven. Enkele uitwerkingen van deze en aanverwante benaderingen worden kort beschreven in §2.2.2. Hiernaast bestaat ook het Euleriaanse raamwerk, waarbij het vastliggende netwerk juist het referentiekader is. Uitwisseling van stof tussen vastliggende punten in het netwerk wordt hierin beschreven. Enkele uitwerkingen hiervan worden genoemd in §2.2.3.

Voor het probleem van stoftransport geldt, dat als men exact weet hoeveel van een stof waar in een systeem aanwezig is, en men het exacte stromingsveld kent van de vloeistof die deze stof meevoert, en bovendien precies weet hoe de stof zich chemisch gedraagt (conservatief of met model voor reacties), dat men dan exact kan voorspellen hoeveel van deze stof zich in de toekomst waar bevindt. Anders gezegd: stoftransport is op deze manier een deterministisch probleem. In de praktijk zijn er echter diverse onzekerheden, bijvoorbeeld in de sensoren (vals-positieven en vals-negatieven, detectielimieten, etc.) en het stromingsveld (en de variatie daarvan in de tijd). We kunnen, om het voorbeeld van de onzekerheid in het stromingsveld te illustreren, wel een heleboel plausibele stromingsvelden bedenken, en daar wellicht ook een kans van realisatie aan hangen. Voor ieder van deze stromingsvelden zal het transportprobleem tot een ander resultaat, namelijk een andere stofverdeling in de toekomst, leiden. Het combineren van al deze resultaten leidt tot een kansverdeling voor de stofverdeling in de toekomst. Deze beschouwing met onzekerheden is een probabilistische benadering. Een uitwerking van een probabilistische benadering voor meerdere parallele waterstromen in een netwerk wordt kort beschreven in §2.2.3.

### 2.2.2 Lagrangiaanse en verwante benaderingen

Voortbouwend op werk van Zierolf et al. (1998) hebben Shang et al. (2003) een *particle backtracing algorithm* (PBA) ontwikkeld, dat in staat is om in een hydraulisch model de diverse paden van waterpakketjes te volgen in de tijd terug van een punt naar allerlei punten in het leidingnet. Dit algoritme gaat uit van een Lagrangiaanse raamwerk.

Preis en Ostfeld (2006) vervangen het hydraulische model dat de basis vormt voor verspreidingsberekeningen door zogenaamde *modelbomen*, die door training en validatie het hydraulische gedrag van een distributienetwerk kunnen representeren. Hoewel de methode robuust en betrouwbaar blijkt, is zij erg duur in termen van rekentijd.

Een manier om alle potentiële besmettingsbronnen (locatie en tijdstip) te identificeren is ontwikkeld door De Sanctis et al. (2010). Zij kijken hierbij niet naar concentraties, maar uitsluitend naar of een sensor wel of geen afwijking meet: een binaire benadering. Alle sensormetingen worden met een *particle backtracking* (PBA)-algoritme teruggerekend in de tijd om alle mogelijke tijdstippen en plaatsen van de oorsprong van iedere sensormeting te bepalen. Vervolgens wordt gekeken welke lokatie-tijdstipcombinaties wel (positieve contaminatiestatus) aan sensormetingen gekoppeld kunnen worden en welke niet (negatieve status). Dit algoritme heet Contamination Status Algorithm of CSA. Dit levert, uitgaande van perfecte

metingen en een perfect hydraulisch model, alle potentiële besmettingsbronnen (locatie en tijdstip). Het resultaat van deze methode kan als opbrengst op zichzelf worden beschouwd, of als invoer (basisselectie) voor een nauwkeurigere bepaling met een inversiebenadering.

Costa et al. (2011) presenteren een benadering waarin 3 algoritmes worden gecombineerd, namelijk een algoritme dat alle mogelijke verspreidingspaden van een (virtuele) besmetting berekent, een algoritme dat alle mogelijke besmettingsbronnen evalueert die een afwijkende sensormeting zouden kunnen verklaren, en tenslotte een algoritme voor de lokalisering van de besmettingsbron op basis van de metingen van meerdere sensoren. De eerste twee algoritmen komen overeen met de benadering van Shang et al. (2002).

### **2.2.3 Euleriaanse/optimalisatiebenaderingen**

Een voor de hand liggende benadering (De Sanctis et al., 2010) om afwijkende sensormetingen die indicatief zijn voor een verontreiniging terug te rekenen naar de bron is in de vorm van een optimalisatieprobleem. Het probleem is doorgaans als volgt opgesteld:

- de besmettingsbron is de te optimaliseren, onafhankelijke variabele;
- de optimalisatiedoelstelling is het minimaliseren van het verschil tussen de gemodelleerde en gemeten stofconcentraties.

Deze benadering kent echter enkele beperkingen, die worden besproken in §2.3. Vanwege deze beperkingen is de benadering verfijnd in Laird et al. (2005, 2006). Zij passen daarbij twee stappen toe:

- In de eerste stap wordt in zekere zin het scheermes van Ockham toegepast door te stellen dat besmettingsscenario's met een klein aantal simultane besmettingen waarschijnlijker zijn dan die met een groter aantal simultane besmettingen, en zich vervolgens uitsluitend te richten op de scenario's met een zo klein mogelijk aantal besmettingspunten.
- In de tweede stap wordt vervolgens gezocht naar alle minimale besmettingsscenario's die in redelijke mate in overeenstemming zijn met de sensormetingen.

Een probabilistische benadering is gezocht door Propato et al. (2010). Deze auteurs beperken eerst het aantal mogelijke besmettingsbronnen met behulp van lineaire algebra. Vervolgens passen zij een Bayesiaanse, stochastische methode toe voor de inversie, de zogenaamde *Minimum Relative Entropy (MRE)*-methode. Deze methode berekent voor iedere potentiële besmettingsbron een waarschijnlijkheid dat deze de daadwerkelijke bron is.

## **2.3 Uitdagingen**

Gezien het beperkte aantal metingen ten opzichte van het grote aantal potentiële besmettingsbronnen is *backtracing* in een optimalisatieformuleringen mathematisch gezien een lastig op te lossen probleem (Laird et al., 2005, De Sanctis et al., 2010). Een veelgebruikte benadering om voor dit soort, zogenaamde slecht-geconditioneerde vraagstellingen tot een mathematische oplossing te komen is *regularisatie* (bijvoorbeeld in Laird et al., 2005, 2006 of Neupauer, 2010), waarmee randvoorwaarden van mathematische aard (dus niet gerelateerd aan het fysieke distributienet en de processen daarin) bijdragen tot het vinden van een oplossing. Er is echter geen garantie dat de zo gevonden oplossing ook daadwerkelijk een relevante oplossing is voor het gestelde fysische probleem (De Sanctis et al., 2010).

Andere uitdagingen bij terugrekenmethoden in het algemeen zijn (De Sanctis et al., 2010):

- het enorme aantal potentiële locaties en tijdstippen van besmetting, waarmee bij een terugrekenactie rekening moet worden gehouden;
- de mate waarin sensoren daadwerkelijk in staat zijn om alle besmettingen op ieder concentratieniveau te meten;
- de correctheid van sensormetingen en -interpretaties (vals-positieven en vals-negatieven);
- onzekerheden in netwerkmodellen (materialen en diameters, wandruwheid, verbruikspatronen, lekken, afsluiters, etc.).

Ook wordt in de regel uitgegaan van een conservatieve stof, terwijl deze aanname in veel gevallen niet terecht is.

## **2.4 Implementaties in commerciële en vrij beschikbare software**

*Backtrace*functionaliteit is reeds beschikbaar in zowel commerciële als vrij beschikbare software. Tabel 1 geeft hiervan een overzicht, inclusief een beknopte weergave van de geboden functionaliteit. Ter vergelijking is ook de functionaliteit van de in dit rapport gebruikte methoden opgenomen.

Tabel 1: Overzicht van software met backtracingfunctionaliteit.

<i>software</i>	<b>Infoworks</b>	<b>InfoWater BTX</b>	<b>Synergie</b>	<b>EPANET-BTX</b>	<b>Netsim</b>	<b>WaterWiSe</b>	<b>dit rapport (1)</b>	<b>dit rapport (2)</b>
<i>beschrijving</i>	hydraulische simulatie-software; functie: upstream network trace	module bovenop ArcGIS	hydraulische simulatie-software	backtracing-extensie van EPANET; dit is een functie-bibliotheek voor gebruik in combinatie met de EPANET-toolkit	In het kader van het Disconto-project is software voor backtracing ontwikkeld voor inpassing in Netsim. Deze integratie heeft echter nog niet plaats-gevonden	Platform voor de real-time monitoring van distributiesystemen	combinatie van voorwaarts stoftransport met EPANET-MSX	probabilistische inversie
<i>geïmplementeerde methodiek</i>	Lagrangiaans?	Lagrangiaans?	Lagrangiaans?	Lagrangiaans	Lagangiaans (gebaseerd op EPANET-BTX)	Lagrangiaans?	voorwaarts Lagrangiaans	Euleriaans
<i>direct toepasbaar door waterbedrijven</i>	ja	ja	ja	vergt schrijven van software	onbekend (integratie in Netsim nog niet voltooid)	ja	niet in huidige vorm	niet in huidige vorm
<i>enkelvoudig eindpunt (sensor)</i>	ja	ja	ja	ja	ja	ja	ja	ja
<i>meervoudige eindpunten (sensoren)</i>	nee	geen gegevens	ja	kan worden geïmplementeerd	geen gegevens	geen gegevens	ja	ja
<i>gecombineerde interpretatie van meerdere eindpunten</i>	nee	geen gegevens	nee	kan worden geïmplementeerd	geen gegevens	geen gegevens	ja	ja
<i>stochastische watervraag</i>	nee	nee	nee	nee	nee	nee?	ja	ja
<i>relatieve bijdrage paden</i>	geen gegevens	ja	ja	kan worden geïmplementeerd	geen gegevens	geen gegevens	ja	nee
<i>opmerking</i>			<i>backtracing</i> alleen steady state					
<i>beschikbaarheid</i>	commercieel	commercieel	commercieel	vrij te downloaden	commercieel/ Disconto	commercieel	BTO	BTO
<i>webadres</i>	www.innovyze.com	www.innovyze.com	http://www.gl-group.com/en/water/SynerGEEW ater.php	http://sourceforge.net/projects/epanet-btx/	www.ureason.com	http://www.visenti.com/waterwise.html		
<i>bron</i>	Infoworks 11.5 handleiding	website	Synergie 4.5.1. User Guide	documentatie	contact met UReason	website		



# 3 Ontwikkelde en toegepaste benaderingen

## 3.1 Inleiding

In dit rapport is een tweetal benaderingen onderzocht. Ten eerste is een Lagrangiaanse benadering uitgewerkt en geïmplementeerd in software. Deze wordt in het volgende hoofdstuk toegepast op twee netwerkmodellen. Tevens is een Euleriaanse inversiebenadering verkend en uitgewerkt. Deze is nog niet toegepast in een praktijkcasus.

## 3.2 Benadering 1: Lagrangiaanse backtracing op basis van deterministisch voorwaarts stoftransport

Stel dat we op ieder punt in een distributienetwerk een voor dat punt  $P_i$  karakteristieke merkstof  $M_i$  aan het water zouden kunnen toevoegen. Wanneer zich ergens anders een verontreiniging voordoet, kan door de aan- of afwezigheid van deze specifieke stof  $M_i$  worden bepaald of (een deel van) het water op deze plek ooit langs het betreffende punt  $P_i$  in het leidingnet is gekomen. Dit geeft dus aan of het punt  $P_i$  een mogelijke bron van de geconstateerde verontreiniging kan zijn.

Hydraulische modellen stellen ons in staat om deze exercitie daadwerkelijk uit te voeren met behulp van de computer. Op ieder moment wordt op ieder knooppunt  $i$  in een netwerkmodel (Figuur 2a,b) zoveel voor dat punt karakteristieke stof  $M_i$  (*locatietracer*) in het water gestopt (Figuur 3a), dat de concentratie uitkomt op 1 mg/l (dit getal is arbitrair; belangrijk is dat het gelijk is voor alle knooppunten ten behoeve van de bepaling van mengverhoudingen). Bovendien wordt er een hoeveelheid van een eveneens voor knooppunt  $P_i$  karakteristieke stof  $T_i$  (*tijdstracer*) toegevoegd (Figuur 3b), zodanig dat de concentratie hiervan uitkomt op  $h$  mg/l, waarbij  $h$  de verstreken tijd in uren is sinds de start van de simulatie. EPANET-MSX transporteert deze stoffen (in een Lagrangiaans raamwerk) gedurende een simulatie door het netwerk en berekent voor ieder simulatietijdstip de concentraties van al deze stoffen  $M_i$  en  $T_i$  op alle knooppunten  $P_i$  van het netwerk.

Als we nu op knooppunt  $P_j$  een verontreiniging constateren, kunnen we, door te kijken welke stoffen  $M_i$  aanwezig zijn op dit tijdstip op dit knooppunt, vaststellen langs welke knooppunten  $P_i$  het water is gestroomd voordat het op knooppunt  $P_j$  aankwam (Figuur 4a). Hiermee hebben we een volledig overzicht van de potentiële locaties van de besmettingsbron verkregen. Bovendien geven de onderlinge verhoudingen van de concentraties van  $M_i$  weer welke fractie van het water langs welk knooppunt is gekomen. Als we veronderstellen dat de *a priori* kans op een besmetting even groot is bij elk van deze punten, representeren de verhoudingen van volumestroomfracties kansverhoudingen van de besmettingsbronlocatie. Zo kan er uit de concentratieverhoudingen direct een (relatieve) herkomstkansdichtheidskaart worden gemaakt.

In de meeste gevallen is het bovendien mogelijk om uit te rekenen op welk tijdstip het water dat nu wordt bemonsterd op  $P_j$  langs  $P_i$  is gekomen door de concentratie van de tijdstracer  $T_i$  te delen door de concentratie van de locatietracer  $M_i$  (Figuur 4b). Hier zit echter een addertje onder het gras. Wanneer water gelijktijdig via verschillende routes van  $P_i$  naar  $P_j$  reist en als mengsel van deze verschillende routes  $P_j$  bereikt, is het op deze manier verkregen tijdstip een naar volumestromen (van de verschillende routes) gewogen gemiddelde van het tijdstip van passeren van  $P_i$  voor deze verschillende routes (let op dat deze beperking specifiek is voor de hier beschreven methode). Hoewel de passagetijden in dit rapport verder niet worden gebruikt, bieden zij potentieel een extra handvat voor het vinden van een besmettingsbron.

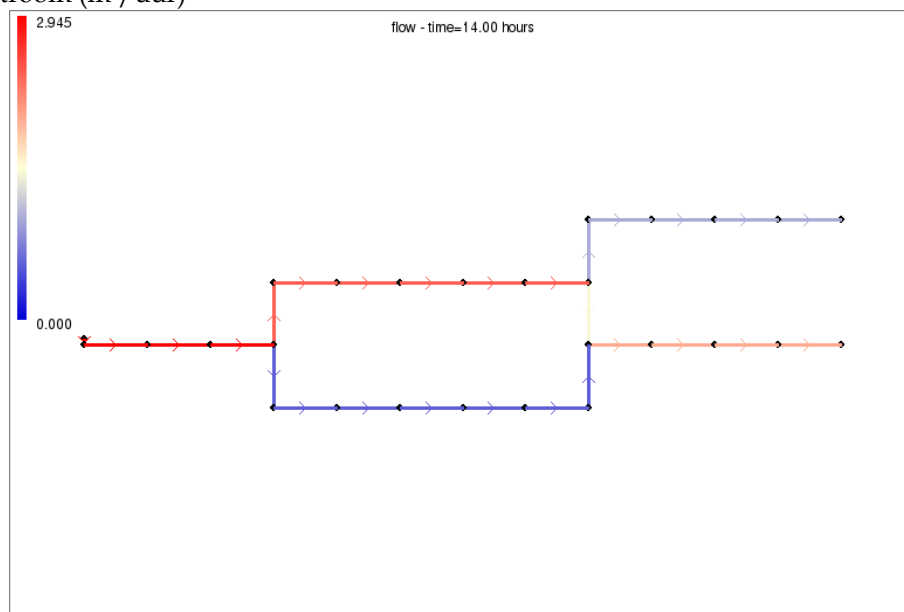
In een stochastische beschouwing van de wereld (Figuur 5a) hebben we voor ieder verbruikspunt een oneindig aantal potentiële vraagpatronen, waarvan er iedere dag maar één wordt gerealiseerd. Door het hydraulische model een aantal keer door te rekenen met steeds andere vraagpatronen uit deze verzameling van plausible patronen met behulp van SIMDEUM®, kan een representatieve verzameling



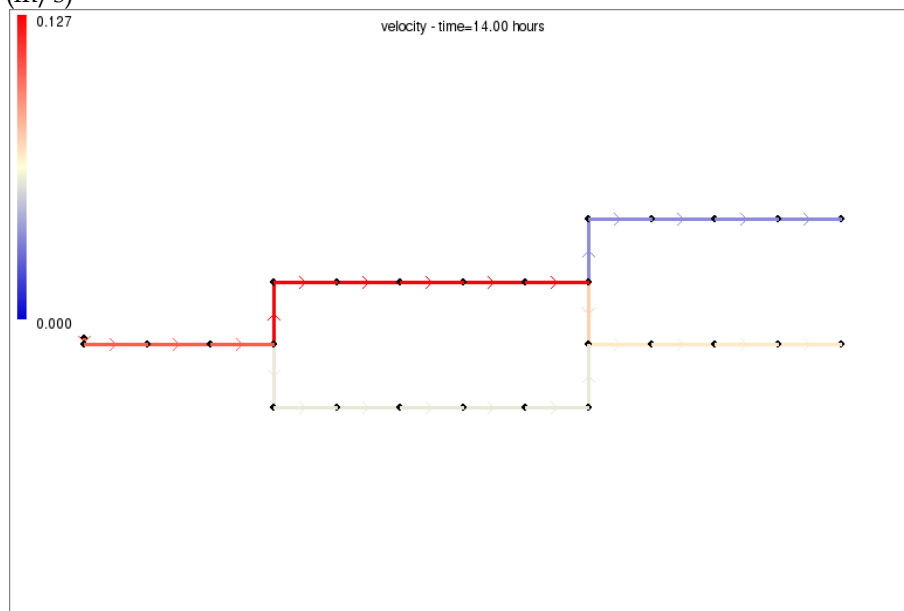
snelheidsvelden worden verkregen. De hierboven beschreven procedure kan voor ieder van deze snelheidsvelden (een zogenaamde *realisatie*) worden herhaald. Als we een uniforme (of andere) kansverdeling voor deze verzameling snelheidsvelden veronderstellen, kunnen we vervolgens de resultaten van de individuele berekeningen combineren tot een stochastische herkomstkansdichtheidskaart (Figuur 5b).

De technische uitvoering van het hierboven beschreven proces is weergegeven in Bijlage II.

a) Volumestroom ( $\text{m}^3/\text{uur}$ )

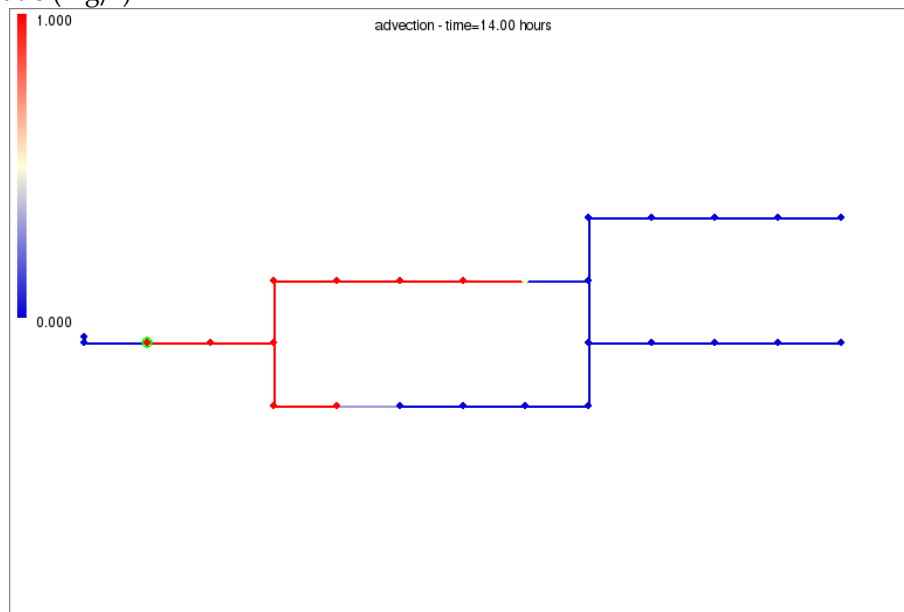


b) Snelheid ( $\text{m/s}$ )

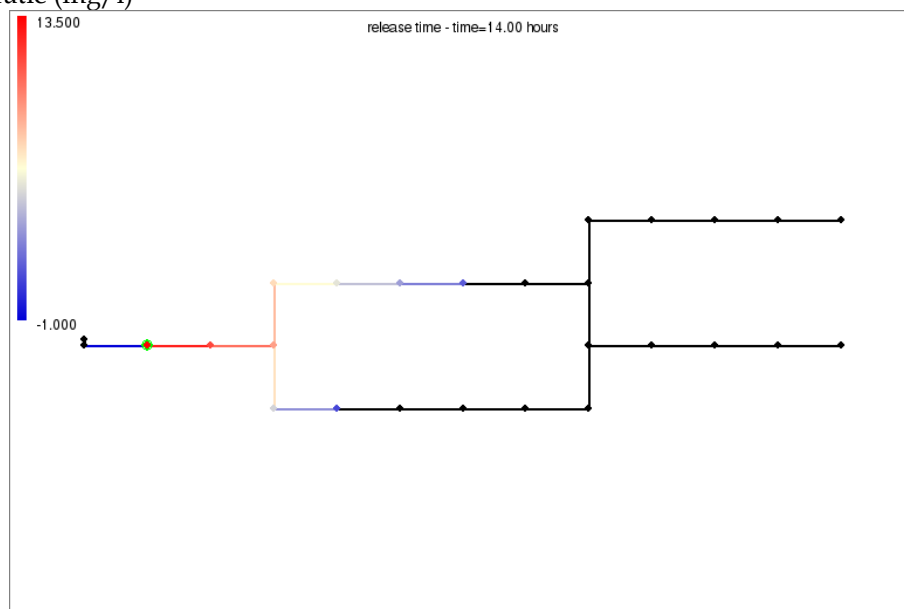


Figuur 2: a) Volumestromen ( $\text{m}^3/\text{uur}$ ) en b) snelheidsveld ( $\text{m/s}$ ) in een netwerkmodel ter illustratie van de uitgewerkte Lagrangiaanse methode voor backtracing.

a) Concentratie (mg/l)

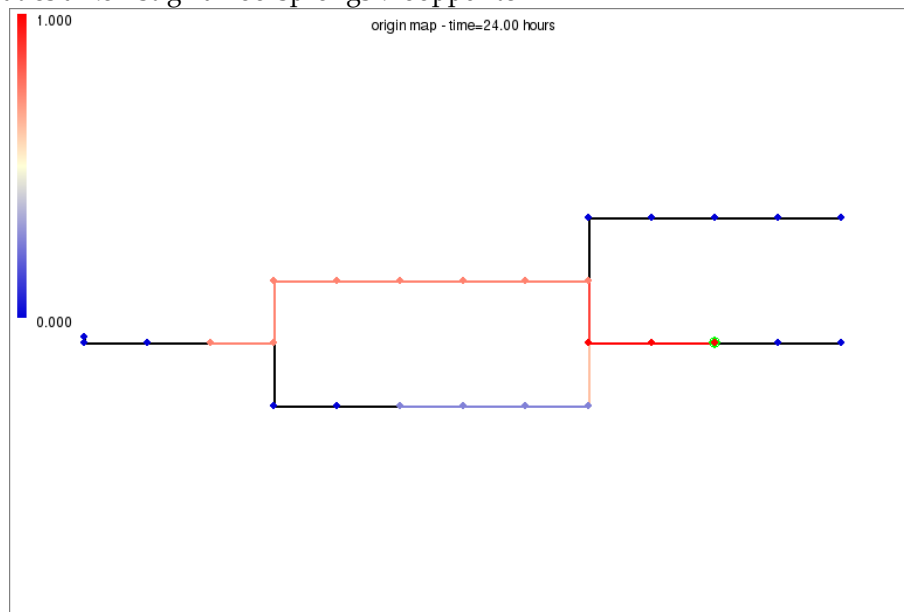


b) Concentratie (mg/l)

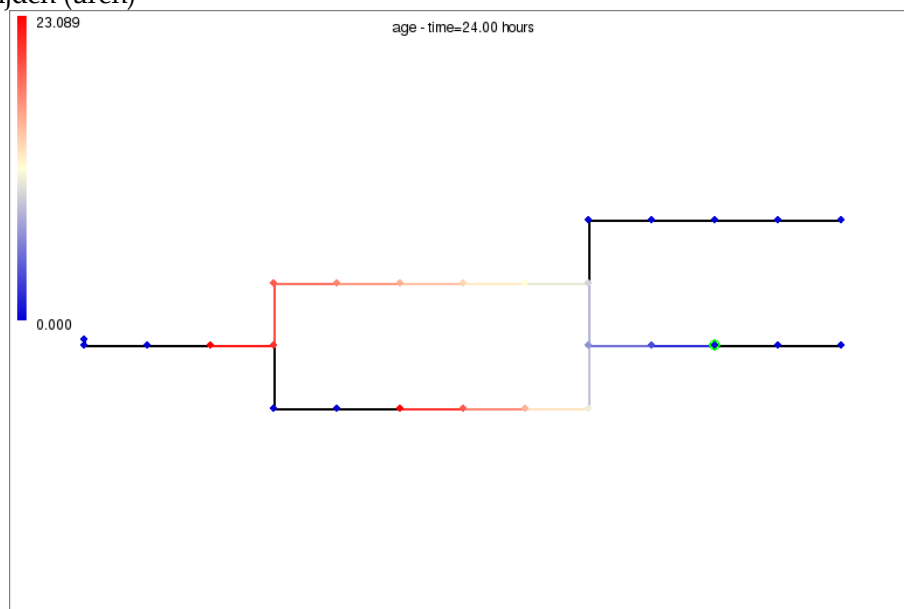


Figuur 3: a) Concentratie (mg/l) van knooppuntspecifieke locatietraceren en b) van knooppuntspecifieke tijdstraceringen op tijdstip  $t=14$  uur in een netwerkmodel ter illustratie van de uitgewerkte Lagrangiaanse methode voor backtracing.

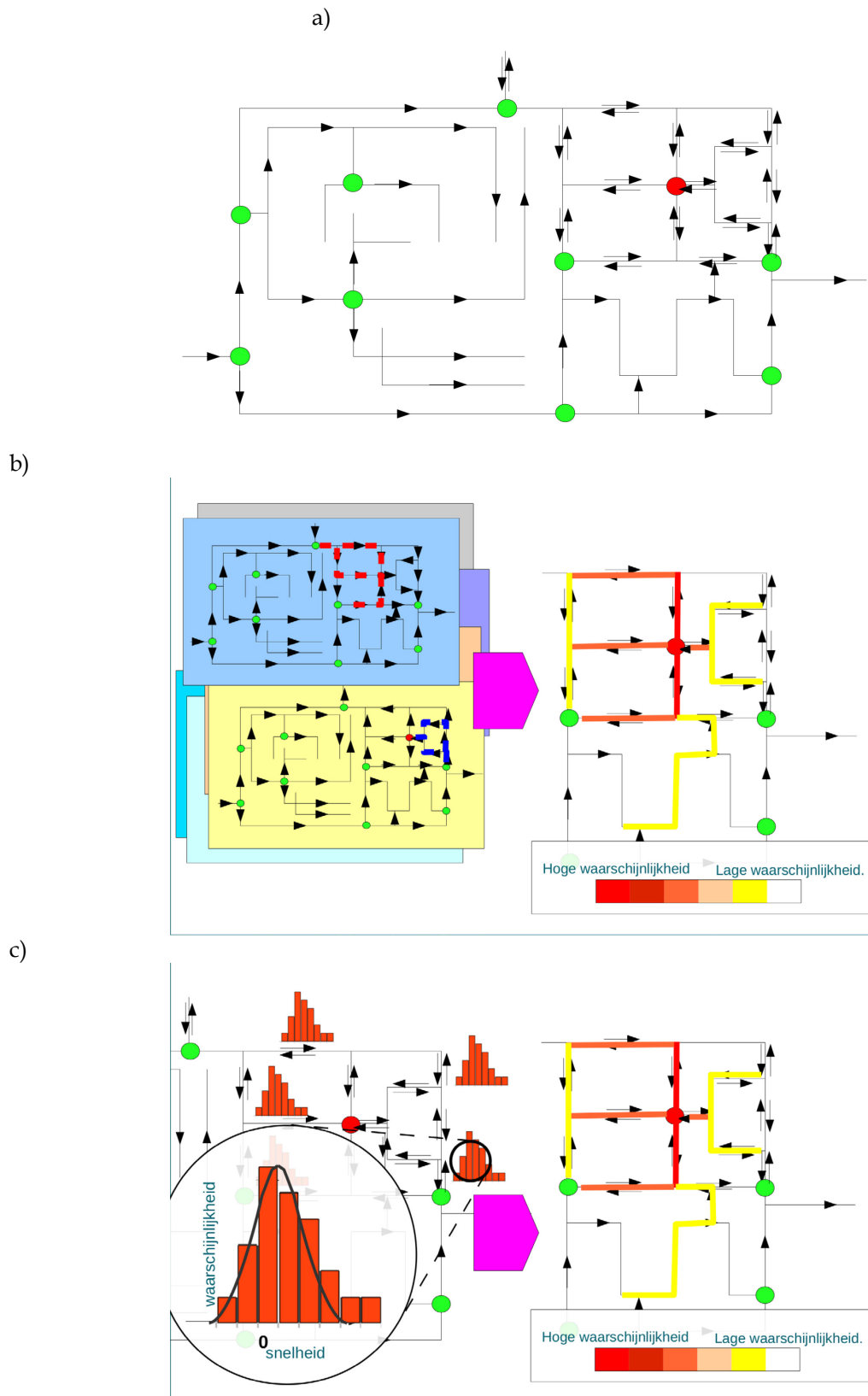
a) Concentraties afkomstig van oorsprongsknooppunten



b) Passagetijden (uren)



Figuur 4: s) Concentraties van knooppuntspecifieke tracers gevonden op het groen omcirkelde punt weergegeven op hun respectievelijke oorsprongsknooppunten en b) teruggerekende passagetijden (ten opzichte van nu, in uren) voor water wat nu langs het groen omcirkelde punt stroomt.



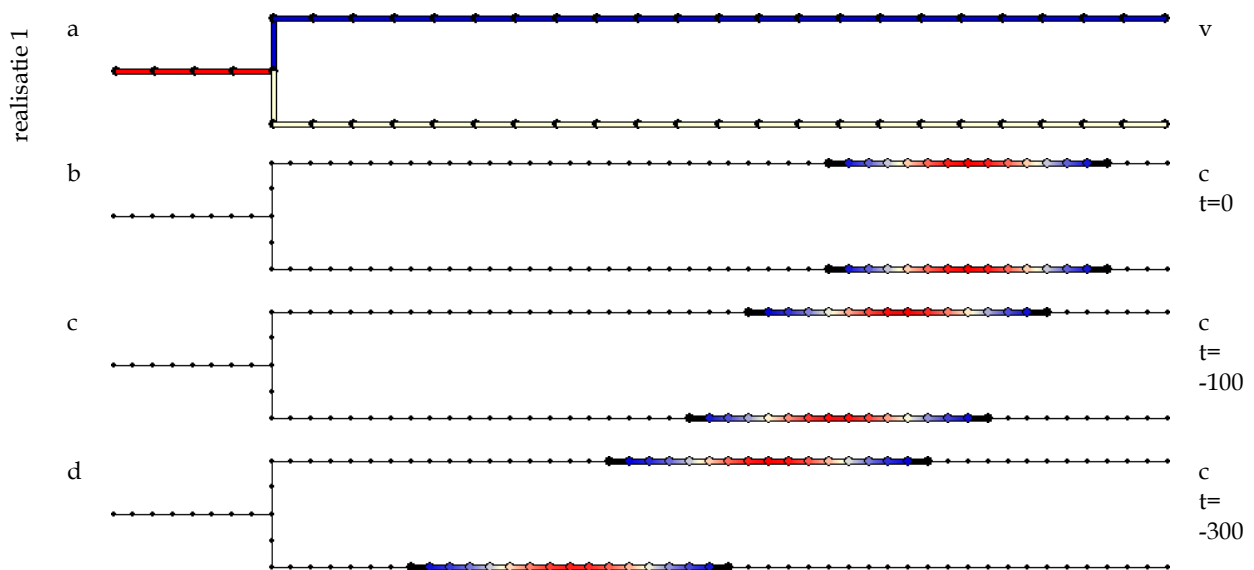
Figuur 5: De onzekerheid het stromingsveld (a) kan op verschillende manieren worden meegenomen in een backtraceactie. Ten eerste kan voor een reeks realisaties van het stromingsveld steeds deterministisch een backtrace worden gemaakt, waarna de resultaten worden gecombineerd tot een herkomstkanskaart (b). Ten tweede kan in een probabilistische inversiebenadering (c) de onzekerheid in het stromingsveld worden beschreven met een kansverdeling en wordt vervolgens met deze kansverdeling in de tijd terug gerekend.

### 3.3 Benadering 2: Backtracing door probabilistische inversie

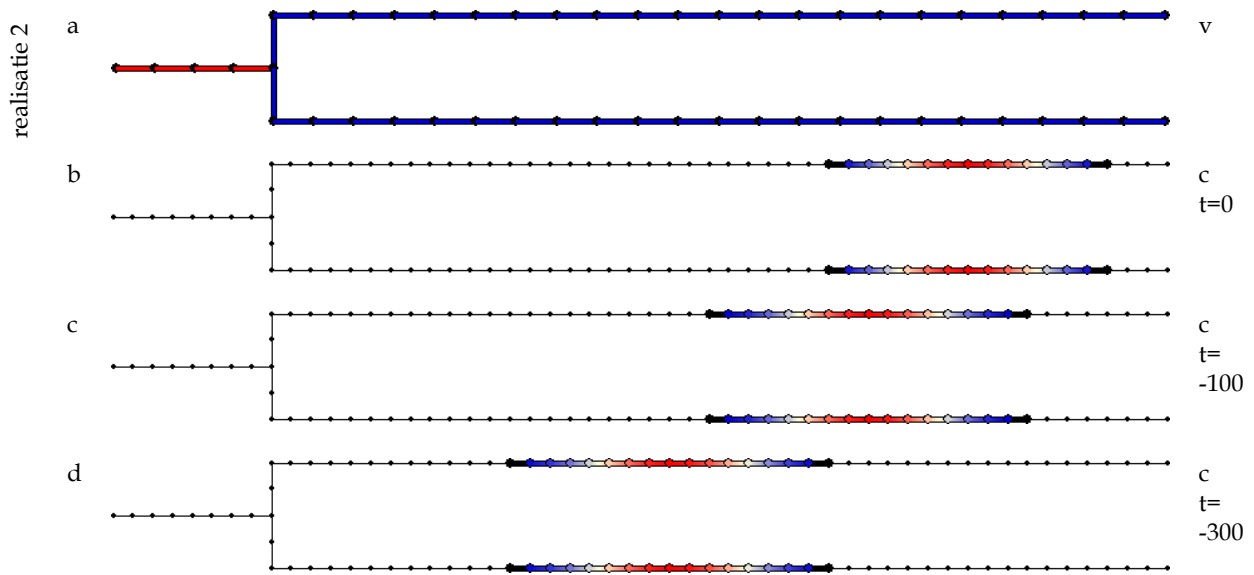
Omdat in een stochastische benadering van het stromingsveld er voor iedere realisatie een nieuwe set berekeningen moet worden uitgevoerd, kan de Lagrangiaanse benadering numeriek duur uitpakken. Dit probleem kan worden omzeild door een gemeenschappelijke beschrijving voor de verzameling van realisaties te bedenken en deze gemeenschappelijke beschrijving te gebruiken om het stoftransport in de tijd terug te berekenen (Figuur 5c). Voor deze benadering ligt de keuze voor een Euleriaans raamwerk, waarin het vastliggende distributienetwerk het referentiekader is, meer voor de hand dan een Lagrangiaanse benadering. We beschrijven nu de concentratie van een stof op ieder punt in het netwerkmodel en de uitwisseling van deze stof tussen deze punten op basis van het stromingsveld.

De mathematische uitwerking van deze benadering is beschreven in Bijlage III. De werking van de benadering wordt geïllustreerd aan de hand van Figuur 8 t/m Figuur 8. In een eenvoudig leidingnetwerkmodel loopt het water van links naar rechts door een enkele leiding, die zich splitst in twee leidingen. Water wordt uitsluitend onttrokken aan de rechter eindpunten van deze twee leidingen. In het netwerk is een variabel snelheidsveld, weergegeven in twee realisaties (Figuur 8a en Figuur 7a). Beide realisaties komen even vaak voor, en ze kunnen worden gecombineerd in een enkele weergave van gemiddelde (kleur van leidingsegmenten en knopen) en variantie (contouren van de knopen), zie Figuur 8a. Voor de individuele realisaties is goed te zien dat de concentratiepuls terug wordt getransporteerd met behoud van vorm en amplitude, al naar gelang de lokale snelheid.

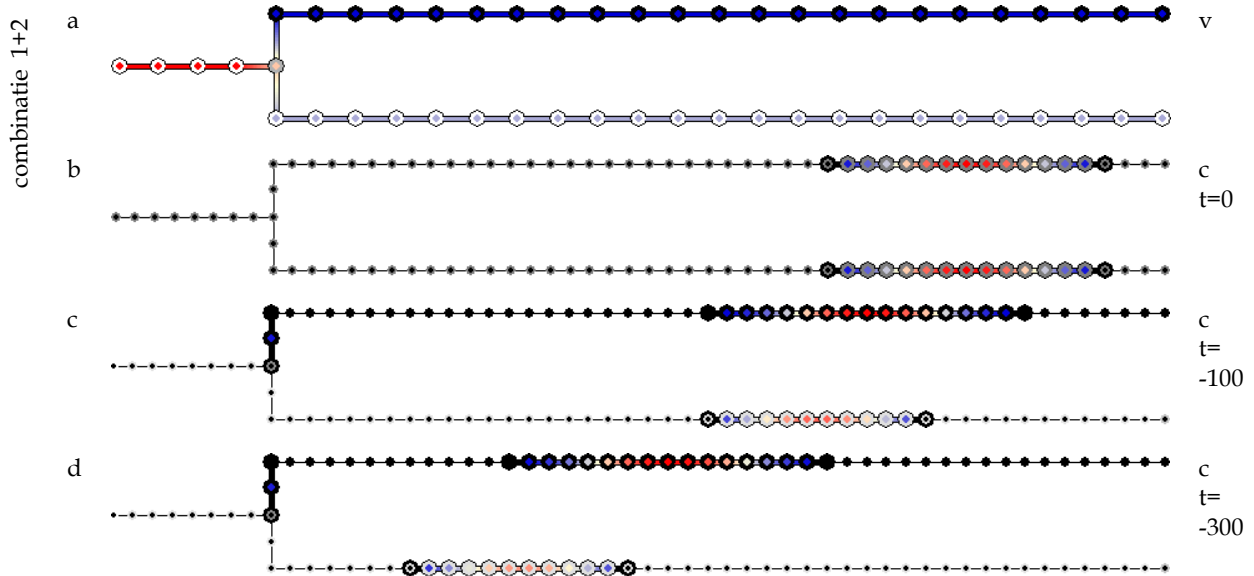
Bij de gecombineerde weergave is het beeld iets complexer. In de bovenste tak zien we een herhaling van het beeld uit de individuele realisaties. In de onderste tak, echter, zien we een sterke uitdemping van de puls door de grotere variantie in het snelheidsveld. Tegen de verwachting in wordt de puls echter ook smaller. Op dit moment hebben wij hier geen verklaring voor. Dit zal in mogelijk vervolgwerk verder moeten worden uitgezocht.



Figuur 6: Illustratie van backtracing d.m.v. probabilistisch inversie voor een realisatie van het stromingsveld. a) snelheidsveld van realisatie 1 (blauw=1, wit=2, rood=3); b) concentratieveld bij begin (zwart/blauw=0, rood=1); c) concentratieveld 100 stappen terug in de tijd; d) concentratieveld 300 stappen terug in de tijd. legenda: v=snelheid, c=concentratie; t=tijd.



Figuur 7: Illustratie van backtracing d.m.v. probabilistisch inversie voor een tweede realisatie van het stromingsveld. a) snelheidsveld van realisatie 2 (blauw=1, wit=2, rood=3); b) concentratieveld bij begin (zwart/blauw=0, rood=1); c) concentratieveld 100 stappen terug in de tijd; d) concentratieveld 300 stappen terug in de tijd. legenda: v=snelheid, c=concentratie; t=tijd.



Figuur 8: Illustratie van backtracing d.m.v. probabilistisch inversie voor een combinatie van de twee realisaties van het snelheidsveld. a) snelheidsveld (blauw=1, wit=2, rood=3); b) concentratieveld bij begin (zwart/blauw=0, rood=1); c) concentratieveld 100 stappen terug in de tijd; d) concentratieveld 300 stappen terug in de tijd. De varianties van de snelheid respectievelijk de concentratie zijn weergegeven in grijsgekleurde contouren van de knopen. Hierbij geeft een zwarte rand een lage en een witte rand een hoge variantie aan. legenda: v=snelheid, c=concentratie; t=tijd.

Zoals te zien in Bijlage III draait de inversie om de vergelijking

$$\mathbf{c}^t = \mathbf{C}_M \mathbf{B}^T (\mathbf{B} \mathbf{C}_M \mathbf{B}^T + \mathbf{C}_D)^{-1} \mathbf{c}^{t+\Delta t}$$

met  $c^t$  het uit te rekenen concentratieveld,  $C_D$  de covariantiematrix van het gehele systeem, met hierin de (co)variantie van het geobserveerde concentratieveld  $c_i^{t+At}$ , en de covariantie behorende bij het transportmodel. De matrix  $B$  beschrijft het stoftransport tussen de knooppunten, en is ijl (veel nullen), waardoor opslag en verwerking in de computer efficiënt is. Hetzelfde geldt voor de matrix  $C_M$ . De matrix  $C_D$  bevat varianties en covarianties van het stochastische snelheidsveld voor alle knooppunten (dimensie  $n \times n$ ). Doordat de variatie van de watervraag door de dag heen in grote lijnen overal gelijk opgaat (in de zin dat we overal o.a. een ochtend- en avondpiek zien), is de covariantie tussen vrijwel alle knooppunten meestal ongelijk aan 0. Dit betekent dat matrix  $C_D$  vrijwel volledig is gevuld. Voor een efficiënte omgang met deze matrix in de berekeningen mag de matrix, en daarmee het aantal knooppunten in het model, niet te groot zijn (maximaal orde  $10^4$  knooppunten). Let op dat het hierbij gaat om knooppunten van de discretisatie van het concentratieveld. Het maximaal aantal knopen in het hydraulisch model is een fractie hiervan. Dit punt beperkt de toepasbaarheid van de methodiek vooralsnog voor grotere netwerkmodellen en praktijktoepassingen.

### 3.3.1 Herkomstkaarten en herkomstkaskaarten voor een enkele sensor

Met de eerstbesproken, Lagrangiaanse benadering kan voor een enkele sensor eenvoudig voor een bepaald tijdstip zowel een herkomstkaart als een herkomstkaskaart worden gemaakt. De eerste (Figuur 9a) geeft het volledige gebied in het netwerk weer, waar het water dat nu langs de sensor stroomt ooit is geweest. De tweede (Figuur 9b) geeft bovendien aan welke fractie van het water dat langs de sensor stroomt door ieder punt van het netwerk is gekomen. Wanneer men veronderstelt dat de kans op een verontreiniging op ieder punt van het netwerk gelijk is, is deze fractie representatief voor de kans dat een door de sensor geconstateerde verontreiniging van het betreffende punt in het netwerk afkomstig is.

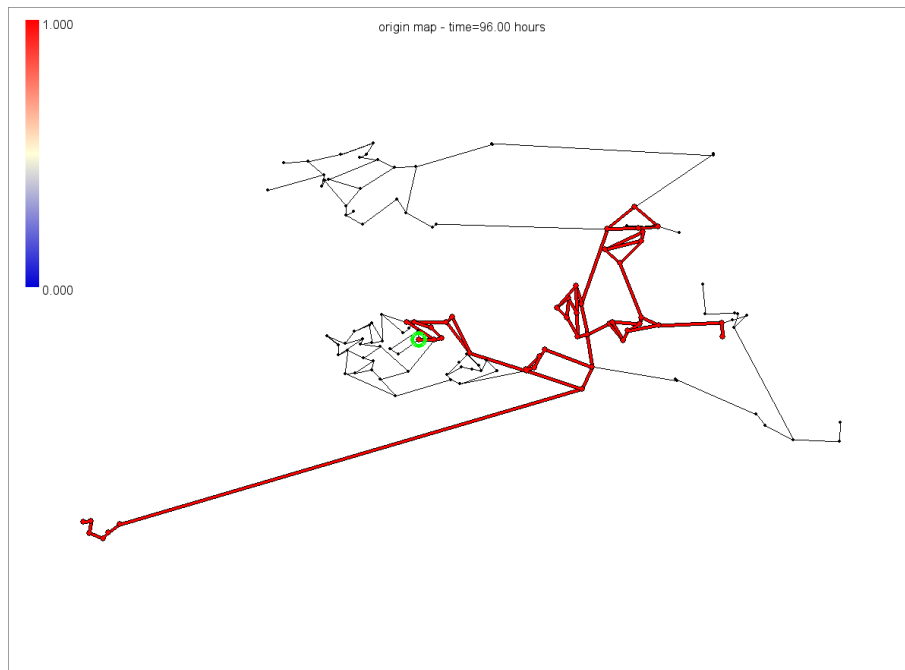
### 3.3.2 Combinaties van sensoren

Door informatie van meerdere sensoren te combineren, kan een veel beter beeld van de mogelijke bronlocatie van een verontreiniging worden verkregen. Dit concept wordt hier uitgewerkt voor de toegepaste Lagrangiaanse benadering van §3.2. Op twee manieren kunnen sensoren bijdragen:

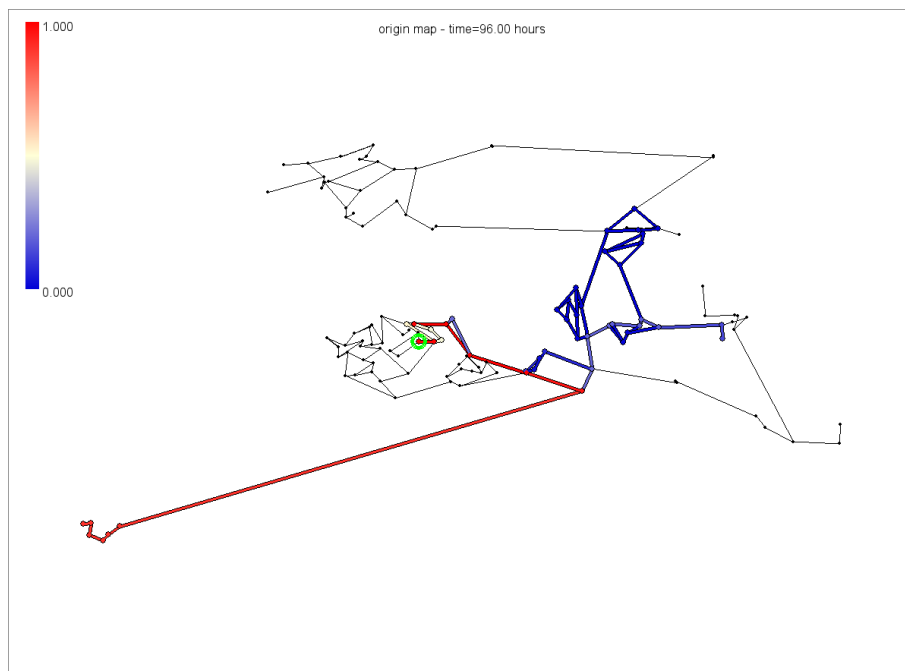
- Positief: als twee (of meer) sensoren (Figuur 10a,b) elk een verontreiniging meten en we veronderstellen dat het dezelfde verontreiniging betreft, moet de bronlocatie zich logischerwijs bevinden in de doorsnede (overlap) van de herkomstkaarten van de sensoren (Figuur 11b).
- Negatief: als een of meer sensoren (Figuur 10a,b) een verontreiniging meten maar een andere sensor (Figuur 11a) niet, en we opnieuw voor de positieve sensoren veronderstellen dat het dezelfde verontreiniging betreft, dan moet de bronlocatie zich bevinden in het verschil van de (gecombineerde) herkomstkaart van positieve en die van de negatieve sensor(en) (Figuur 12).

Let op dat we hierbij steeds veronderstellen dat de benodigde meetwaarden op hetzelfde tijdstip (*synchron*) verkregen zijn. Deze benadering leent zich echter eveneens voor de combinatie van metingen die niet op hetzelfde tijdstip zijn uitgevoerd (*diachroon*), maar dit concept wordt hier niet verder uitgewerkt. Met betrekking tot het gebruik van de negatieve bepalingen geldt bovendien een belangrijk voorbehoud: de aanname bij het beschreven gebruik is dat er voldoende tijd is verstreken sinds de aanvang van de (doorlopende) besmetting om de betreffende sensoren te bereiken.

a)



b)



Figuur 9: a) Herkomstkaart voor tijdstip  $t=72$  uur voor een enkele sensorlocatie; b) Herkomstkanskaart voor hetzelfde tijdstip en dezelfde locatie. Het gebruikte netwerkmodel is het eerste model uit de Battle of the Water Sensor Networks (Ostfeld et al., 2008).

Let op dat hierbij de factor tijd volledig buiten beschouwing wordt gelaten. In feite gaan we ervanuit dat de verontreinigingsbron reeds lange tijd (voor ieder punt in het netwerk langer dan het oudste water dat langs een van de beschouwde sensoren stroomt) in het netwerk aanwezig is.



Verfijning van de benadering door het combineren van gegevens over de tijdspanne sinds het passeren van ieder punt voor iedere beschouwde sensor leidt in principe tot een verdere inperking. We bepalen voor iedere sensorlocatie op welk tijdstip het passerende water langs alle knooppunten van het netwerk is gestroomd. Vervolgens eisen we dat voor de sensoren die een veronderstelde verontreiniging detecteren de passagetijden voor de potentiële brongebieden binnen een bepaalde tijdspanne van elkaar moeten liggen. Deze eis komt voort uit de veronderstelling dat de verschillende sensoren dezelfde verontreiniging constateren, die dus in dezelfde periode in het netwerk is gekomen. De gekozen tijdspanne is bijvoorbeeld een half uur voor een veronderstelde kortdurende injectie, of langer voor een verondersteld langer lopend probleem. Bovendien eisen we dat de passagetijden voor niet-detecterende sensoren buiten deze tijdspanne liggen. Het nu gevonden potentiële brongebied zal kleiner van omvang zijn dan alleen de overlap van herkomstgebieden. Deze verfijning is in het volgende nog niet toegepast, maar wordt aanbevolen voor verdere ontwikkeling.

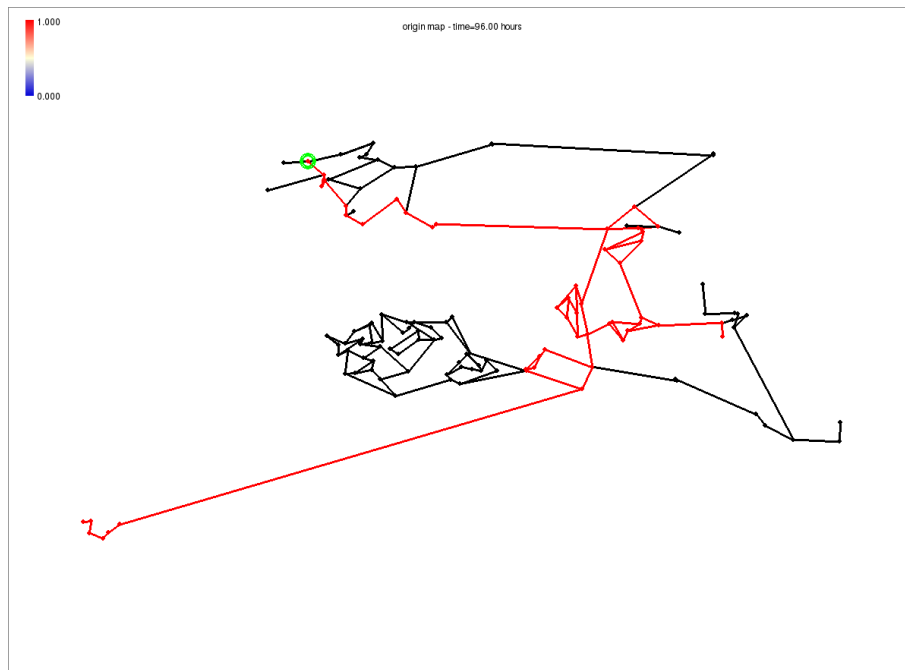
### **3.3.3 Dekkingskaarten**

Een herkomstkaart voor een sensor voor een bepaald tijdstip geeft weer welk deel van het netwerk wordt gezien (gedekt) op het bijbehorende observatietijdstip. Let op dat hierbij besmettingstijdstippen horen die in het verleden liggen en verschillen per deel van het netwerk. Herkomstkaarten kunnen worden gebruikt om dekkingskaarten voor een of meerdere sensoren te maken. Bijlage IV geeft een volledig overzicht van de methodiek, die in hoofdstuk 0 zal worden toegepast. Hieronder worden de belangrijkste toepassingen met enkele voorbeelden geïllustreerd.

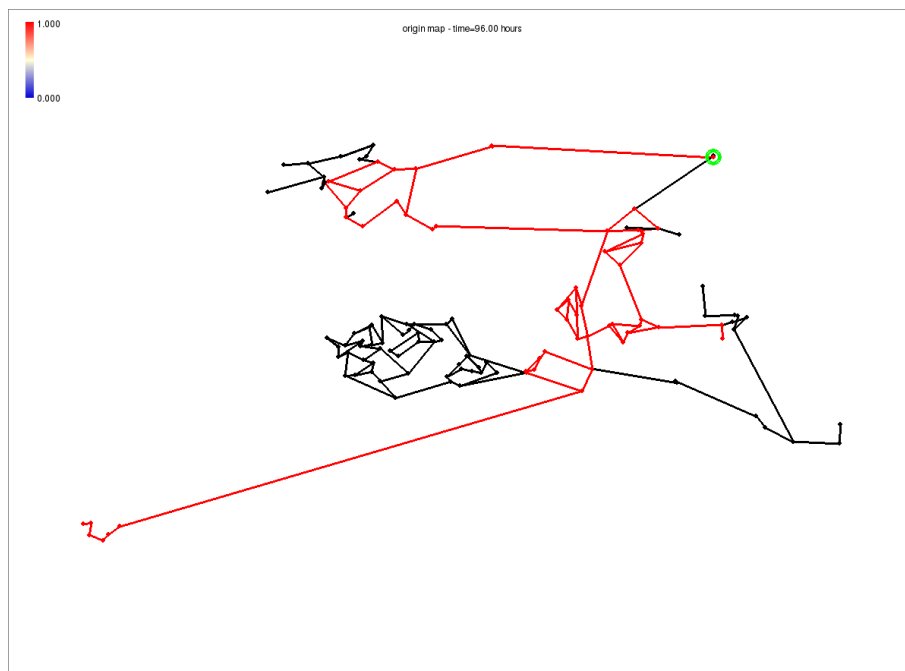
Doordat het snelheidsveld in het netwerk door de dag heen varieert, kunnen herkomstkaarten voor een punt (sensorlocatie) er op verschillende momenten van de dag anders uitzien (Figuur 13a,b). Door de doorsnede (overlap) van alle herkomstkaarten in een aaneengesloten periode van 24 uur te berekenen, wordt het effectieve dekkingsgebied van een sensor bepaald (Figuur 13c). Hiernaast zorgen stochastische variaties in de watervraag voor variaties in het snelheidsveld, die op dezelfde manier kunnen worden behandeld.

Ook kan worden weergegeven gedurende welke fractie van een etmaal er in een bepaald gebied dekking is (Figuur 14). Dit kan slechts een deel van de dag zijn ( $<1$ ), de gehele dag (1), of zelfs meervoudige dekking (door meer dan een sensor). Let op dat een waarde van 1 niet wil zeggen dat de locatie de gehele dag in het zichtveld van dezelfde sensor ligt.

a)

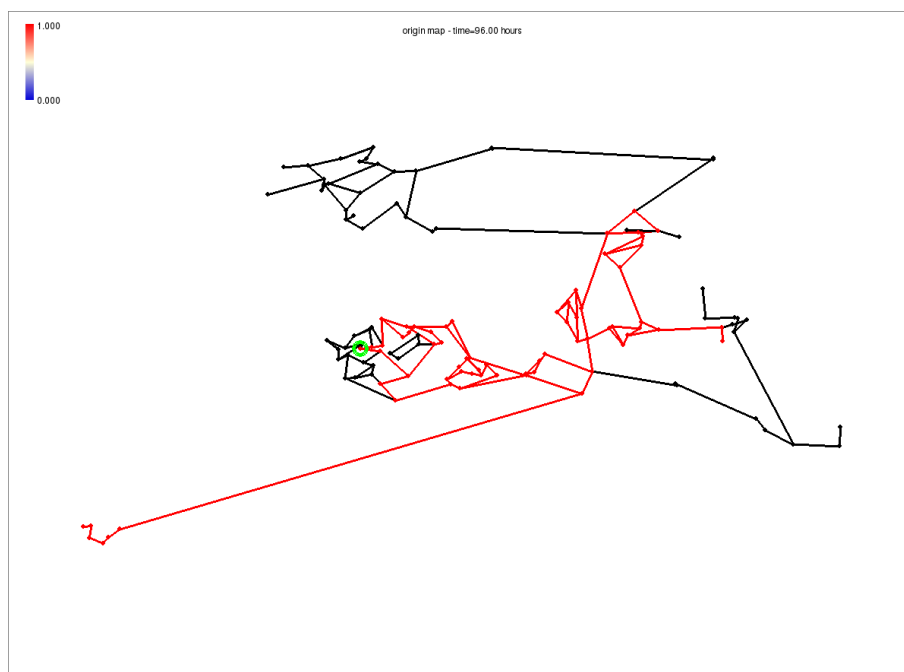


b)

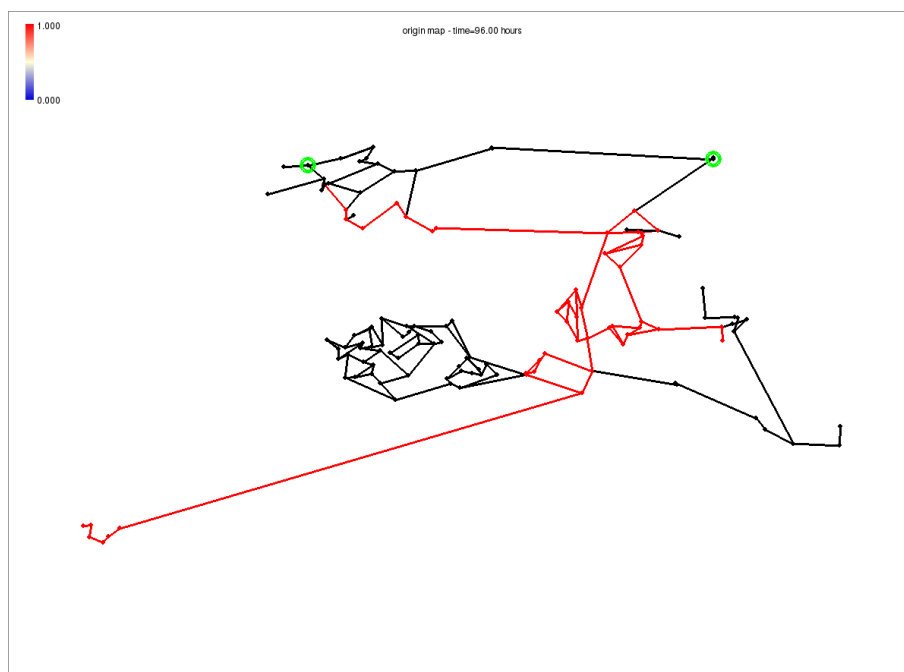


*Figuur 10: Herkomstkaarten voor individuele sensoren.*

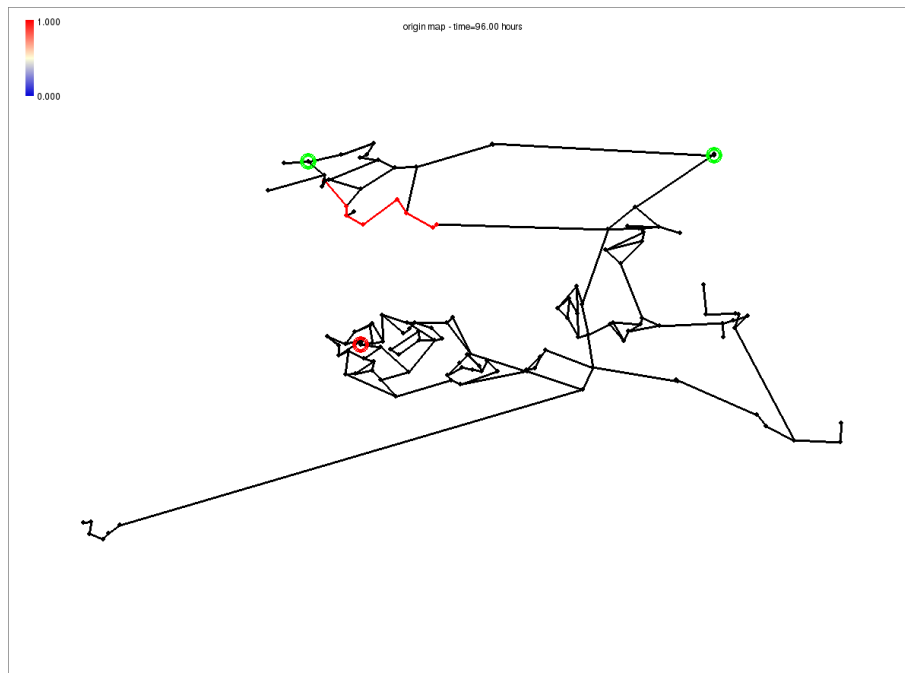
a)



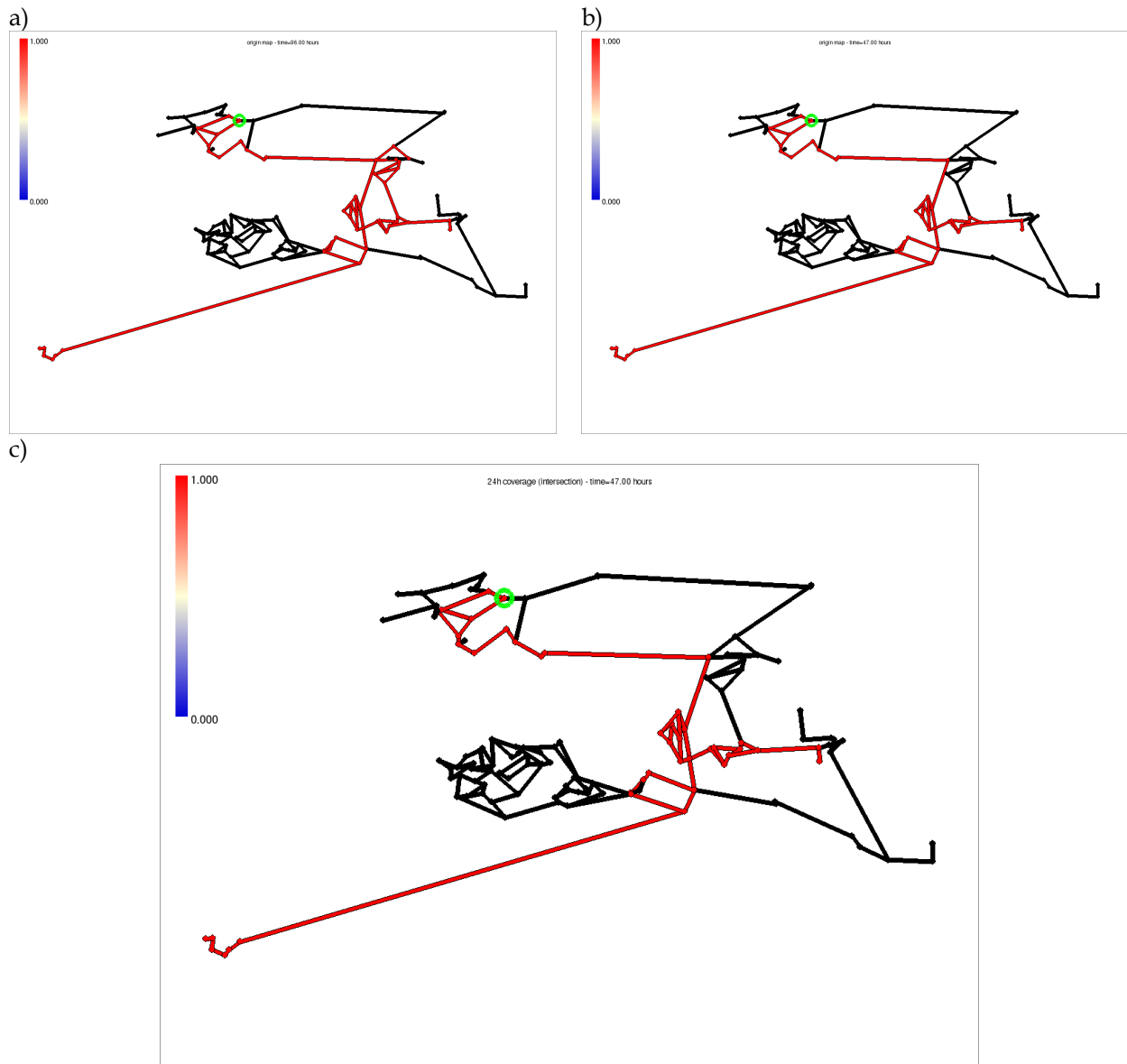
b)



Figuur 11: a) Herkomstkaart voor individuele sensor. b) Gecombineerde herkomstkaart voor de twee sensoren van Figuur 10.

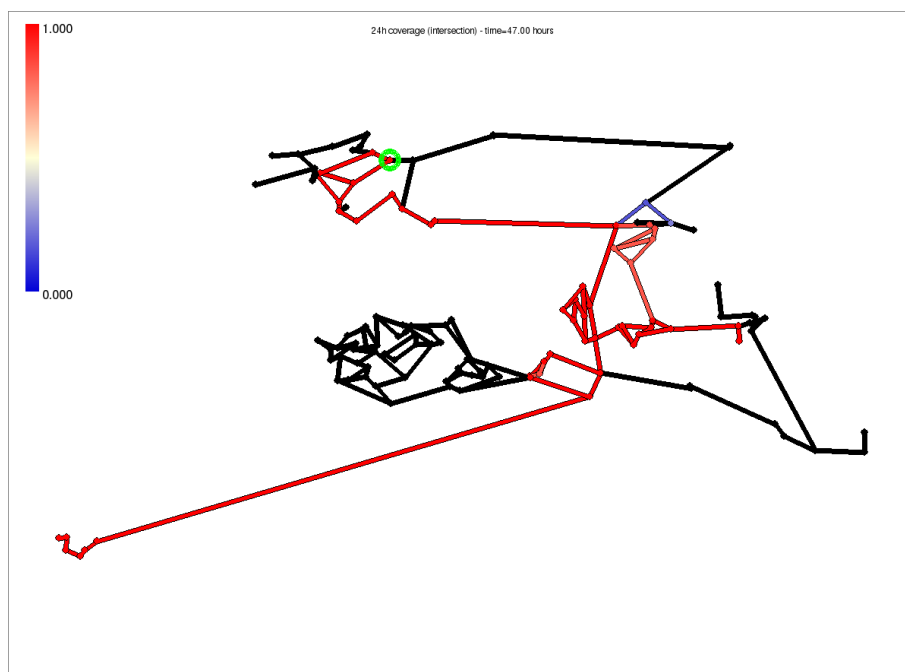


*Figuur 12 Combinatie van herkomstkaarten van individuele sensoren van Figuur 10. Toevoeging van een derde, negatieve sensormeting (Figuur 11a), perkt het potentiële brongebied verder in. Let op dat beschouwing van de potentiële bronlocatie in dit geval uitsluitend op de knooppunten plaatsvindt.*

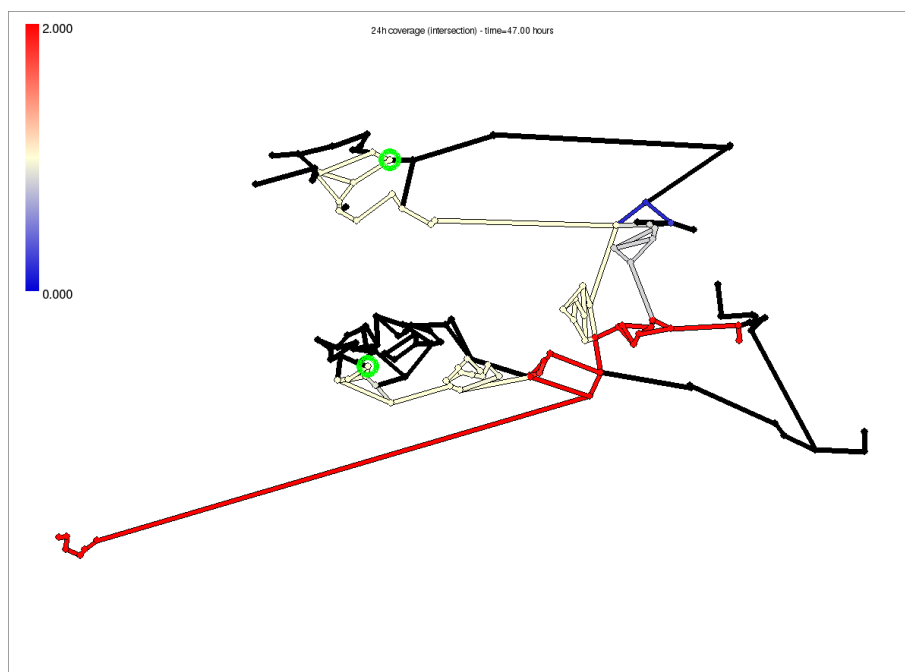


Figuur 13: a) Herkomstkaart voor tijdstip 1, b) herkomstkaart voor tijdstip 2, c) effectieve dekkingsgebied op basis van herkomstkaarten voor een periode van 24 uur.

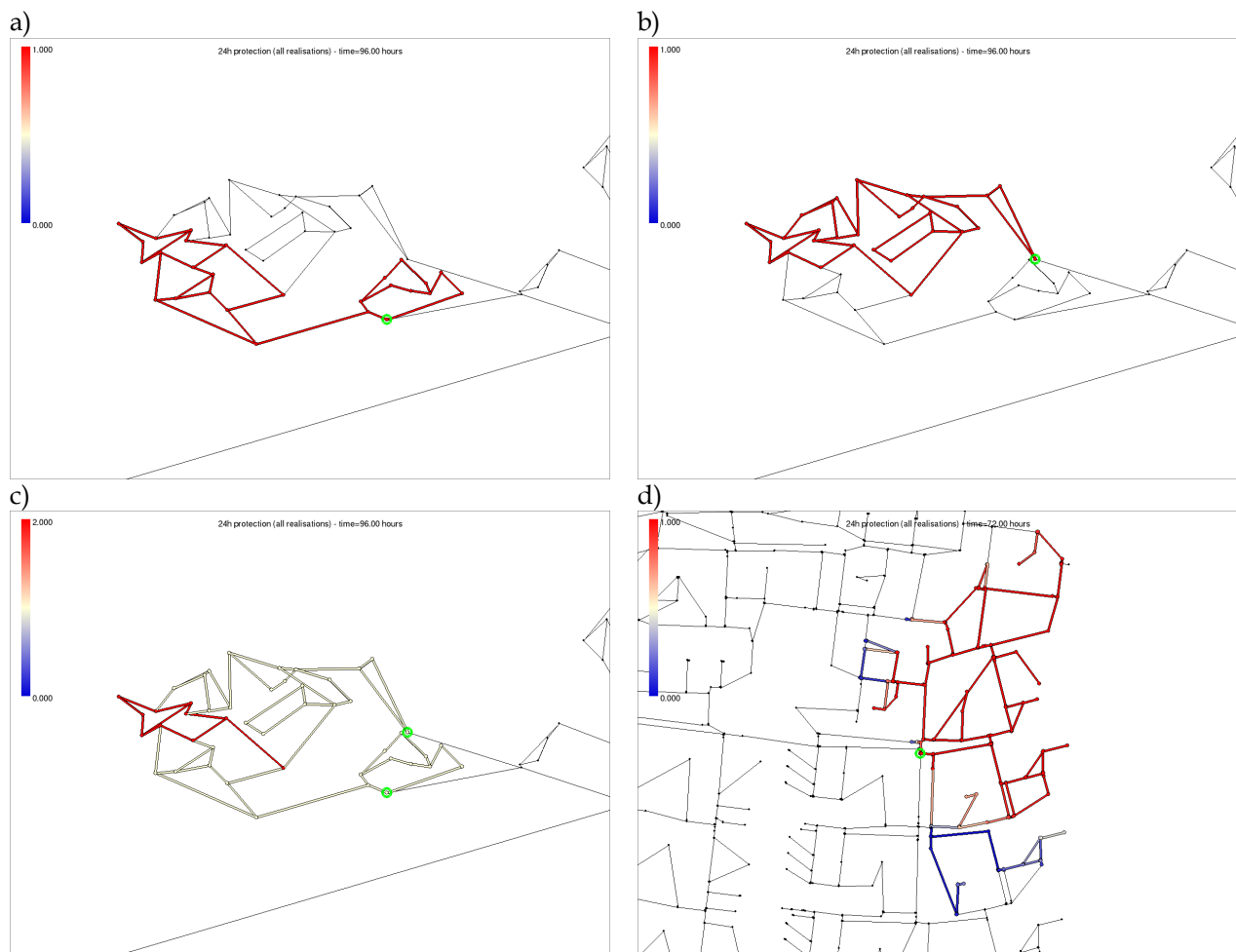
a)



b)



Figuur 14: Fractie van de tijd dat delen van het leidingnet in het zichtveld van de gemarkeerde sensor(en) ligt. a) enkele sensor, b) meerdere sensoren. Waarden groter dan 1 geven aan dat een locatie in het gezichtsveld van meer dan 1 sensor ligt.



Figuur 15: 24-Uurs beschermingsgebieden voor sensorlocaties. a,b) beschermingsgebieden voor individuele sensorlocaties; c) gecombineerd beschermingsgebied; d) 24-uurs beschermingsgebied bij grote variaties in het snelheidsveld door de dag heen en van dag tot dag. De fractie van de dag dat een punt binnen het beschermingsgebied ligt is aangegeven met de kleurschaal.

### 3.3.4 Beschermingsgebieden

Door te kijken waar tracerstoffen vanaf een sensorlocatie naar toe worden getransporteerd, kan worden vastgesteld welk gebied in feite door een sensor wordt beschermd tegen een besmetting stroomopwaarts van de sensor (Figuur 15a,b). Voor combinaties van sensoren kan eveneens de redundantie worden vastgesteld (Figuur 15c). Ook hiervoor geldt dat door variaties in de watervraag en het snelheidsveld door de dag heen en van dag tot dag het dekkingsgebied enigszins kan variëren (Figuur 15d).

### 3.4 Gebruikte en ontwikkelde software

Om de hierboven beschreven methoden toe te passen, is gebruik gemaakt van bestaande software (EPANET-MSX) en is bovendien enige software geschreven. Deze is geschreven in Python (Python, 2013), hetgeen een vrij beschikbare programmeertaal is die op tal van platformen kan worden geïnstalleerd. Een overzicht hiervan is weergegeven in Tabel 2 en Bijlage V. Tests hiervan zijn weergegeven in Bijlage VI.

Tabel 2: Overzicht van software waarin de beschreven benadering zijn geïmplementeerd.

<i>naam</i>	<i>doel</i>	<i>type</i>	<i>status</i>
EPANETinp2msx	Maakt op basis van een EPANET-invoerbestand een reeks invoerbestanden voor EPANET-MSX t.b.v. <i>backtrace</i> berekeningen	command line	klaar en werkend
joinmsxresults	Voegt de resultaten van een reeks berekeningen met EPANET-MSX samen tot een enkel bestand met interne compressie. Hiermee wordt de omvang van de data met een factor 25 teruggebracht en het verwerken van de resultaten aanzienlijk versneld.	command line	klaar en werkend
wxreadepanet	Leest de resultaten van EPANET/EPANET-MSX/joinmsxresults in en combineert deze tot herkomstkaarten, dekkingskaarten, beschermingskaarten, etc. (Lagrangiaanse benadering met probabilistische interpretatie).	grafische gebruikersschil	klaar en werkend
stochasticbacktracer	Voert stoftransport terug in de tijd uit op synthetische of ingelezen netwerken op basis van de hierboven beschreven probabilistische inversie benadering.	grafische gebruikersschil	experimenteel





# 4 Toepassing van de ontwikkelde methodieken

## 4.1 Inleiding

In het voorgaande zijn twee methodieken voor *backtracing* van verontreiniging besproken. Vanwege het bereikte ontwikkelstadium van de in hoofdstuk 3.3 beschreven inversiemethodiek, is er echter voor gekozen om uitsluitend de ontwikkelde Lagrangiaanse methode van hoofdstuk 3.2 toe te passen op de voor dit project geselecteerde testgebieden. De resultaten hiervan worden in de volgende paragrafen beschreven.

## 4.2 Uitgangspunten en onderzoeksvragen

De volgende uitgangspunten zijn gekozen:

- Binaire beschouwing van detectie van afwijkingen/besmettingen: òf wel, òf niet (aan/uit).
- Er worden herkomstkaarten gemaakt voor een selectie van reële sensorlocaties op basis van bestaande monsterlocaties.
- Voor de verschillende testgebieden (Almere, Eindhoven) wordt uitgegaan van gelijke sensordichtheden.
- Het model moet minimaal drie dagen in de tijd terugkijken.
- Historische monsterlocaties worden beschouwd als sensorlocaties.

Een antwoord wordt gezocht op de volgende onderzoeksvragen:

- Hoe goed is de dekking van de voorzieningsgebieden voor de gekozen sets van locaties.
- Hoe goed is de locatie van de bron van een besmetting in te perken?
- Hoe verhouden de resultaten voor de verschillende testgebieden zich tot elkaar?

Bovendien wordt de volgende vraag verkend:

- Wat is de gevoeligheid van de berekende herkomstkaarten voor het tijdstip van bemonstering/meting?

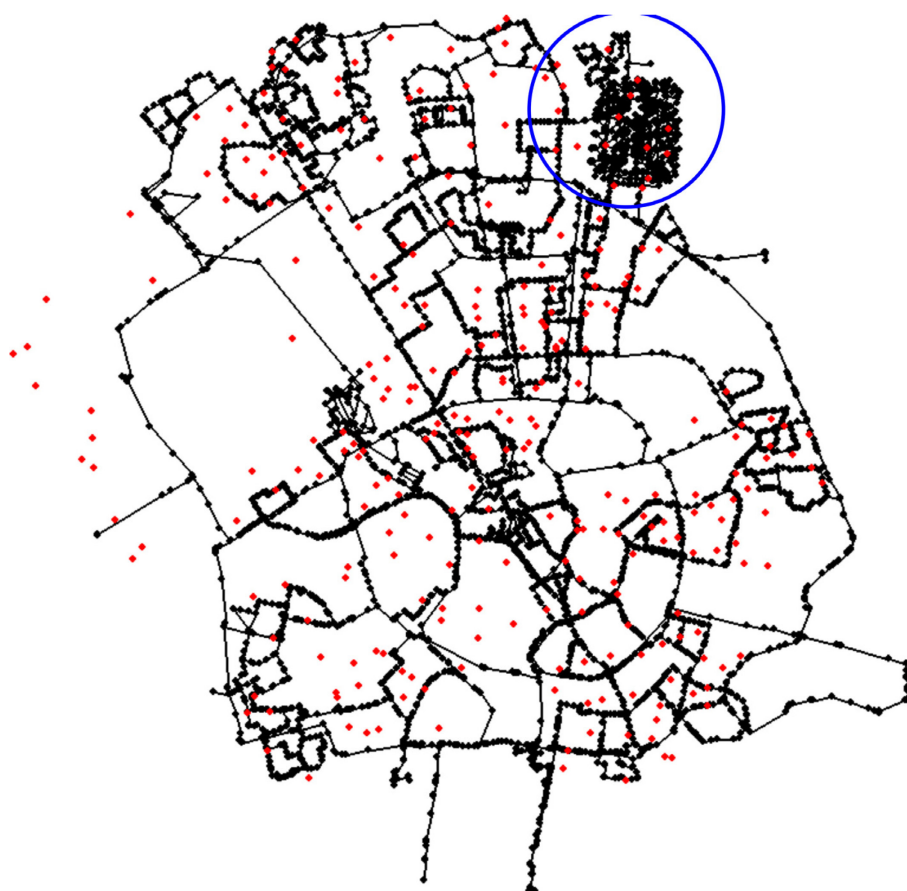
## 4.3 Testgebieden

Voor toepassing van de uitgewerkte methoden is gekozen voor twee steden waarvan reeds netwerkmodellen met bijbehorende SIMDEUM-patronen beschikbaar waren vanuit andere projecten, namelijk Eindhoven en Almere (Figuur 16 en Figuur 17). Hiervoor zijn potentiële sensorlocaties geselecteerd op basis van reguliere monsterlocaties. Voor Eindhoven is gekozen voor alle 286 locaties die in 2011 zijn bezocht. Gekozen is om in Almere dezelfde sensorlocatiedichtheid toe te passen. De maat voor de sensorlocatiedichtheid die gebruikt is, is het aantal sensoren per lus in een onverdund netwerkmodel. Aangezien het onverdunde model van Eindhoven 2485 lussen telt en dat van Almere 1473, zijn voor Almere de 170 meest bezochte (minimaal 13 keer) monsterlocaties in de jaren 2009-2012 gekozen. De locaties zijn in Figuur 18 weergegeven met rode stippen.

Om de benodigde rekentijd en opslagruimte te beperken, is ervoor gekozen om de modellen deels te verdunnen. Om ook effecten in de tertiaire structuren te kunnen onderzoeken, is in beide modellen een wijk onverdund gelaten (Figuur 18c,d). Voor Eindhoven betreft het de buurten Vaartbroek en Heesterakker; bij Almere gaat het om Almere-Haven. Beide onverdunde wijken worden vanaf verschillende punten gevoed. Details van het voorbereidingsproces van de modellen zijn weergegeven in Bijlage VII. Voor beide modellen zijn de resulterende aantallen knooppunten en leidingen weergegeven in Tabel 3.

Tabel 3: Aantallen knooppunten en leidingen in de toegepaste hydraulische modellen van Eindhoven en Almere, inclusief onverdunde wijken. De gerapporteerde rekentijd geldt per realisatie van het snelheidsveld op een Intel Core2 Q8200-processor.

model	volledig model			onverdunde wijk	
	aantal knopen	aantal leidingen	rekentijd (cpu-uren)	aantal knopen	aantal leidingen
Eindhoven	6673	6694	ca. 72	1330	1435
Almere	6856	7884	ca. 92	3962	4282



Figuur 16.: Netwerkmodel van proefgebieden Eindhoven. Rode stippen geven monsterlocaties van 2011 aan, de blauwe cirkel markeert de detailwijk.

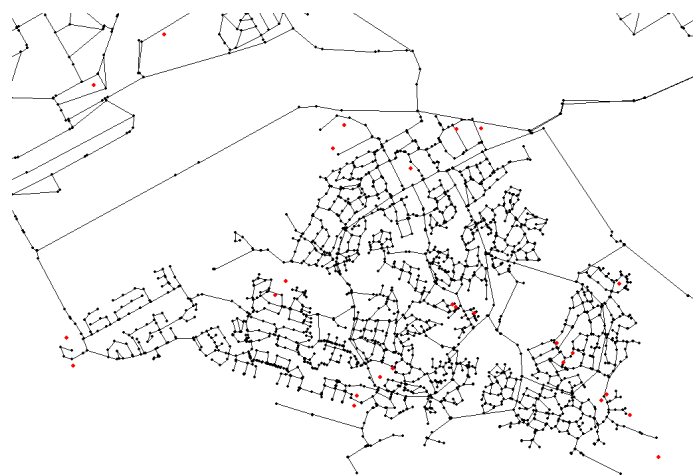


Figuur 17: Netwerkmodel van proefgebieden Eindhoven. Rode stippen geven geselecteerde monsterlocaties aan, het blauwe ovaal markeert de detailwijk.

a)



b)



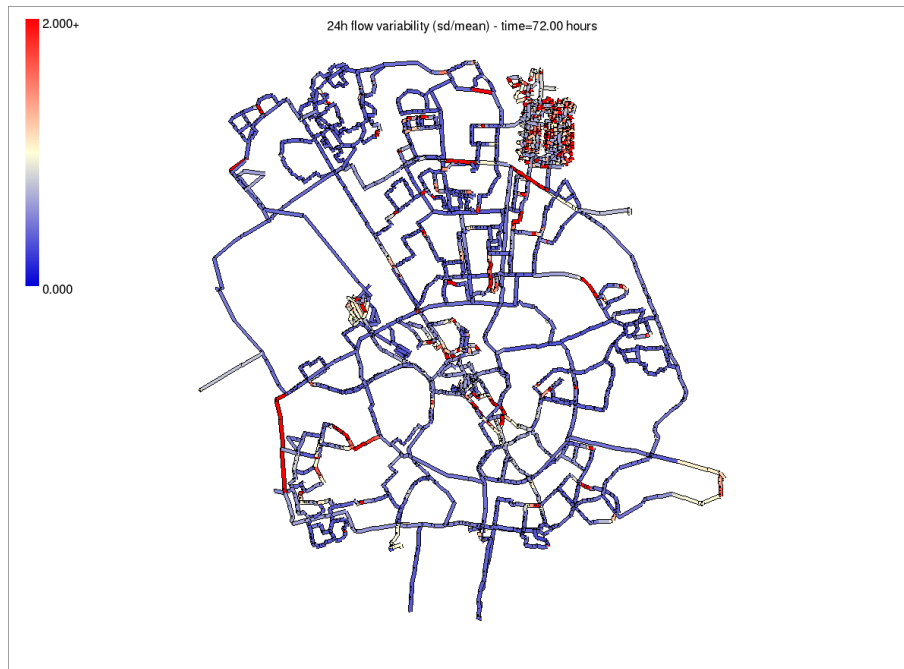
Figuur 18: Uitsnedes van wijken die in volledig detail zijn beschouwd voor Eindhoven (a) en Almere (b).

## 4.4 Resultaten Eindhoven

### 4.4.1 Variabiliteit van het stromingsveld

Zoals besproken in §2.1 is de belangrijkste moeilijkheid bij het *backtracen* van verontreinigingen richting de bron de variabiliteit van het stromingsveld in het netwerk. Voor het netwerk van Eindhoven wordt deze variabiliteit geïllustreerd in Figuur 19. De variatie in de volumestroom (Figuur 19a) is duidelijk in zowel enkele segmenten van het verdunde deel van het netwerk, dat we bij benadering het secundaire

a)



b)



Figuur 19: Illustratie van de variabiliteit van het stromingsveld in het netwerk van Eindhoven gedurende een periode van 24 uur: a) Standaarddeviatie van de volumestroom gedeeld door de gemiddelde volumestroom; b) fractie van de tijd dat de stroomrichting omgekeerd is ten opzicht van de meestvoorkomende stroomrichting. Beide figuren zijn gebaseerd op 12 tijdstippen (interval 2 uur) in 5 realisaties van het stromingsveld en daarom slechts indicatief.

netwerk zouden kunnen noemen, als in een groot deel van het tertiaire netwerk in de detailwijk. Ook vinden we regelmatige omkeringen van de stroomrichting (Figuur 19b) zowel op secundair als op tertiair niveau terug.

#### **4.4.2 Herkomst en dekking**

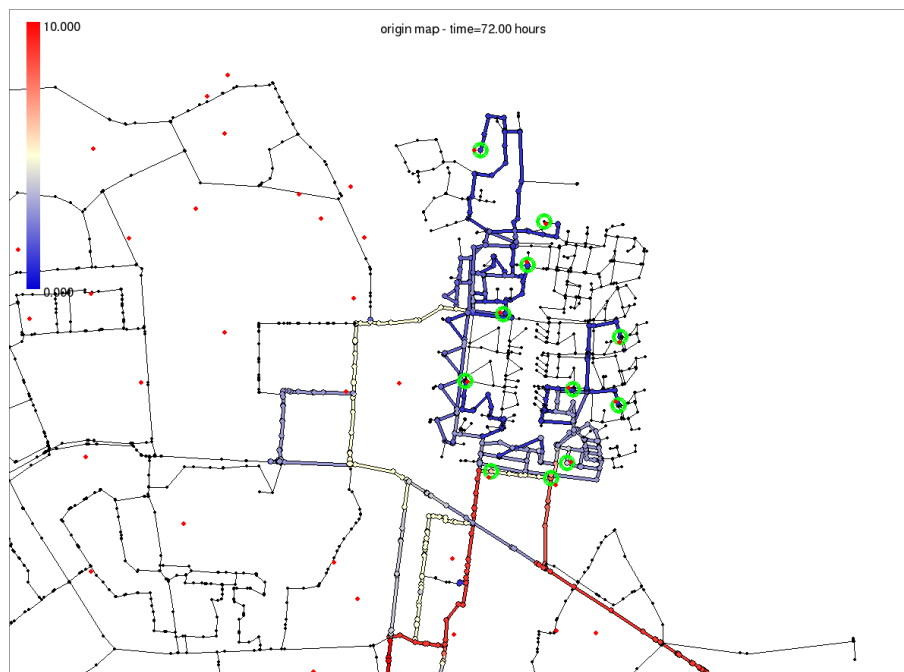
Uitgaande van de monsterlocaties als sensorlocaties kan een grote hoeveelheid herkomstkaarten voor sensorlocaties worden gegenereerd voor ieder (gesimuleerd) tijdstip van de dag. Samenvoeging van de kaarten voor een periode van 24 uur geeft een dekkingskaart voor ieder van deze locaties. Wanneer ook nog combinaties van sensoren worden beschouwd, wordt het aantal mogelijkheden bijna oneindig. Om die reden beperken we ons hier tot een selectie.

Figuur 20 en Figuur 21 tonen herkomstkaarten voor een selectie van 11 sensorlocaties overeenkomend met 11 monsterlocaties in de onverdunde wijk. Uitgangstijdstip is middernacht na 72 uur simulatietijd. Weergegeven zijn herkomstkaarten in een detailweergave van de wijk en zijn directe omgeving (Figuur 20a,b) en een weergave van het relevante deel van het volledige netwerk (Figuur 21a,b). Kaarten zijn gemaakt voor één realisatie van het snelheidsveld (Figuur 21a,c) en voor een combinatie van 5 realisaties (Figuur 20b en Figuur 21b). Elk van de kaarten geeft voor iedere locatie in het netwerk met een kleur aan door hoeveel sensoren de betreffende locatie voor dit tijdstip wordt gedekt. Figuur 22 en Figuur 23 tonen 24-uurs dekkingskaarten voor deze zelfde combinaties.

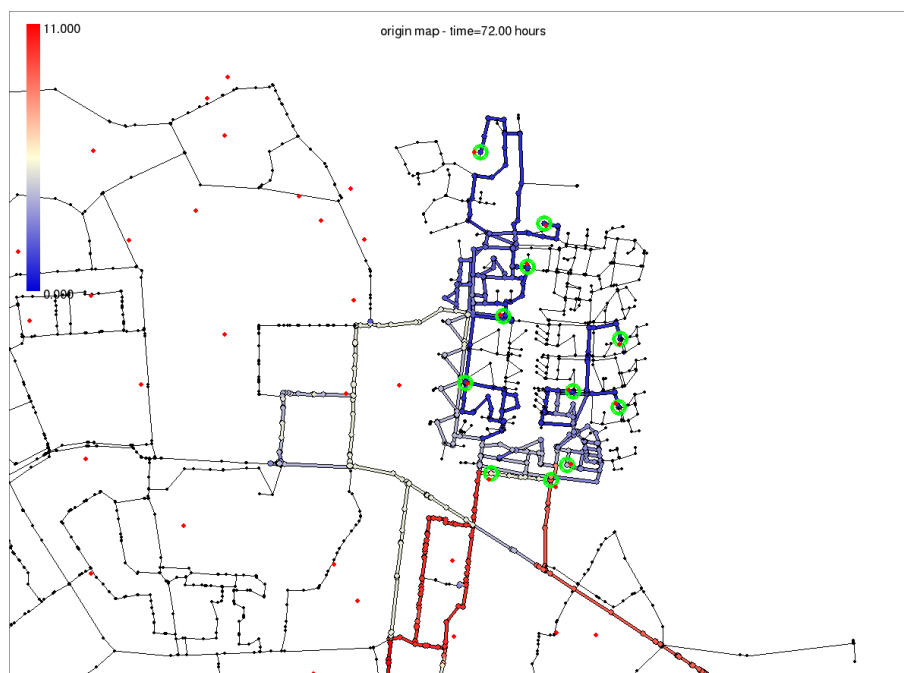
Voor deze specifieke wijk vallen drie zaken op. Ten eerste is het aantal monsterpunten onvoldoende om besmettingen die binnen een aanzienlijk deel van de wijk optreden op te pikken. Vermoedelijk geldt dit punt eveneens voor andere wijken waar door verdunning de fijne structuur van het netwerk is verdwenen in het model. Ten tweede zijn de verschillen tussen de niet-stochastische benadering (Figuur 20a, Figuur 21a, Figuur 22a en Figuur 23a, een enkel snelheidsveld) en de stochastische benadering (Figuur 20b, Figuur 21b, Figuur 22b en Figuur 23b, combinatie van 5 realistische snelheidsvelden) gering. Slechts op een enkele locatie verschilt het stromingsveld zodanig van realisatie tot realisatie dat er andere paden worden gevolgd. Ten derde valt op dat iedere dekkingskaart een deelverzameling representeert van de punten die in de herkomstkaart zijn weergegeven. Dit ligt voor de hand, omdat water vanaf sommige punten slechts gedurende een deel van de dag de sensorlocatie bereikt.

Wanneer we uitzoomen naar het niveau van het volledige netwerkmodel (Figuur 24a), en alle 236 virtuele sensorlocaties beschouwen in een gecombineerde herkomstkaart, is opnieuw te zien dat er delen in het netwerk zijn die niet door sensoren worden gezien, zelf in de secundaire structuur (Figuur 24b,c).

a)

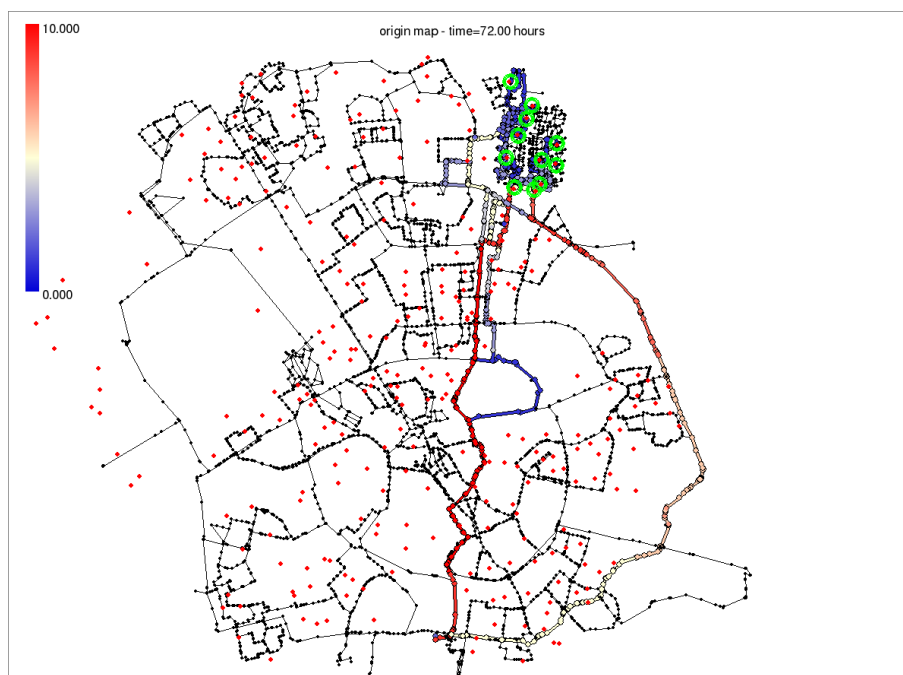


b)

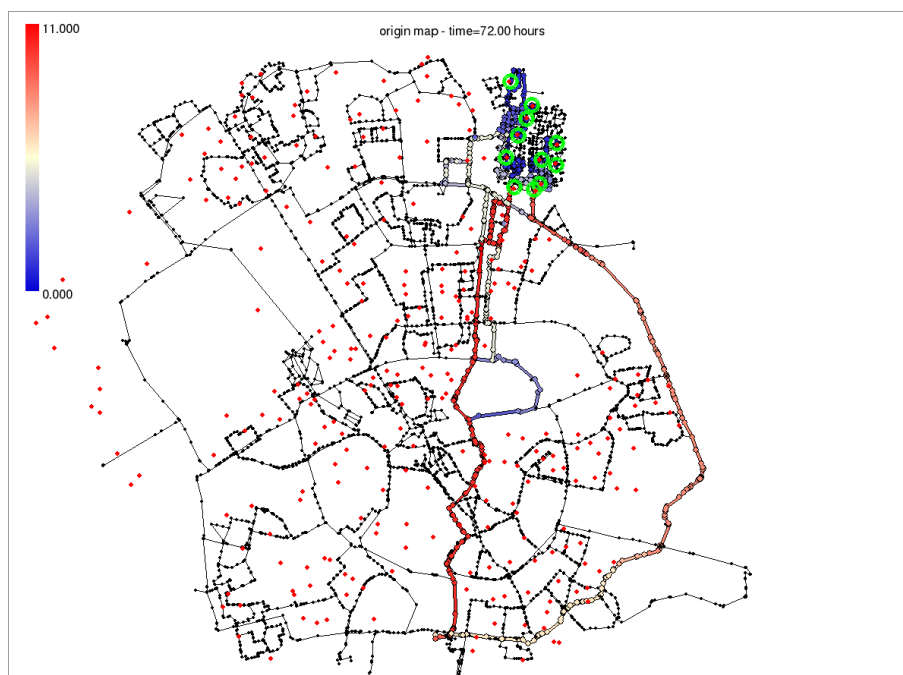


Figuur 20: Herkomstkaarten (referentietijdstip 00:00) voor Eindhoven voor een combinatie van 11 sensoren gekoppeld aan 11 monsterlocaties in de onverdunde wijk: detailweergaves van de wijk. Kaarten zijn gebaseerd op 1 realisatie van het snelheidsveld (a) en een combinatie van 5 realisaties (b). De kleurschaal geeft aan hoeveel sensorherkomstgebieden elkaar op ieder punt overlappen. Monsterpunten zijn in rood weergegeven, sensorlocaties zijn groen omcirkeld.

a)



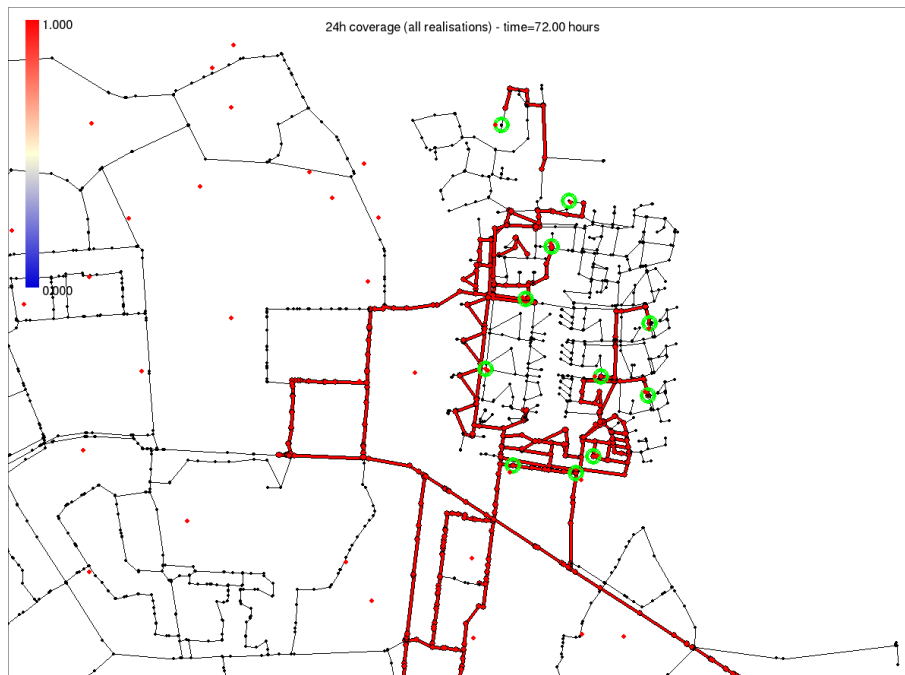
b)



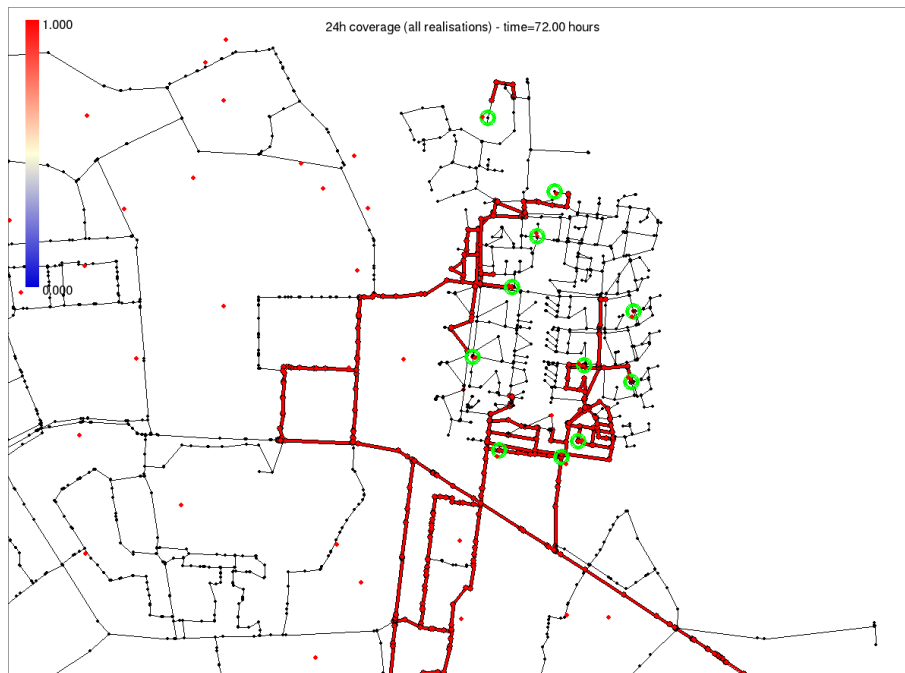
*Figuur 21: Herkomstkaarten (referentietijdstip 00:00) voor Eindhoven voor een combinatie van 11 sensoren gekoppeld aan 11 monsterlocaties in de onverdunde wijk: volledige weergave van het relevante deel van het netwerk. Kaarten zijn gebaseerd op 1 realisatie van het snelheidsveld (a) en een combinatie van 5 realisaties (b). De kleurschaal geeft aan hoeveel sensorherkomstgebieden elkaar op ieder punt overlappen. Monsterpunten zijn in rood weergegeven, sensorlocaties zijn groen omcirkeld.*



a)

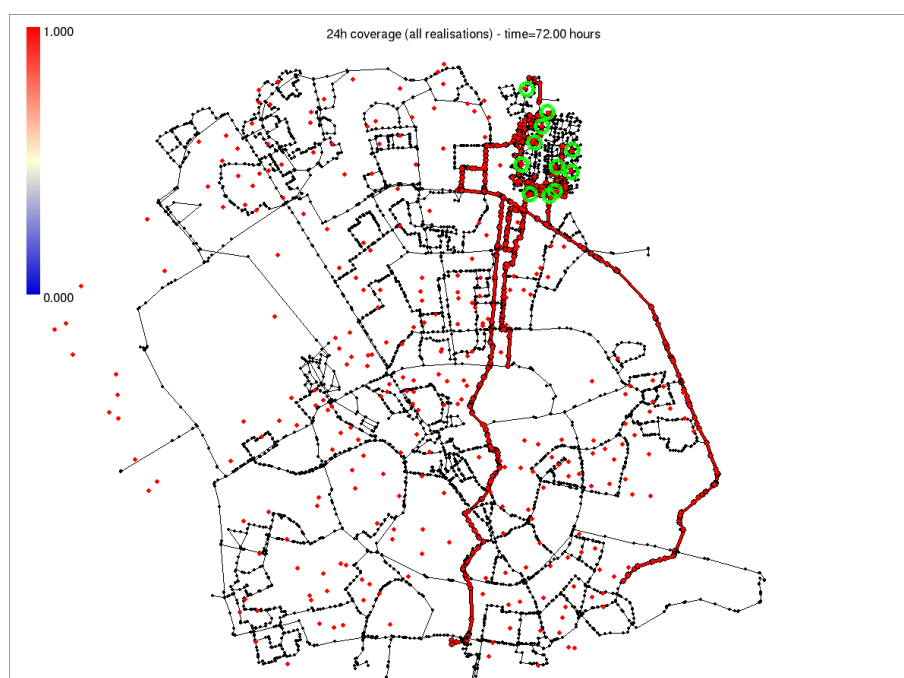


b)

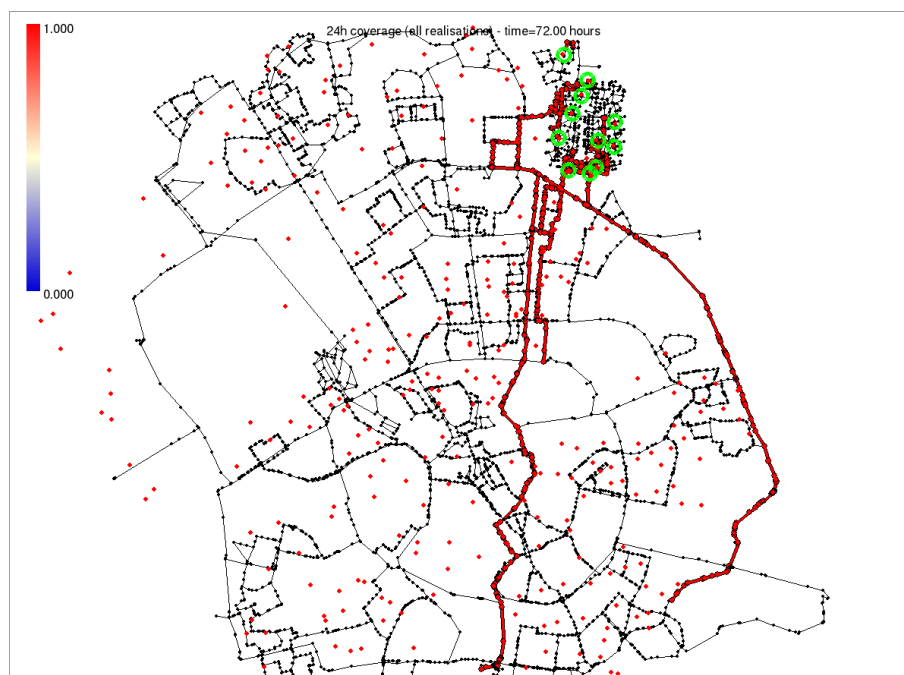


Figuur 22: 24-Uurs dekkingskaarten (referentietijdstip 00:00) voor Eindhoven voor een combinatie van 11 sensoren gekoppeld aan 11 monsterlocaties in de onverdunde wijk: detailweergave van de wijk. Kaarten zijn gebaseerd op 1 realisatie van het snelheidsveld (a) en een combinatie van 5 realisaties (b). Gebieden met dekking gedurende 24 uur (niet noodzakelijkerwijs continu door dezelfde sensor) zijn rood gemarkeerd.

a)

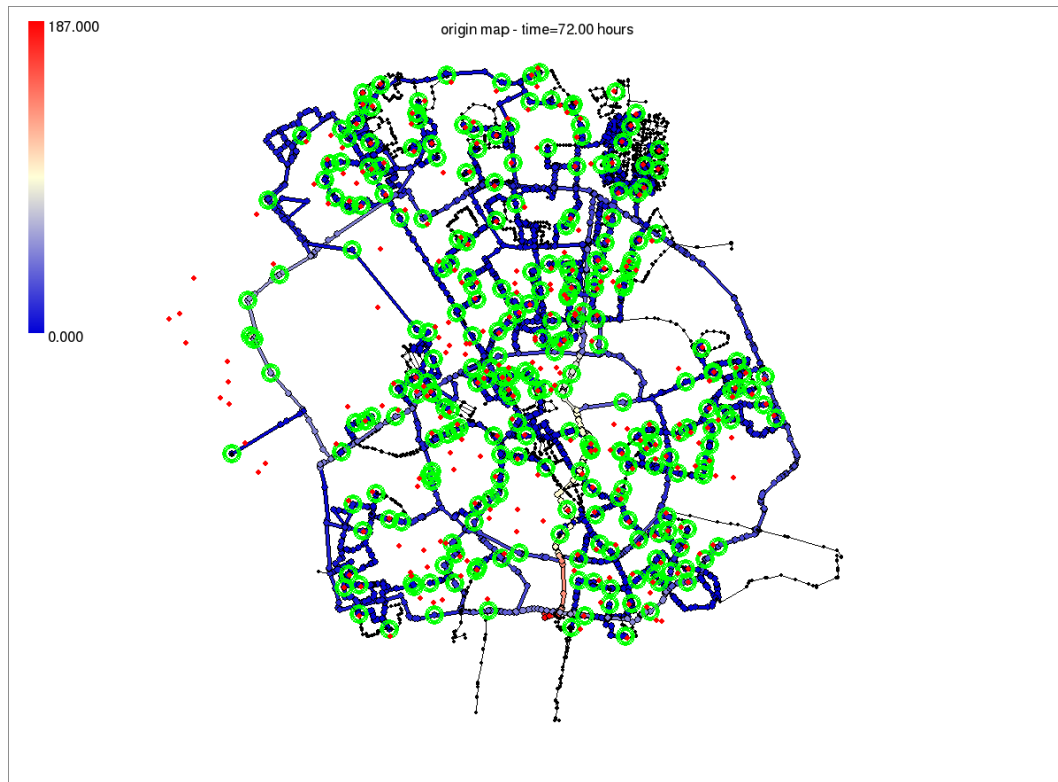


b)

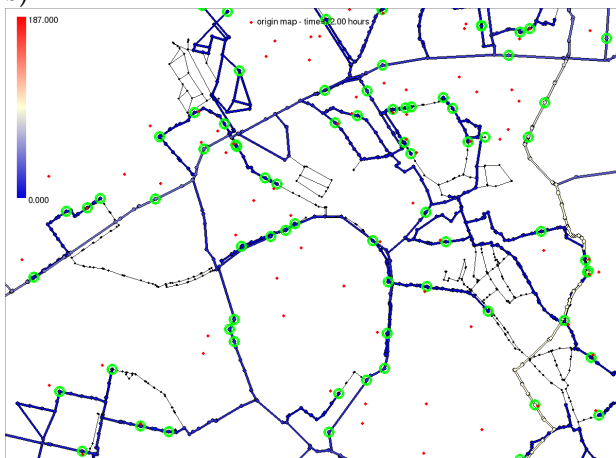


Figuur 23: 24-Uurs dekkingskaarten (referentietijdstip 00:00) voor Eindhoven voor een combinatie van 11 sensoren gekoppeld aan 11 monsterlocaties in de onverdunde wijk: een volledige weergave van het relevante deel van het netwerk. Kaarten zijn gebaseerd op 1 realisatie van het snelheidsveld (a) en een combinatie van 5 realisaties (b). Gebieden met dekking gedurende 24 uur (niet noodzakelijkerwijs continu door dezelfde sensor) zijn rood gemarkeerd.

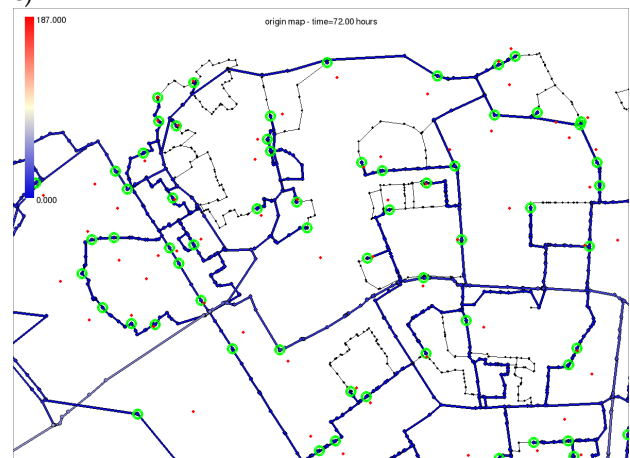
a)



b)



c)



Figuur 24: Herkomstkaarten (referentietijdstip 00:00) voor een combinatie van 236 sensoren corresponderend met alle monsterlocaties in het model voor 1 realisatie van het snelheidsveld. a) volledig model, met monsterlocaties in rood en bijbehorende sensorlocaties in groen. b,c) deelgebieden in detail. De kleurschaal geeft aan hoeveel sensorherkomstgebieden elkaar op ieder punt overlappen. Bij afwezigheid van enige dekking zijn leidingen en punten dun en zwart weergegeven.

#### 4.4.3 Nadere beschouwing van stochastiek

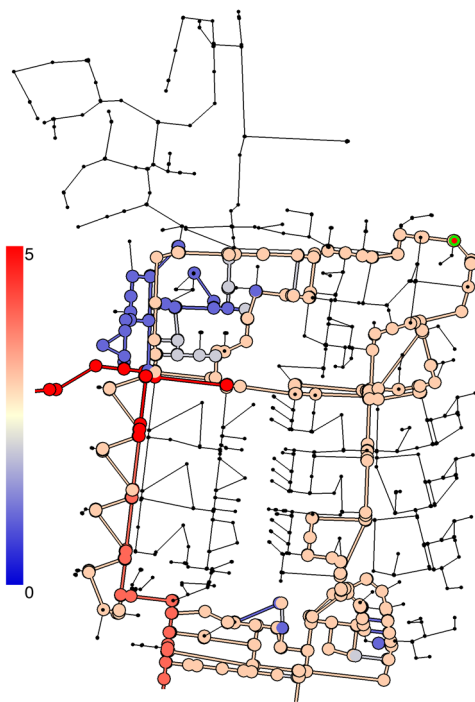
Een iets andere benadering van de evaluatie van effect van stochastische vraag op *backtracing* is door te kijken naar de hoeveelheid overlap tussen *traces* voor individuele realisaties van het snelheidsveld. Figuur 25b toont een selectie van 12 knopen in het netwerkmodel, waarbij *backtraces* werden berekend voor alle 5 realisaties van het snelheidsveld. De hoeveelheid overlap tussen de sporen van deze vijf realisaties wordt gevisualiseerd in Figuur 25c met behulp van cirkeldiagrammen. Deze taartdiagrammen

tonen welke fractie (in termen van aantallen netwerkknopen, zie Figuur 25a) van de gecombineerde *backtraces* wordt gedekt door een, twee, drie, vier of vijf realisaties. We beschouwen alleen die delen van de *traces* die binnen het gedetailleerde deel van de netwerkmodel vallen (zie Figuur 25a). Merk op dat deze benadering de *backtraces* beschouwt op een binaire manier: een knooppunt is onderdeel van de *backtrace* of niet. Grotendeels of geheel blauwe ( $n = 5$ ) taartdiagrammen geven aan dat voor deze knooppunten de *backtrace* hetzelfde is voor alle realisaties van het snelheidsveld. Knooppunten waarbij de blauwe ( $n = 5$ ) fractie kleiner is tonen echter een aanzienlijke variabiliteit in de *backtraces* voor dit knooppunt voor andere realisaties van het stromingsveld. *Backtracing* vanaf deze knooppunten is daarom zeer gevoelig voor de stochastische aard van watervraag. Zoals te zien in Figuur 25b, zijn deze gevoelige knooppunten meestal te vinden in de periferie van het tertiaire netwerk.

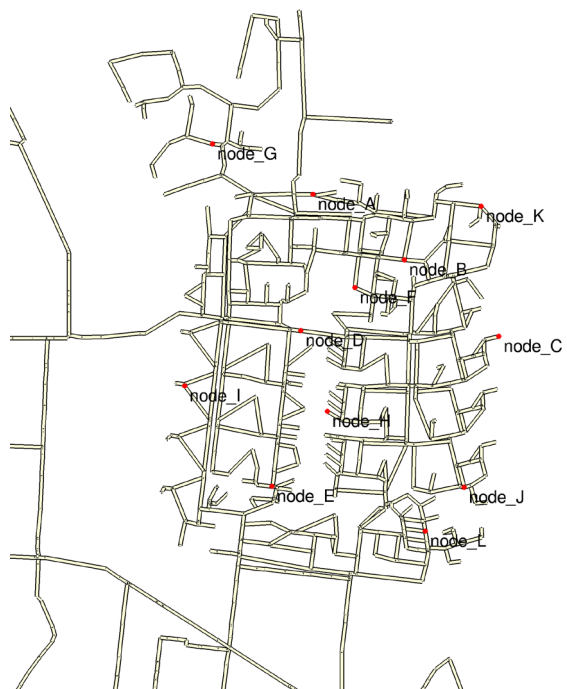
#### **4.4.4 Bronbepaling**

Door combinatie van positieve sensormetingen (d.w.z. sensormetingen waarin een verontreiniging wordt gedetecteerd) en negatieve sensormetingen (waarin deze niet te zien is) kan het potentiële herkomstgebied van een verontreiniging sterk worden ingeperkt. Ook hier zijn teveel combinaties mogelijk om deze allemaal te tonen – opnieuw wordt een selectie gepresenteerd. Figuur 26 en Figuur 27 tonen twee voorbeelden van inperking van een potentieel brongebied van een bij een sensor geconstateerde verontreiniging door toevoeging van informatie van meerdere sensoren, zowel positief (ook gemeten) als negatief (niet gemeten). De omvang van het potentiële brongebied krimpt hiermee van de doorsnede van de stad tot slechts enkele leidingsegmenten.

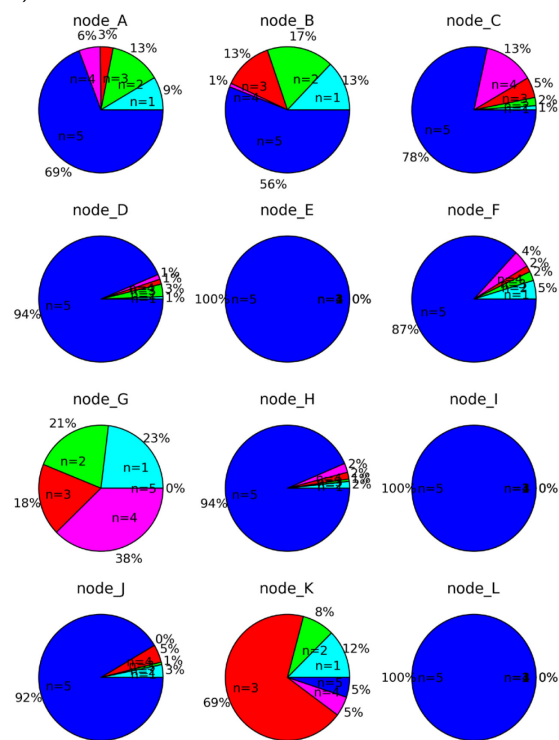
a)



b)

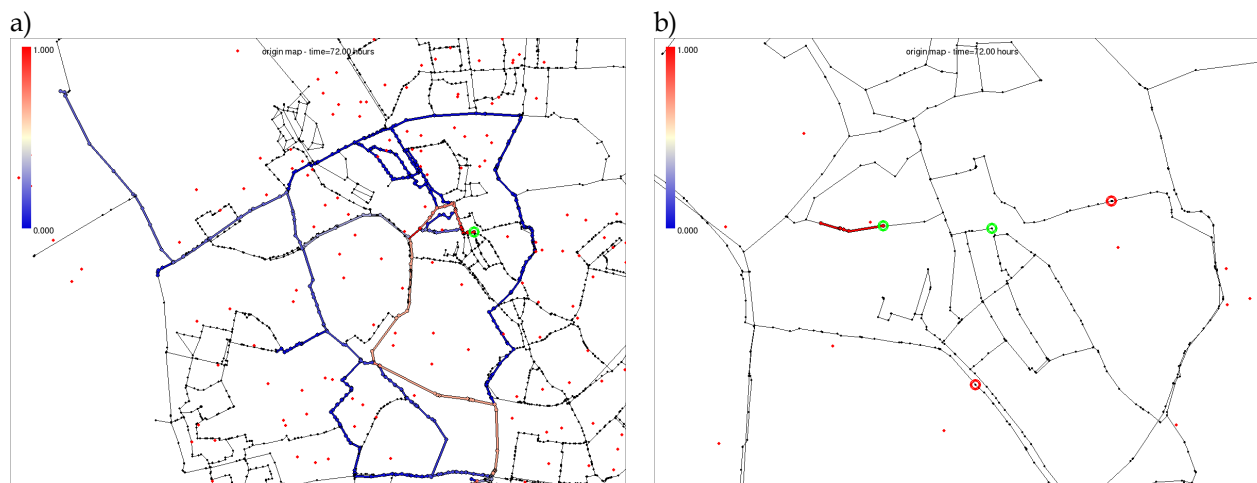


c)



Figuur 25: Padverdeling voor backtraces in 5 verschillende realisaties van het stromingsveld voor twaalf geselecteerde knooppunten in de detailwijk. a) Gestapelde herkomstkaarten voor knooppunt node\_K; b) locaties van geselecteerde knooppunten; c) Taartdiagrammen voor de gestapelde herkomstkaarten voor de twaalf geselecteerde knooppunten. Deze taartdiagrammen geven aan welke fractie van alle knooppunten die in de backtraces voor de 5 realisaties zijn aangetroffen in welk aantal backtraces voorkomt. NB: in deze beschouwing worden de backtraces slechts tot aan de rand van de detailwijk gevolgd.

a)



Figuur 27: Herkomstkaart van een enkele sensor (groen omcirkeld) in een verdund gedeelte van het netwerkmodel, met omvangrijk potentieel herkomstgebied (kleuren geven de relatieve bijdrage van de herkomstgebieden in de volumestroom ter plekke van de sensor weer). b) Inperking van de omvang van het potentiële herkomstgebied door combinatie van positieve (groen omcirkeld) en negatieve (rood omcirkeld) sensormetingen voor hetzelfde tijdstip (binaire weergave: mogelijk of niet mogelijk).

## 4.5 Almere

### 4.5.1 Variabiliteit van het stromingsveld

Voor het netwerk van Almere wordt de variabiliteit van het stromingsveld gedurende een periode van 24 uur geïllustreerd in Figuur 28. Het beeld komt overeen met dat van Eindhoven: de variatie in de volumestroom (Figuur 28a) is duidelijk in zowel enkele segmenten van het (secundaire) verdunde deel van het netwerk als in een groot deel van het tertiaire netwerk in de detailwijk. Regelmatige omkeringen van de stroomrichting (Figuur 28b) zowel op secundair als op tertiair niveau terug.

### 4.5.2 Herkomst en dekking

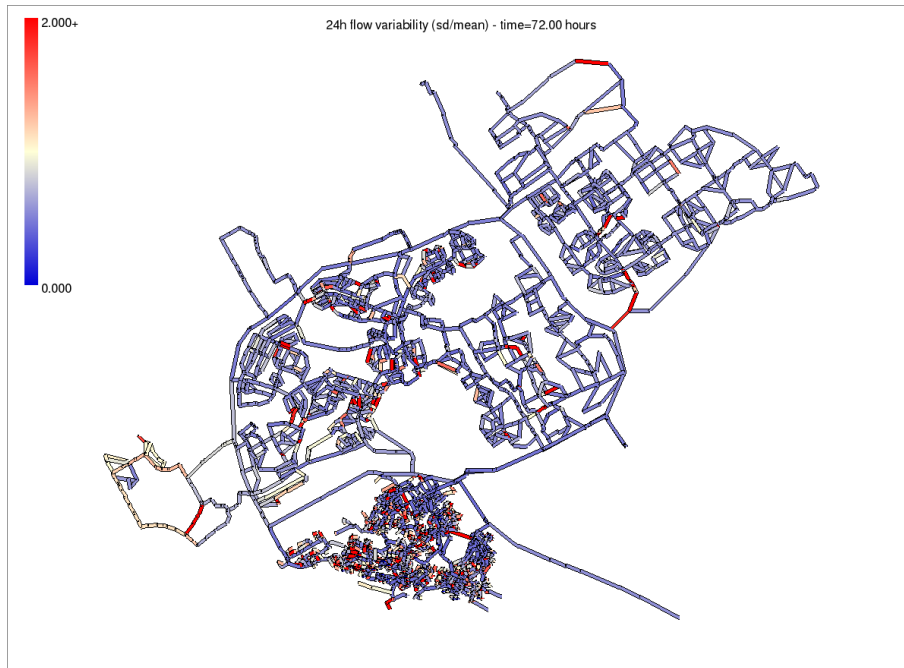
Ook voor Almere wordt, vanwege het schier oneindig grote aantal combinaties, slechts een selectie van herkomst- en dekkingskaarten gepresenteerd.

Figuur 29 toont herkomstkaarten voor een selectie van 24 sensorlocaties overeenkomend met 24 monsterlocaties in de onverdunde wijk. Uitgangstijdstip is middernacht na 72 uur simulatietijd. Weergegeven zijn herkomstkaarten in een detailweergave van de wijk en zijn directe omgeving. Kaarten zijn gemaakt voor één realisatie van het snelheidsveld (Figuur 29a) en voor een combinatie van 5 realisaties (Figuur 29b). Elk van de kaarten geeft voor iedere locatie in het netwerk met een kleur aan door hoeveel sensoren de betreffende locatie voor dit tijdstip wordt gedekt. Het rood weergegeven deel van het netwerk loopt direct door naar het pompstation. Hier vallen slechts 21 i.p.v. 24 paden samen, doordat op enkele van de aan monsterpunten gekoppelde knooppunten geen watervraag in het model blijkt te zitten. Figuur 30 toont 24-uurs dekkingskaarten voor deze zelfde combinaties. Voor beide kaartsoorten blijkt de combinatie van 5 realisaties niet tot een wezenlijk ander beeld te leiden; wel is het effectieve 24-uurs dekkingsgebied opnieuw iets kleiner dan voor een enkele realisatie.

Uitgaande van de 24 sensorlocaties is de dekking van de detailwijk beperkt. Om deze reden is ook een herkomstkaart gemaakt voor alle 52 netwerkknopen die aan alle monsterlocaties uit de wijk uit de periode 2009-2012 (Figuur 29c). Deze toont een zeer vergelijkbaar beeld. Dit komt doordat de monsterlocaties sterk geclusterd zijn.

De verschillen tussen de instantane herkomstkaart van Figuur 29a en de 24-uurs dekkingskaart van Figuur 30 zijn minimaal. Dit duidt erop dat de monsterpunten via niet nauwelijks door de dag

a)



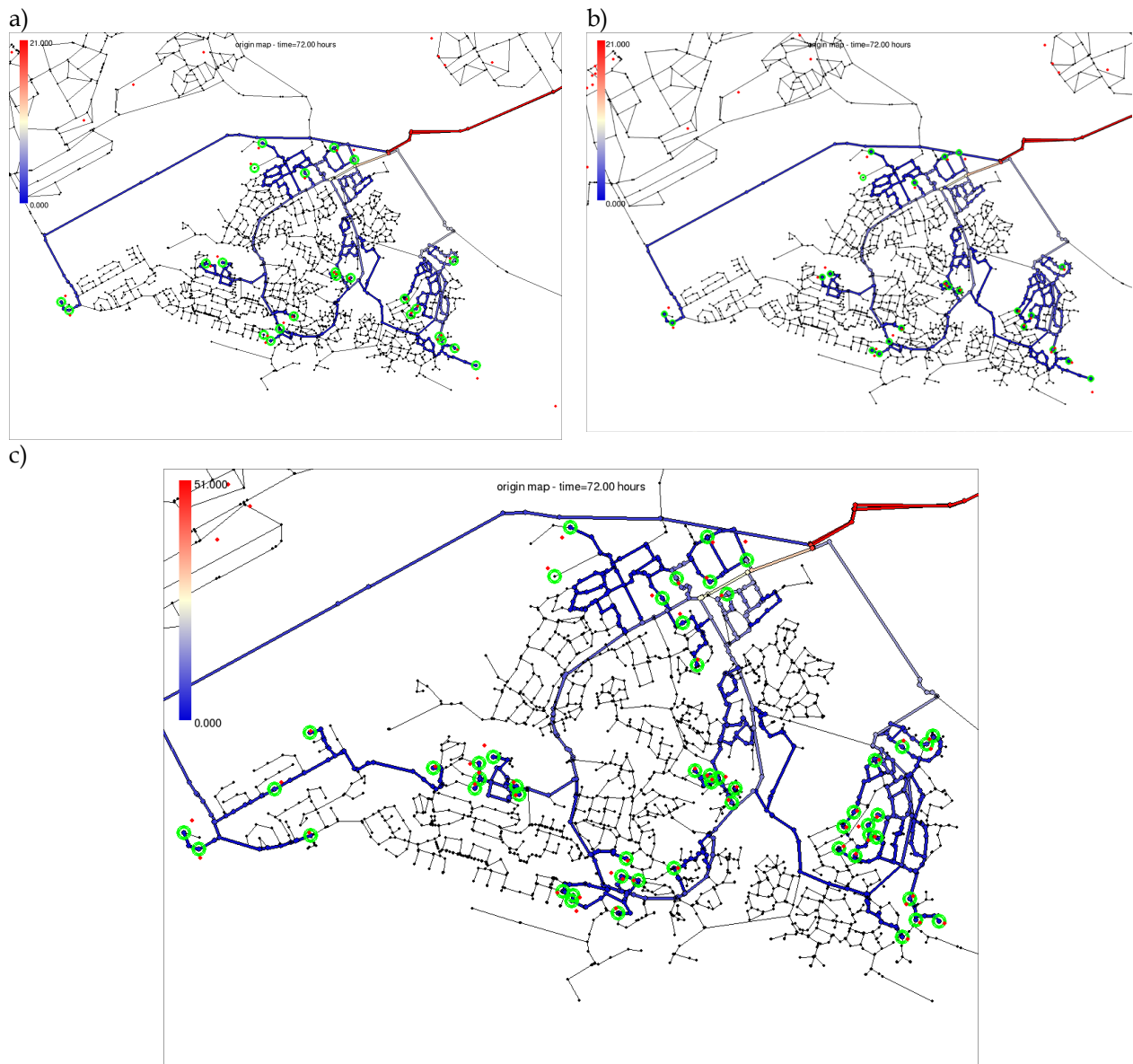
b)



Figuur 28: Illustratie van de variabiliteit van het stromingsveld in het netwerk van Almere gedurende een periode van 24 uur: a) Standaarddeviatie van de volumestroom gedeeld door de gemiddelde volumestroom; b) fractie van de tijd dat de stroomrichting omgekeerd is ten opzicht van de meestvoorkomende stroomrichting. Beide figuren zijn gebaseerd op 12 tijdstippen (interval 2 uur) in 1 realisatie van het stromingsveld en daarom slechts indicatief.

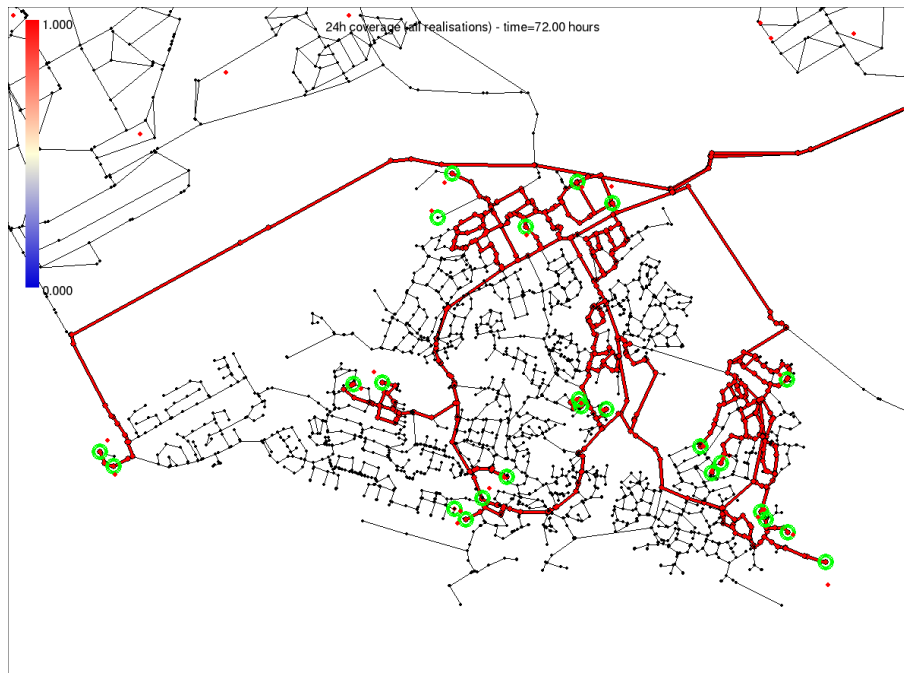
veranderende paden door het netwerk worden gevoed. Deze eigenschap maakt *backtracing* van verontreinigingen minder gevoelig voor het tijdstip op de dag en daarmee in principe eenvoudiger.



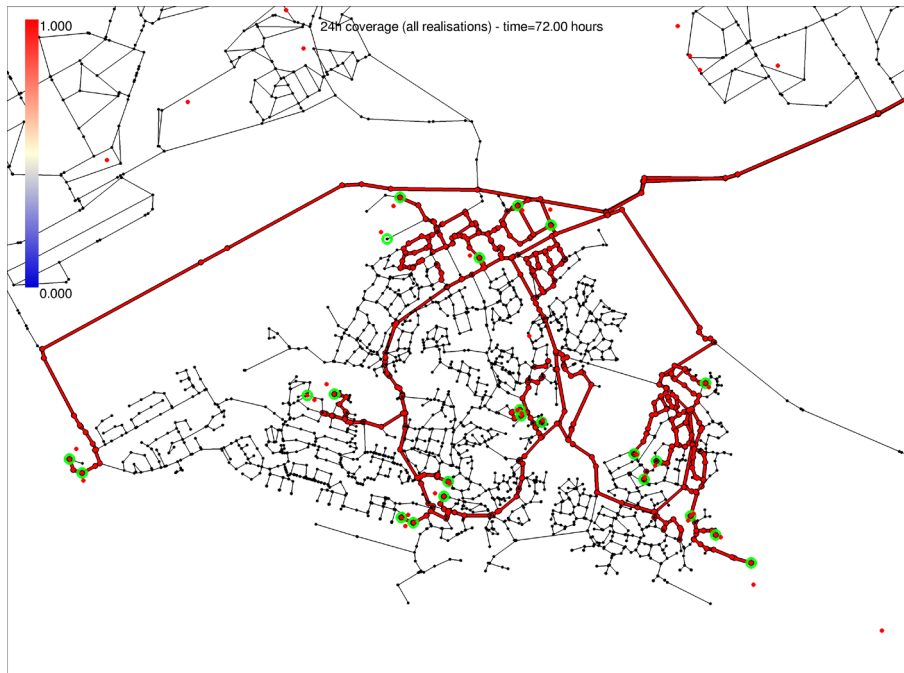


Figuur 29: Herkomstkaarten (referentietijdstip 00:00) voor Almere voor een combinatie van 24 sensoren gekoppeld aan 24 monsterlocaties in de onverdunde wijk. a) gebaseerd op 1 realisatie van het snelheidsveld; b) gebaseerd op een combinatie van 5 realisaties; c) herkomstkaart voor 52 monsterlocaties. De kleurschaal geeft aan hoeveel sensorherkomstgebieden elkaar op ieder punt overlappen. Monsterpunten zijn in rood weergegeven, sensorlocaties zijn groen omcirkeld.

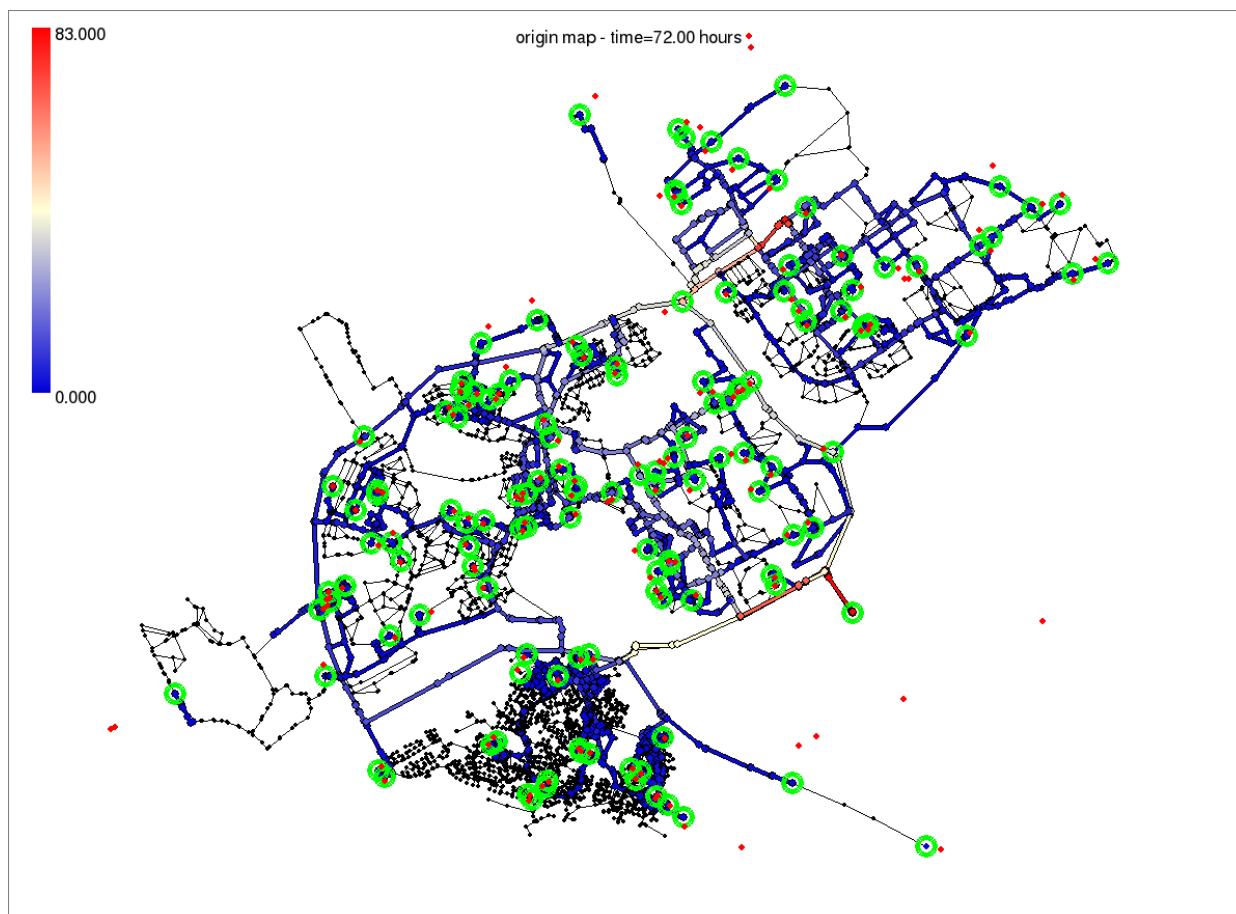
a)



b)



Figuur 30: 24-Uurs dekkingskaarten (referentietijdstip 00:00) voor Almere voor een combinatie van 24 sensoren gekoppeld aan 24 monsterlocaties in de onverdunde wijk. Een detailweergave van de wijk (a,b) en een volledige weergave van het relevante deel van het netwerk (c,d) zijn opgenomen. Kaarten zijn gebaseerd op 1 realisatie van het snelheidsveld (a,c) en een combinatie van 5 realisaties (b,d). Gebieden met dekking gedurende 24 uur (niet noodzakelijkerwijs continu door dezelfde sensor) zijn rood gemarkeerd.



Figuur 31: Herkomstkaart voor Almere (referentietijdstip 00:00) voor een combinatie van 170 sensoren corresponderend met alle geselecteerde monsterlocaties in het model voor 1 realisatie van het snelheidsveld. De kleurschaal geeft aan hoeveel sensorherkomstgebieden elkaar op ieder punt overlappen.

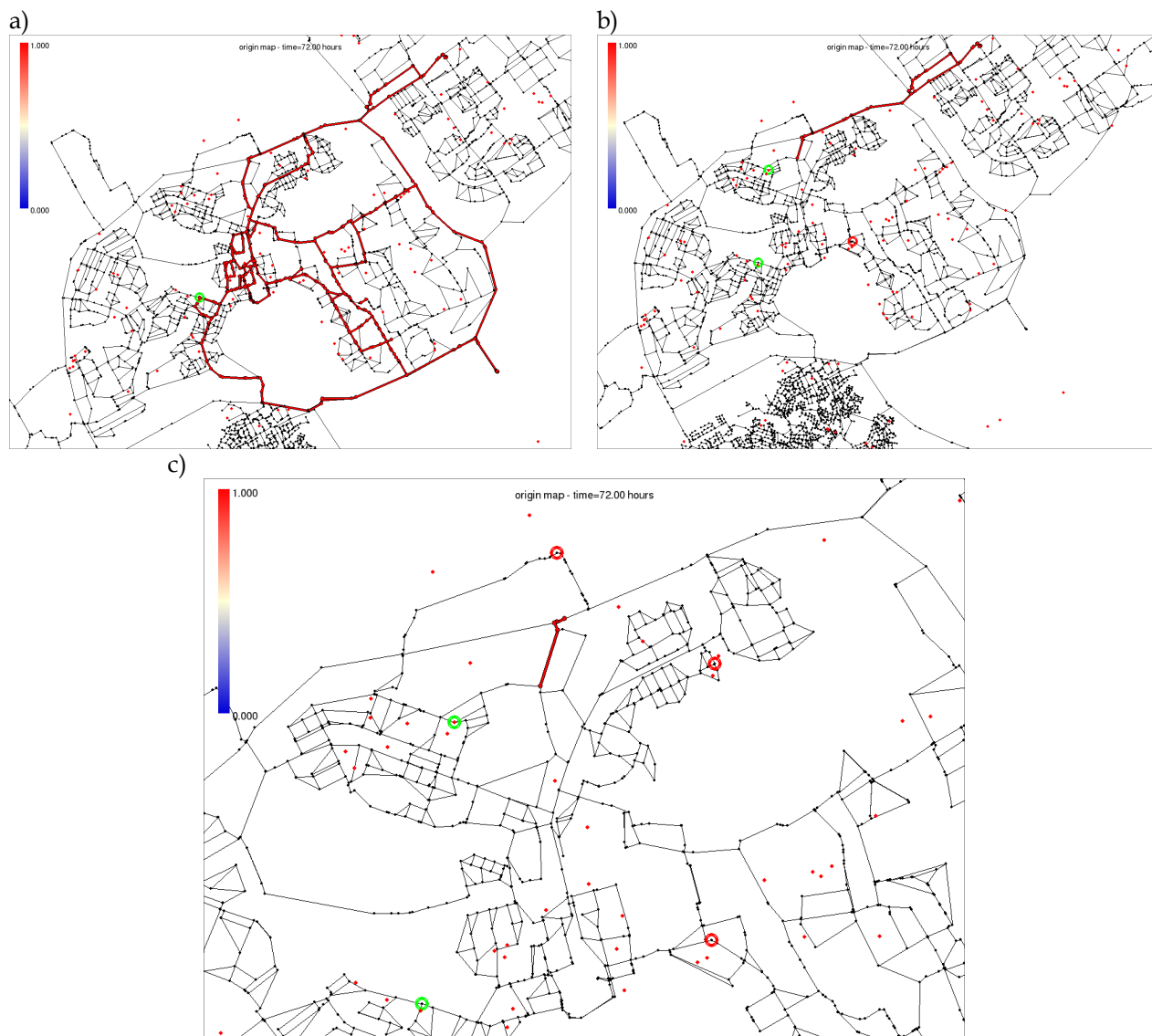
#### 4.5.3 Bronbepaling

Figuur 32 toont een voorbeeld van de inperking van het potentiële brongebied door combinatie van positieve en negatieve sensormetingen. Let op dat we hierbij aannemen dat de “positieve” sensoren dezelfde verontreiniging (in de zin van dezelfde bron) meten. Met een beperkt aantal sensoren blijkt het potentiële brongebied opnieuw in te perken tot slechts enkele segmenten van het leidingenmodel.

Het is denkbaar dat een combinatie van positieve en negatieve bepalingen niet tot een potentieel herkomstgebied leidt. Dit kan duiden op hetzij een probleem met de sensorbepalingen (vals-positief danwel vals-negatief) hetzij op een afwijking van het gebruikte hydraulische model t.o.v. de werkelijke situatie, bijvoorbeeld ten gevolge van incorrect geregistreerde afsluiterstanden.

#### 4.6 Vergelijking van Eindhoven en Almere

Tabel 4 geeft een overzicht van de belangrijkste karakteristieken en resultaten van netwerkmodellen van Eindhoven en Almere. Hoewel de lusedichtheid in de detailwijken zeer vergelijkbaar is (circa 3 lussen per kilometer leiding), tonen de gemaakte dekkingskaarten een grotere mate van dekking in de Eindhovense detailwijk (Figuur 23) dan in de Almeerse (Figuur 30). Dit is waarschijnlijk te wijten aan de clustering van monsterpunten in Almere. In beide gevallen blijkt inperking van een potentieel brongebied goed mogelijk.



Figuur 32: Inperking van het potentiële brongebied van een verontreiniging door successieve toevoeging aan de beschouwing van sensoren die de verontreiniging wel (groen) of niet (rood) detecteren op hetzelfde tijdstip.

Tabel 4: Vergelijking van karakteristieken en resultaten van de onderzochte netwerkmodellen van Almere en Eindhoven.

	Almere	Eindhoven
leidinglengte in netwerkmodel - totaal	493 km	281 km
leidinglengte in netwerkmodel - detailwijk	99 km	35 km
aantal lussen in netwerkmodel - totaal	1029	322
aantal lussen in netwerkmodel - detailwijk	316	105
lussen / km - totaal	2,3	1,1
lussen / km - detailwijk	3,2	3,0
dekking monsterpunten	39% van knooppunten in netwerkmodel	67% van knooppunten in netwerkmodel
inperkingsmogelijkheid brongebied	goed	goed

## **4.7 Gevoeligheidsbeschouwing**

### **4.7.1 Afsluitermanipulatie**

*Backtrace*berekeningen zijn alleen betekenisvol bij een correcte weergave van de werkelijkheid in het netwerkmodel. Wanneer bijvoorbeeld de netwerkconfiguratie wordt veranderd door het sluiten van een afsluiter, kan dit, afhankelijk van de locatie van de afsluiter, beperkte tot aanzienlijke gevolgen hebben voor herkomstberekeningen. Dit wordt geïllustreerd aan de hand van een voorbeeld in Figuur 33. Een van de drie toevoerleidingen van de detailwijk wordt afgesloten (Figuur 33a). Het gebied dat in de reguliere situatie door deze leiding van water wordt voorzien (Figuur 33b) moet nu via de andere twee toevoerleidingen gevoed worden; het oorspronkelijke pad (Figuur 33c) is niet meer geldig. Dit betekent dat ook de oorspronkelijke herkomstbepaling (Figuur 33d) verschuift (Figuur 33e).

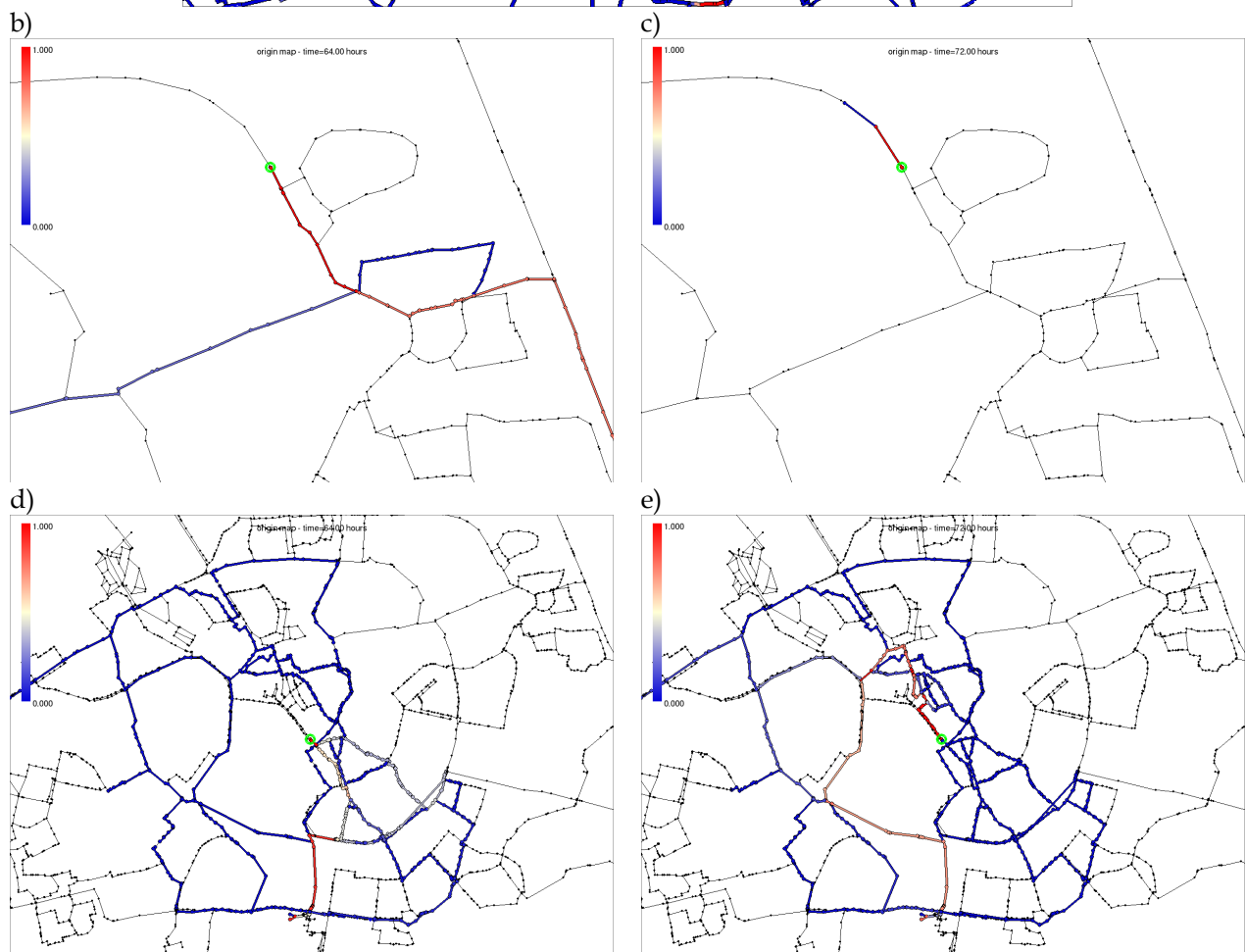
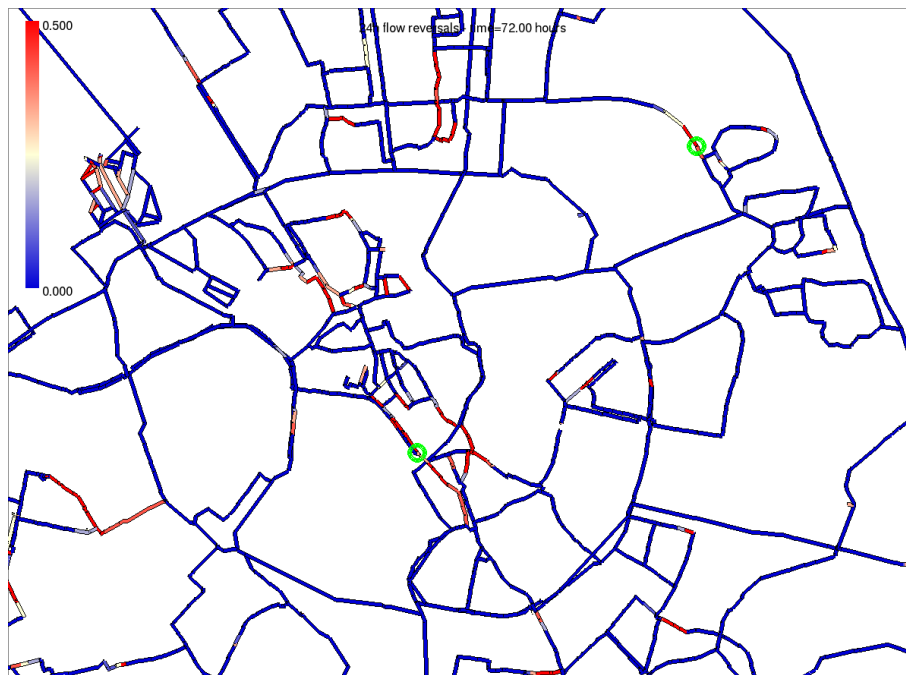
### **4.7.2 Tijdstip van meting**

Het ligt voor de hand dat *backtracing* in delen van het netwerk waar de stroming(srichting) niet sterk varieert door de dag heen (zie Figuur 19 en Figuur 28) gemakkelijker en zekerder is dan in gebieden waar deze variatie groter is. Om deze reden wordt de variatie van de bepaling van het potentiële brongebied met het tijdstip op de dag geïllustreerd aan de hand van een deel van het netwerk waar deze variatie relatief groot is. Gebieden waar omkeringen van de stroomrichting plaatsvinden kunnen het ene moment vanaf de ene kant worden gevoed en het andere moment van de andere kant, waarbij een geheel andere herkomstkaart hoort (zie Figuur 34). Hier zou een combinatie van de waterkwaliteitssensor met een debietmeter uitsluitel kunnen geven.



Figuur 33: Consequenties van het sluiten van een afsluiter op een gevoelige locatie voor bronbepalingsberekeningen. a) locatie waar een afsluiter dicht wordt gezet; b) gebied dat in de oorspronkelijke configuratie (open afsluiter) (deels) door deze leiding van water wordt voorzien; c) herkomst van het water dat door deze leiding stroomt; d) herkomstbepaling van een locatie (groen) in het kritische gebied voor de oorspronkelijke configuratie; e) herkomstbepaling van een locatie (groen) in het kritische gebied voor de alternatieve configuratie met gesloten afsluiter.

a)



Figuur 34: Verschillende herkomstkaarten voor verschillende tijdstippen (b,d: 64 uur, c,e: 72 uur) op twee locaties in Eindhoven die worden gekenmerkt door regelmatige omkeringen in de stroomrichting (a).

#### **4.8 Nut en noodzaak**

Hoewel er sterke, o.a. vanuit de technologie gedreven, belangstelling is voor het gebruik van sensing in het distributienet, is de discussie wat een waterbedrijf er daadwerkelijk mee gaat doen en waarvoor sensing nuttig is, van groot belang. Ontwikkelingen in kwaliteit en prijs van de sensoren, onderzoek naar methodieken om informatie uit sensordata te halen, maar ook vragen en enthousiasme vanuit de waterbedrijven om sensoren optimaal te gebruiken zijn hierbij bepalende randvoorwaarden.

Het nut en de noodzaak van backtracing laten zich vatten in de volgende redenering:

- backtracing is een relatief goedkope en betrouwbare methode die de mogelijkheid biedt om paden van verontreinigingen in het net te berekenen, en daarmee ook het nut van sensorlocaties en beschermingsgebieden van sensoren aan te tonen;
- buiten een methode als bovenstaand bestaat er geen betrouwbare en tegelijkertijd relatief goedkope benadering die dezelfde informatie oplevert;
- zonder de informatie die backtracing kan leveren, kunnen er geen garanties, uitspraken of vuistregels over besmettingsrisico's of dekkingsgebieden worden gemaakt.

Vanuit economisch standpunt is het op enkele plaatsen "op gevoel" sensoren plaatsen wellicht de gunstigste oplossing, maar vanuit het oogpunt van garanties voor een (nagenoeg) onberispelijke waterkwaliteit en betrouwbare levering voor alle aansluitingen, volstaat deze benadering niet. En juist met de hier beschreven methodieken kunnen de risico's helder in kaart worden gebracht bij plaatsing van een zo gering mogelijk aantal sensoren.

De resultaten uit dit onderzoek laten zien hoe goed een rekenmodel en slimme plaatsing van sensoren kunnen bijdragen aan het vinden van besmettingsbronnen en hoe groot deze bijdrage kan zijn. Hiermee vormen de resultaten een motivatie voor plaatsing van waterkwaliteitsensoren, en zeker in combinatie met de methodiek als in dit rapport beschreven. De methodiek om met behulp van backtracen een besmetting te herleiden naar de bron is waardevol, omdat hiermee sneller en effectiever ingegrepen kan worden wanneer een verontreiniging wordt geconstateerd (isolatie, informeren van de consument en verwijdering).

Ook zonder het plaatsen van fysieke sensoren biedt de ontwikkelde methodiek een handvat om, in het geval van een geconstateerde verontreiniging (een meting en/of klachten van klanten) 'eenvoudig' en goedkoop te bepalen waar aanvullend gemeten moet worden om de bron te kunnen opsporen. Op dit moment is dat in de Nederlandse drinkwatersector nog niet mogelijk.





# 5 Conclusies en aanbevelingen

## 5.1 Conclusies

### 5.1.1 Benaderingen

- Er is een Lagrangiaanse benadering van *backtracing* (combinatie van voortwaarts-Lagrangiaanse berekeningen) uitgewerkt, geïmplementeerd, getest en toegepast op de netwerken van Almere en Eindhoven. Deze benadering werkt goed, ook in de stochastische toepassing met verschillende vraagpatronen en realisaties van het stromingsveld.
- Er is een Euleriaanse inversiebenadering voor het *backtracen* voorgesteld. Deze is mathematisch uitgewerkt en geïmplementeerd in een experimentele code. De benadering is interessant, maar nog niet voldoende getest voor toepassing in echte distributienetten. Bovendien lijkt de maximaal haalbare grootte van de door te rekenen netwerkmodellen beperkt.
- Momenteel heeft de expliciete Lagrangiaanse benadering van *backtracing* de voorkeur boven de Euleriaanse inversiebenadering, ook voor stochastisch benaderde problemen.
- Op basis van de beschikbare informatie kan worden geconcludeerd dat de hier ontwikkelde methoden verder gaan en meer informatie opleveren dan verschillende beschikbare software voor *backtracing*.

### 5.1.2 Toepassing van Lagrangiaanse methodiek op Eindhoven en Almere

- Uitgaande van alle monsterlocaties uit 2011 in Eindhoven als sensorlocaties moet worden vastgesteld dat in de dekking van sensoren op deze locaties onvolledig is in de periferie van het netwerk (NB: dit is geen uitspraak over de wenselijkheid van volledige dekking).
- Voor Almere wordt met de gekozen monsterlocaties een vrij beperkte dekking gevonden (lager dan in Eindhoven) voor de detailwijk als men uitgaat van dezelfde sensordichtheid (in aantal sensoren per netwerklus). Dit verandert niet wanneer alle monsterlocaties uit de periode 2009-2012 worden bekeken, aangezien deze sterk geclusterd blijken te zijn.
- Het betreft in beide gevallen echter een fictief sensornetwerk. Een betere dekking voor deze wijken met hetzelfde aantal sensorlocaties is goed denkbaar, wanneer dekking als optimalisatiecriterium voor de plaatsing wordt toegepast. Dit is onderwerp van gepland onderzoek.
- Een stochastische benadering van de watervraag en daarmee het stromingsveld leidt lokaal tot iets andere transportpaden en daarmee resultaten. In de regel zijn de verschillen echter klein en volstaat een niet-stochastische benadering op de meeste onderzochte ruimtelijke en tijdschalen (bij de veronderstelling dat het gebruikte aantal van 5 realisaties van het snelheidsveld voldoende groot is). Uitzondering hierop vormt de periferie van het tertiaire netwerk, waar knooppunten worden gevonden die voor andere realisaties van het stromingsveld compleet andere *backtraces* laten zien.
- Op een beperkt aantal locaties in het netwerk is er grote gevoeligheid van de herkomstkaarten voor het tijdstip van bepaling, doordat water op verschillende momenten van de dag de plek via verschillende routes bereikt. Dit illustreert het belang van een goed hydraulisch model.
- De verandering van de effectieve netwerkconfiguratie door het dicht- of opendraaien van een afsluiter kan in sommige gevallen grote gevolgen hebben voor de oorsprong van water op diverse punten in het netwerk. Dit betekent dat een correcte weergave van de netconfiguratie essentieel is voor het correct kunnen *backtracen* van verontreinigingen.
- Door het combineren van positieve (verontreiniging wordt vastgesteld) en negatieve (vastgesteld wordt dat er geen verontreiniging is) sensormetingen is het gebied waar de besmettingsbron moet worden gezocht significant in te perken, soms tot slechts enkele leidingsegmenten in het netwerkmodel.

## 5.2 Nut en noodzaak

De ontwikkelde methodiek ondersteunt de bepaling van een zinvolle toepassing van waterkwaliteitssensoren in het distributienetwerk. Op basis van metingen van slim geplaatste sensoren kan met toepassing van de methodiek in geval van calamiteiten worden bepaald op welk deel van het netwerk de aandacht geconcentreerd dient te worden. Hiermee kan de verspreiding van een verontreiniging sneller worden beperkt en worden de gevolgen geminimaliseerd. Dit vermogen vormt de toegevoegde waarde van waterkwaliteitssensoren in het distributienet in combinatie met hydraulische modellen.

## 5.3 Aanbevelingen

### 5.3.1 Voor toepassing

- Waterbedrijven wordt aanbevolen om, wanneer zij waterkwaliteitssensoren in hun netwerk plaatsen, een *backtrace* methodiek en een procedure voor het combineren van bepalingen voor individuele sensoren, zoals hier uitgewerkt, te implementeren. Zo kan in geval van een geconstateerde verontreiniging snel het potentiële brongebied worden ingeperkt. Hiermee blijft enerzijds de toepassing van correctieve maatregelen tot een klein gebied beperkt en wordt anderzijds de bron zo snel mogelijk gevonden. Momenteel in gebruik zijnde software biedt hiervoor onvoldoende functionaliteit.
- Om deze inperking zo effectief mogelijk te maken, wordt aanbevolen om dit aspect in de locatiekeuze voor de sensoren mee te nemen. Hiervoor wordt in gepland BTO-onderzoek een methodiek ontwikkeld.
- Ook voor het interpreteren van andere bronnen van informatie over waterkwaliteitsproblemen (klantenklachten, “*The Human Sensor*”), biedt de gepresenteerde methodiek een hulpmiddel om snel het potentiële brongebied van de gemelde problemen in te perken.
- *Backtrace* berekeningen hebben echter alleen waarde wanneer zij uitgaan van een correcte weergave van de werkelijkheid. Dit betekent dat zij moeten worden gebaseerd op een goed, actueel hydraulisch model van het leveringsgebied met correcte afsluiterstanden en bedrijfsvoering. Voor snelle toepassing in crisissituaties vergt dit een continue controle en actualisering van het hydraulische model en een volledige en correcte administratie van afsluitermanipulaties en bedrijfsvoeringsacties.
- De ontwikkelde methodiek is opgezet rondom EPANET. Voor toepassing bij en door de waterbedrijven zelf is implementatie in daar gangbare hydraulische software noodzakelijk.

### 5.3.2 Voor vervolgonderzoek

De volgende punten worden aanbevolen voor verder onderzoek:

- verdere uitwerking en het zoeken van een oplossing voor de opgemerkte problemen met grote netwerken in Euleriaanse inversiebenadering, aangezien toepassing van deze methode naar verwachting grote voordelen in termen van rekentijd en benodigde opslagruimte biedt;
- het uitdrukken van (de omvang van) beschermingsgebied van een sensor in termen van het aantal aansluitingen of afgenomen volume;
- optimalisatie van sensorplaatsing op basis van omvang en verdeling van potentiële herkomstgebieden bij multisensormetingen - dit idee zal verder worden uitgewerkt in het BTO-onderzoek van 2013-2014;
- gebruik maken van passagetijdstippen om tot betere inperking en bepaling van tijdstip van aanvang verontreiniging te komen;
- gebruik maken van daadwerkelijk gemeten concentraties in plaats van binaire benadering om tot betere inperking te komen;
- gebruik van detectielimieten van sensoren;
- inschatting van de effecten van vals-positieve en vals-negatieve metingen;
- uitwerking van concept voor diachrone (niet-gelijktijdige) observaties;
- combinatie van meerdere soorten metingen (sensormetingen, labanalyses, klantmeldingen, etc.);
- hoeveel realisaties van het snelheidsveld zijn nodig voor een goede representatie van de daadwerkelijk optredende variabiliteit?;
- op welk detailniveau is het nog belangrijk de stochastische benadering toe te passen?;

- wat zijn een geschikte maat en richtlijn voor sensordichtheid (per aantal aansluitingen, per aantal lussen, etc)?;
- toepassing van volumestroom- en drukmetingen voor de controle van afsluiterstanden in een hydraulisch model.



## 6 Referenties

- Berry, J., W.E. Hart, C.E. Phillips, J.G. Uber en J.-P. Watson (2006), *Sensor placement in municipal water networks with temporal integer programming models*. J. Water Resour. Plann. Manage., 132\_4\_, 218–224.
- Berry, J., E. Boman, L. A. Riesen, W. E. Hart, C. A. Phillips, J.P. Watson, R. Murray (2010) *User's Manual TEVA-SPOT Toolkit Version 2.4*. Office of Research and Development, U.S. Environmental Protection Agency, EPA 600/R-08/041B.
- Blokker, E.J.M. (2010) *Stochastic water demand modelling for a better understanding of hydraulics in water distribution networks*. Proefschrift, TU Delft.
- Costa, D.M., Melo, L.F. , Martins, F.G. (2011) *Localization of Contamination Sources in a Drinking Water Distribution System: A Method Based on the Residence Time of Water in Pipes*. 11th International Chemical and Biological Engineering Conference, Lisbon.
- De Sanctis, A. E., Shang, F., Uber, J. G. (2010) *Real-time identification of possible contamination sources using network backtracking methods*. Journal of Water Resources Planning and Management, 136 (4), 444-453.
- Kim, T-H, White, H. On more robust estimation of skewness and kurtosis: Simulation and application to the S & P 500 index, discussion paper (grey literature), 2003.
- Laird, C.D., Biegler, L.T., van Bloemen Waanders, B.G., en Bartlett, R.A. (2005) *Contamination source determination for water networks*. Journal of Water Resources Planning and Management, 131(2), 125-134.
- Laird, C.D., Biegler, L.T., en van Bloemen Waanders, B.G. (2006) *Mixed-integer approach for obtaining unique solutions in source inversion of water networks*. Journal of Water Resources Planning and Management, 132(4), 242-251.
- Muntendam-Bos, A. G., Kroon, I. C., and Fokker, P. A. (2008) *Time-dependent Inversion of Surface Subsidence due to Dynamic Reservoir Compaction*, Math. Geosci., 40, 159–177, doi:10.1007/s11004-007-9135-3.
- Ostfeld, A. , James G. Uber, Elad Salomons, Jonathan W. Berry, William E. Hart, Cindy A. Phillips, Jean-Paul Watson, Gianluca Dorini, Philip Jonkergouw, Zoran Kapelan, Francesco di Pierro, Soon-Thiam Khu, Dragan Savic, Demetrios Eliades, Marios Polycarpou, Santosh R. Ghimire, Brian D. Barkdoll, Roberto Gueli, Jinhui J. Huang, Edward A. McBean, William James, Andreas Krause, Jure Leskovec, Shannon Isovitsch, Jianhua Xu, Carlos Guestrin, Jeanne VanBriesen, Mitchell Small, Paul Fischbeck, Ami Preis, Marco Propato, Olivier Piller, Gary B. Trachtman, Zheng Yi Wu en Tom Walski (2008) *The Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms*, J. Water Resour.Plann. Manage., 134(6), 556-568.
- Preis, A., en Ostfeld, A. (2006) *Contamination source identification in water systems: A hybrid model trees-linear programming scheme*. Journal of Water Resources Planning and Management, 132(4), 263-273.
- Python (2013) <http://www.python.org>, laatst benaderd op 24 april 2013.
- Rossman, L. A. (2000) *EPANET 2 Users Manual*. U.S. Environmental Protection Agency, report EPA/600/R-00/057.
- Shang, F., Uber, J. G., Polycarpou, M. M. (2002) *Particle backtracking algorithm for water distribution system analysis*. Journal of Environmental Engineering, 128 (5),441-450.

Shang, F., Uber, J. G., Rossman, L. (2008) *EPANET Multi Species Extension Software*. U.S. Environmental Protection Agency, Washington D.C., EPA/600/C-10/002.

Tarantola, A. (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, Philadelphia.

Van Thienen, P. en R. Floris (2012) *Optical tomography for drinking water discolouration research – Methodology, hardware, software, and validation*. BTO 2012.240(s).

# I Verklarende woordenlijst

<i>24-uurs beschermingskaart</i>	Een kaart die voor een of meer meetlocaties in het netwerk aangeeft voor een of meer realisaties van het stromingsveld welke delen van het netwerk gedurende de afgelopen 24 uur continu stroomafwaarts gepositioneerd zijn geweest van deze meetlocatie(s); d.w.z. dat gedurende deze 24 uur continu water van de meetlocatie naar deze gebieden is gestroomd.
<i>24-uurs dekkingskaart</i>	Een kaart die voor een of meer meetlocaties in het netwerk aangeeft voor een of meer realisaties van het stromingsveld welke delen van het netwerk gedurende de afgelopen 24 uur continu in beeld is geweest van deze meetlocatie(s); d.w.z. dat gedurende deze 24 uur continu water van de gemarkeerde gebieden langs de meetlocatie(s) is gestroomd.
<i>deterministische benadering discretisatie</i>	Op basis van de huidige toestand van een systeem is met een goed model exact een toekomstige toestand te voorspellen. Het (wiskundig) beschrijven van (een eigenschap van) een continu object in een eindig (discreet) aantal punten. Tussen deze punten wordt een geleidelijke (lineair of anders) variatie van de eigenschap verondersteld.
<i>Euleriaans referentiekader</i>	Bij het transport van water wordt het vastliggende leidingnet als referentiekader aangenomen. Het leidingnet wordt gediscrètiseerd in een groot aantal dicht op elkaar liggende knooppunten. De concentratie van een tracerstof wordt op elk van deze knooppunten beschreven. Telkens wordt de uitwisseling van de stof tussen knooppunten als functie van het stromingsveld berekend.
<i>herkomstkaart</i>	Een kaart die voor een bepaald tijdstip voor een bepaalde meetlocatie aangeeft in welke delen van het netwerk het water is geweest dat op dit tijdstip op deze meetlocatie wordt bemonsterd.
<i>herkomstkanskaart</i>	Een kaart die voor een bepaald tijdstip voor een bepaalde meetlocatie aangeeft welke fractie van het bemonsterde water uit welke delen van het netwerk afkomstig is. Wanneer men een uniforme verontreinigingskans veronderstelt, is deze fractie een maat voor de herkomstkans.
<i>Lagrangiaans referentiekader</i>	Bij het transport van water wordt een bewegend 'pakketje' water (met daarin een of meerdere tracerstoffen) als referentiekader aangenomen. Voor het pakketje wordt steeds een nieuwe locatie berekend op basis van het stromingsveld in het netwerk.
<i>probabilistische benadering</i>	Vanuit een huidige situatie, die mogelijk niet volledig bekend is, wordt met een model dat zelf ook onzekerheden kent een kansverdeling voor een toekomstige toestand bepaald.
<i>realisatie</i>	Een van (vele) mogelijke of plausibele verschijningsvormen van een stochastisch systeem. Voorbeeld: Het herhaald gooien van een dobbelsteen is een stochastisch proces. Elk van de zes mogelijke uitkomsten van de worp is een realisatie.
<i>regularisatie</i>	Het toepassen van randvoorwaarden van mathematische aard (dus niet gerelateerd aan het fysieke distributienet en de processen daarin) op een complex probleem om het oplossingsproces te vergemakkelijken (betekenins in de context van dit rapport).
<i>stochastisch systeem</i>	Een systeem dat zich in enige mate willekeurig gedraagt. Van dit systeem is niet te voorspellen in welke toestand het zich in de toekomst zal bevinden, hooguit de kans dat het zich in een bepaalde toestand bevindt. Voorbeeld: Het gooien van een dobbelsteen.





## II Opzet van backtracing met voorwaarts Lagrangiaanse benadering

<i>stap</i>	<i>process</i>	<i>programma</i>
1	Maak reguliere EPANET-inputfile (.INP) met gewenste verbruikspatronen (meerder files voor meerdere sets stochastische patronen).	bijvoorbeeld via export uit Infoworks of Synergiee
2	Voer de reguliere hydraulische modelberekening uit voor een of alle bestanden.	EPANET
3	Genereer invoerbestanden voor EPANET-MSX voor ieder EPANET-invoerbestand. Voor ieder knooppunt in het netwerkmodel worden twee invoerbestanden voor EPANET-MSX aangemaakt, een voor de locatitracer $M_i$ en een voor de tijdstracer $T_i^+$ .	EPANETinp2msx.py*
4	Voer transportberekening uit voor ieder gegenereerd invoerbestand	EPANET-MSX
5	Voeg de resultaten van de individuele transportberekeningen samen (met compressie)	joinmsxresults.py*
6	Interactieve visualisatie en interpretatie	wxreadepanet.py*

+ Tijdens het project is gebleken dat wanneer er meerdere tracers in een bestand worden gedefinieerd, er soms discontinuïteiten ontstaan in herkomstkaart. De oorzaak van dit probleem is niet achterhaald, maar deze *workaround* biedt een werkbare oplossing.

\* in het kader van dit project ontwikkeld



# III Discreet mathematisch-fysisch model voor stoftransport met probabilistische benadering

## Toegepaste basis

We gaan uit van een netwerkmodel waarop een discreet aantal knooppunten  $p_i$  is gedefinieerd op punten  $x_i$ . Op deze knooppunten is een snelheidsveld  $u_i(t)$  als functie van de tijd bekend (uit simulaties). Op deze punten is ook een stofconcentratie  $c_i$  gedefinieerd. Het transport van deze stof wordt in een Euleriaans raamwerk beschreven met de volgende vergelijking (zie Rossman, 2000):

$$\frac{\partial c}{\partial t} = -u_i \frac{\partial c}{\partial x} \quad (1)$$

Waarbij  $u$  negatieve of positieve waarden kan aannemen naar gelang de richting van het snelheidsveld, en  $c$  en  $t$  positief zijn gedefinieerd ( $\{u \leq 0, u \geq 0\}, c \geq 0, t \geq 0$ ).

Een discrete variant van deze vergelijking op basis van een derde orde upwind-schema in een eindige differentieformulering is

$$c_i^{n+1} = c_i^n - \Delta t [u^+ c_x^- + u^- c_x^+] \quad (2)$$

Hierin is

$$\begin{aligned} u^+ &= \max(u, 0) \\ u^- &= \min(u, 0) \\ c_x^- &= \frac{2c_{i+1} + 3c_i - 6c_{i-1} + c_{i-2}}{6\Delta x} \\ c_x^+ &= \frac{-c_{i+2} + 6c_{i+1} - 3c_i - 2c_{i-1}}{6\Delta x} \end{aligned} \quad (3)$$

Voor de volledige set punten  $p_i$  kan een stelsel van vergelijkingen worden opgesteld:

$$c^{t+\Delta t} = Bc^t \quad (4)$$

Let op dat vergelijking (4) slechts geldt voor de interne knooppunten van een reeks  $p_i$  met oplopend knooppuntnummer. Deze zijn slechts met twee andere punten zijn verbonden. Bij een groter aantal verbindingen, complexere geometrie en/of onregelmatige nummering moeten geschikte varianten van vergelijking (4) worden toegepast.

Wanneer we uitgaan van een concentratieveld op een bepaald tijdstip, zoals bijvoorbeeld (lokaal) vastgesteld door een beperkt aantal sensoren, kan door vergelijking (3) te inverteren het concentratieveld

voor een stapje  $\Delta t$  terug in de tijd worden bepaald, rekening houdend met het transport van de stof. Dit proces kan in principe vele malen worden herhaald. Een klassieke inversie volgens

$$\mathbf{c}^t = \mathbf{B}^{-1} \mathbf{c}^{t+\Delta t} \quad (5)$$

leidt echter meestal tot instabiele oplossingen.

Onder aanname van o.a.:

- een Gaussische verdeling van de werkelijke waarden van de concentratieparameters  $c_i^t$  en  $c_i^{t+\Delta t}$
- een Gaussische verdeling van de werkelijke waarden van het snelheidsveld  $u_i^t$
- een a priori-schatting van de oplossing  $c_{i,a priori}=0$

kan  $c_i^t$  worden teruggerekend uit  $c_i^{t+\Delta t}$  door gebruik te maken van het inversieraamwerk van Tarantola (2005, vergelijking (1.106)) met de volgende uitdrukking:

$$\mathbf{c}^t = \mathbf{C}_M \mathbf{B}^T (\mathbf{B} \mathbf{C}_M \mathbf{B}^T + \mathbf{C}_D)^{-1} \mathbf{r}^{t+\Delta t} \quad (6)$$

Hierin is  $\mathbf{C}_M$  de a priori modelcovariantiematrix en  $\mathbf{C}_D$  de “uitgebreide” data-covariantiematrix. In de formulering van Tarantola (2005) is sprake van een a priori model  $c_0$ , dat in onze benadering als nulvector (geen afwijking) wordt verondersteld en daarom wegvalt uit de vergelijking.

De a priori modelcovariantiematrix  $\mathbf{C}_M$  wordt gevuld door een reeks plausibele modellen voor het systeem te genereren en daarvoor een covariantiematrix te bepalen (zie bijvoorbeeld Muntendam-Bos et al., 2008). Hiermee worden de te verwachten grenzen en waarschijnlijkheden voor de oplossing  $c_i^t$  voorgeschreven. De “uitgebreide” datacovariantiematrix  $\mathbf{C}_D$  is een sommatie van de covariantiematrices voor respectievelijk de “observaties” (in ons geval  $c_i^{t+\Delta t}$ )  $\mathbf{C}_d$  en het transportmodel  $\mathbf{C}_T$  (Tarantola, 2005, vergelijking (1.101)):

$$\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_T \quad (7)$$

De matrix  $\mathbf{C}_d$  bevat de covarianties van het concentratieveld  $c_i^{t+\Delta t}$ , dat wellicht het resultaat van een eerdere berekening vanaf een vorige tijdstap is, waarbij deze matrix zondermeer beschikbaar is. De matrix  $\mathbf{C}_T$  wordt gevuld met de covarianties behorende bij een representatieve verzameling van snelheidsvelden voor het systeem, zoals verkregen uit simulaties, vermenigvuldigd met een factor  $\Delta t/\Delta x$ , zie uitdrukking (4).

De a posteriori covariantiematrix die hoort bij de oplossing van uitdrukking (6) wordt als volgt geformuleerd (Tarantola, 2005):

$$\tilde{\mathbf{C}}_M = \mathbf{C}_M - \mathbf{C}_M \mathbf{G}^T (\mathbf{B} \mathbf{C}_M \mathbf{B}^T + \mathbf{C}_D)^{-1} \mathbf{B} \mathbf{C}_M \quad (8)$$

Deze kan gebruikt worden als  $\mathbf{C}_d$  in vergelijking (7) in een nieuwe iteratie van het terugrekenproces.

*Implementatie*

Om te voorkomen dat er een expliciete inversie van een matrix (duur!) moet worden uitgevoerd, kan vergelijking  $\mathbf{c}_+^t = C_M^+ B^T (B_+ + C_M^+ B_+^T + C_D^+)^{-1} \mathbf{r}_+^{t+\Delta t}$  (29) als volgt worden herschreven:

$$\mathbf{c}^t = C_M B^T (LU)^{-1} \mathbf{r}_+^{t+\Delta t} \quad (9)$$

met

$$LU = (BC_M B^T + C_D) \quad (10)$$

De LU-decompositie is numeriek gezien relatief goedkoop uit te voeren.

Vervolgens wordt (9) omgeschreven:

$$LU(C_M B^T)^{-1} \mathbf{c}^t = \mathbf{r}_+^{t+\Delta t} \quad (11)$$

oftewel

$$LU\mathbf{x} = \mathbf{b} \quad (12)$$

met

$$\begin{aligned} \mathbf{r}_+^{t+\Delta t} &= (C_M B^T)^{-1} \mathbf{r}_+^{t+\Delta t} \\ \mathbf{b} &= \mathbf{r}_+^{t+\Delta t} \end{aligned} \quad (13)$$

Met  $Ux=y$  wordt (12)  $Ly=b$ , waaruit, aanzien  $L$  een onderdriehoeksmatrix is,  $y$  gemakkelijk op te lossen is. Vervolgens kan op vergelijkbare wijze  $x$  opgelost worden uit  $Ux=y$ . Uiteindelijk verkrijgen we:

$$\mathbf{c}^t = (C_M B^T)^{-1} \mathbf{r}_+^{t+\Delta t} \quad (14)$$

Dezelfde benadering wordt gekozen op de a posteriori covariantiematrix  $C_p$  te verkrijgen, uitgaande van de reeds uitgevoerde LU-decompositie van uitdrukking (10).

$$C_p = (I - R)C_M \quad (15)$$

$$\begin{aligned} R &= C_M B^T (BC_M B^T + C_D)^{-1} B \\ &= C_M B^T (LU)^{-1} B \end{aligned} \quad (16)$$

$$LU(C_M B^T)^{-1} R = B \quad (17)$$

Met

$$X = (C_M B^T)^{-1} R \quad (18)$$

kan opnieuw  $Y=UX$  (let op:  $X$  en  $Y$  zijn nu matrices) worden opgelost uit  $LY=B$ , waarna  $X$  wordt opgelost uit  $UX=Y$  en de resolutiematrix  $R$  wordt bepaald:

$$R = (C_M B^T)^{-1} X \quad (19)$$

en vervolgens uit (15) de a posteriori covariantiematrix.

## Oriëntatie van gradiënt in inversie

Omdat we door het uitvoeren van de inversie in feite de richting van het stromingsveld omdraaien, is het voor de stabiliteit van het probleem beter om ook de oriëntatie van de discrete gradiëntoperator in het upwind-schema om te draaien, waardoor vergelijking (2) verandert in

$$c_i^{n+1} = c_i^n - \Delta t [u^+ c_x^+ + u^- c_x^-] \quad (20)$$

### Verder uitbreiding

De hieronder beschreven uitbreiding is mathematisch uitgewerkt maar nog niet geïmplementeerd en getest in een toepassing.

Indien  $u$  wordt gemiddeld over een bepaalde tijdsperiode, dan kan het voorkomen dat het gemiddelde de waarde nul of bijna nul aanneemt. Het stoftransport wordt dan gemiddeld ook (bijna) nul. Gegeven een maximale waarde voor de snelheid  $u$  en het gemiddelde stoftransport is (bijna) nul, dan zal de variantie van het stoftransport groter zijn dan wanneer het gemiddelde snelheidsveld over alleen positieve ( $u^+$ ) of alleen negatieve ( $u^-$ ) snelheidswaarden is berekend. De variantieberekening in Tarantola (2005) gaat echter uit van een lineaire benadering tussen data (hier  $c^t$ ) en modelparameters (hier  $u$ ), zie bijvoorbeeld vergelijking 1.105 in Tarantola (2005). We hebben hier echter te maken met een bilineaire term (product van lineaire termen, hier  $u$  en  $c$ ), wat doorwerkt in de oplossing van het inversieprobleem en de bijbehorende covariantieberekening:

$$\mathbf{r}^t = C_M \mathbf{B}^T (\mathbf{B} C_M \mathbf{B}^T + C_D)^{-1} \mathbf{r}^{t+\Delta t} \quad (21)$$

$$C_D = C_d + C_T \quad (22)$$

$$\tilde{C}_M = C_M - C_M \mathbf{G}^T (\mathbf{B} C_M \mathbf{B}^T + C_D)^{-1} \mathbf{B} C_M \quad (23)$$

$$C_d \leftarrow \tilde{C}_M \quad (24)$$

Met  $C_D$  de covariantiematrix van het gehele systeem,  $C_d$  de covariantiematrix van het geobserveerde concentratieveld  $c_i^{t+\Delta t}$ , en  $C_T$  de covariantie behorende bij het transportmodel.

Uit vergelijking (21)-(23) volgt dat de variantie terug in de tijd niet verandert indien  $C_T = 0$  of indien  $B=0$ . De covariantie vanuit het transportmodel zal niet gelijk zijn aan nul. Echter, de bilineariteit in  $u$  en  $c$  zorgt ervoor dat er bij gemiddelde snelheid nul (of dichtbij) nul de covariantie niet (of nauwelijks) toeneemt omdat  $B \approx 0$ , terwijl er wel degelijk grote snelheidsverschillen (dus hogere variantie) ten grondslag kunnen liggen.

Het singuliere probleem kan omzeild worden door het opknippen van bovenstaand transportprobleem in twee verschillende soorten realisaties van het model (een zgn. *mixture model*): enerzijds een situatie die correspondeert met tijdsperioden waarbij er positieve snelheidsvelden zijn gemeten of gesimuleerd (aangegeven met  $\Delta_i t^+$ ,  $i=1 \dots N$ ) en anderzijds de tijdsperioden waarbij het snelheidsveld precies de andere kant op wijst ( $\Delta_i t^-$ ,  $i=1 \dots M$ ). Merk verder op dat  $\Delta_i t^+$  en  $\Delta_i t^-$  verwijzen naar historische (simulatie)gegevens, de  $\Delta t$  in (2) heeft betrekking op de (fictieve) tijd in de inversieberekeningen.

Zonder verlies van algemene geldigheid in vergelijking (3), gaan we ervan uit dat

$$u^+ = \frac{1}{N^+} \sum_i^{N^+} u_i^+; u^+ > 0$$

$$u^- = \frac{1}{N^-} \sum_i^{N^-} u_i^-; u^- < 0$$
(25)

Vergelijking (21) heeft direct implicaties voor de uitkomst van vergelijking (2), omdat er namelijk twee concentratieprofielen kunnen worden uitgerekend, afhankelijk van welk snelheidsveld als uitgangspunt is genomen.

In matrix-vorm, leidt vergelijking voor een derdegraads upwind schema (3) tot

$$V_{c_+}^{t+\Delta t} = \begin{pmatrix} b_{1,1} & b_{1,2} & 0 & L & L & L & 0 \\ -6 & 3 & 2 & 0 & 0 & 0 & M \\ 1 & -6 & 3 & 2 & 0 & 0 & M \\ 0 & 0 & 0 & 0 & 0 & 0 & M \\ M & 0 & 1 & -6 & 3 & 2 & 0 \\ 0 & L & 0 & 1 & -6 & 3 & 2 \end{pmatrix} c^t + \begin{pmatrix} \beta \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} c_0^t$$
(26)

waar  $c_0$  de concentratie is van het punt waar vandaan wordt gerekend ( $i=0$ ), en  $\beta$  en  $b$  specifieke waarden aannemen naar gelang de randvoorwaarde. Vergelijking (6) kan, indien er vanuit één tijdstip (m.a.w. gesteld dat er geen tijdsvariërende randvoorwaarde is) met bijbehorende specifieke randvoorwaarde wordt gekeken ( $\beta=0$ ), verder beknopt worden weergegeven als:

$$V_{c_+}^{t+\Delta t} = B^+ r^t \quad (27)$$

en

$$V_{c_-}^{t+\Delta t} = B^- r^t \quad (28)$$

Onder de in het eerste deel van deze bijlage genoemde aannames kan  $c_i^t$  worden teruggerekend uit  $c_i^{t+\Delta t}$  door gebruik te maken van het inversieraamwerk van Tarantola (2005, vergelijking (1.106)) met de volgende uitdrukkingen:

$$V_{c_+}^t = C_M^+ B_+^T (B_+ + C_M^+ B_+^T + C_D^+)^{-1} r^{t+\Delta t} \quad (29)$$

en

$$r_{c_-}^t = C_M^+ B_-^T (B_- C_M^+ B_-^T + C_D^-)^{-1} r^{t+\Delta t}$$

Waarbij er kan worden aangenomen dat een schatting van de concentratie  $c^t$ , gegeven de data over  $\Delta t$  en  $\Delta t^+$ , overeenkomt met de verwachtingswaarde van de stochast  $c$  (die een uitkomst is van een trekking uit de kansdichtsverdeling behorende aan  $c_-^t$  en een verdeling behorende bij  $c_+^t$ ). De totale kansdichtsverdeling voor  $X$  zal in dat geval een bimodale verdeling aannemen, met



kansdichtheidsfunctie:

$$f(c) = \lambda g_-(c) + (1 - \lambda) g_+(c)$$

en een mix-parameter met waarde  $\lambda=0.5$  indien beide kansverdelingen even zwaar meetellen in de bimodale kansverdeling. Het eerste moment van  $f(x)$  (de verwachtingswaarde van de stochast, ofwel het gemiddelde) is gedefinieerd als:

$$E(c) = \mu = \lambda \mu_- + (1 - \lambda) \mu_+ \quad \text{met} \quad \mu = \int cf(c)dc$$

Met bovenstaand instrumentarium kunnen we spreken van twee verwachtingswaarden die zullen corresponderen met  $c_+$  en  $c_-$ . Met andere woorden, gegeven de notatie  $\langle c \rangle$  als gemiddelde van  $c$  verkrijgen we:

$$\langle c \rangle = E(c) = \lambda \langle c_- \rangle + (1 - \lambda) \langle c_+ \rangle$$

De totale variantie ( $\sigma^2$ ) behorende bij vergelijking (2) kan worden gereconstrueerd (Kim en White, 2003) met behulp van de a posteriori covariaantiematrixes  $\tilde{C}_M^-$  en  $\tilde{C}_M^+$ :

$$\sigma^2 = \lambda \left( \left( \tilde{C}_M^- \right)^2 + \delta_-^2 \right) + (1 - \lambda) \left( \left( \tilde{C}_M^+ \right)^2 + \delta_+^2 \right)$$

met

$$\delta_- = \mu_- - \mu \quad \text{en} \quad \delta_+ = \mu_+ - \mu$$



# IV Dekkingskaarten

## Definities

$S_{ij}$  is de verzameling netwerkknoten dievoor snelheidsveldrealisatie  $j$  worden teruggevonden als potentiële bronknoten voor een verontreiniging die wordt gedetecteerd op tijdstip  $t$  bij sensorlocatie  $i$ .

## Formuleringen

	<i>kaart</i>	<i>bepaling</i>
1)	gezichtsveld van een sensor op tijdstip $t$	$S_{i1}$
2)	gezichtsveld van een sensor op tijdstip $t$ met $n$ realisaties van het stochastische snelheidsveld	$S_{i1} \cap \dots \cap S_{in}$
3)	gezichtsveld van $m$ sensoren op tijdstip $t$ , samen dekkend	$S_{11} \cup \dots \cup S_{m1}$
4)	gezichtsveld van $m$ sensoren op tijdstip $t$ , alle dekkend	$S_{11} \cap \dots \cap S_{m1}$
5)	gezichtsveld van $m$ sensoren op tijdstip $t$ met $n$ realisaties van het stochastische snelheidsveld, samen dekkend	$(S_{11} \cup \dots \cup S_{m1}) \cap \dots \cap (S_{1n} \cup \dots \cup S_{mn})$
6)	gezichtsveld van $m$ sensoren op tijdstip $t$ met $n$ realisaties van het stochastische snelheidsveld, alle dekkend	$(S_{11} \cap \dots \cap S_{m1}) \cap \dots \cap (S_{1n} \cap \dots \cap S_{mn})$
7)	gezichtsveld van een sensor gedurende een periode van 24 uur, gedurende enige tijd in beeld	$(S_{i1})_{t=1} \cup \dots \cup (S_{i1})_{t=24}$
8)	gezichtsveld van een sensor gedurende een periode van 24 uur, gedurende volledige periode in beeld	$(S_{i1})_{t=1} \cap \dots \cap (S_{i1})_{t=24}$
9)	gezichtsveld van een sensor gedurende een periode van 24 uur met $n$ realisaties van het stochastische snelheidsveld, gedurende enige tijd in beeld	$(S_{i1} \cap \dots \cap S_{in})_{t=1} \cup \dots \cup (S_{i1} \cap \dots \cap S_{in})_{t=24}$
10)	gezichtsveld van een sensor gedurende een periode van 24 uur met $n$ realisaties van het stochastische snelheidsveld, gedurende volledige periode in beeld	$(S_{i1} \cap \dots \cap S_{in})_{t=1} \cap \dots \cap (S_{i1} \cap \dots \cap S_{in})_{t=24}$
11)	gezichtsveld van $m$ sensoren gedurende een periode van 24 uur met $n$ realisaties van het stochastische snelheidsveld, gedurende enige tijd in beeld	$((S_{11} \cup \dots \cup S_{m1}) \cap \dots \cap (S_{1n} \cup \dots \cup S_{mn}))_{t=1} \cup \dots \cup ((S_{11} \cup \dots \cup S_{m1}) \cap \dots \cap (S_{1n} \cup \dots \cup S_{mn}))_{t=24}$
12)	gezichtsveld van $m$ sensoren gedurende een periode van 24 uur met $n$ realisaties van het stochastische snelheidsveld, gedurende volledige periode in beeld	$((S_{11} \cup \dots \cup S_{m1}) \cap \dots \cap (S_{1n} \cup \dots \cup S_{mn}))_{t=1} \cap \dots \cap ((S_{11} \cup \dots \cup S_{m1}) \cap \dots \cap (S_{1n} \cup \dots \cup S_{mn}))_{t=24}$



# V Gebruikte software

## EPANET

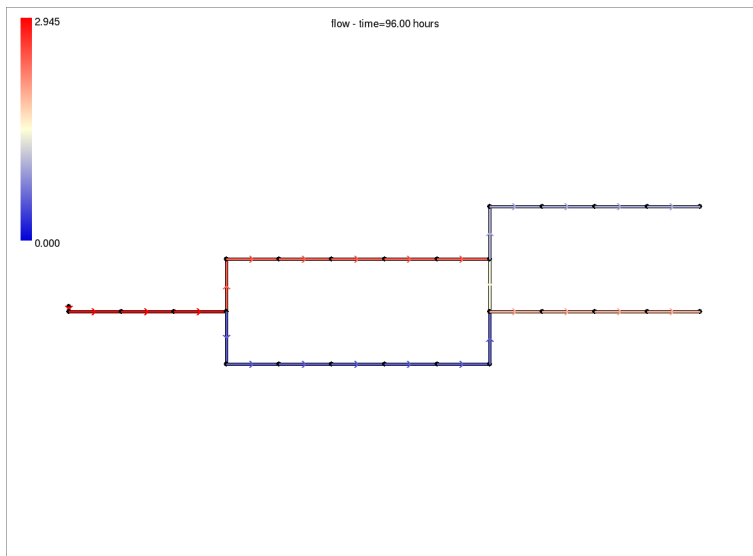
- EPANET 2.0.12 (linux 64 bit gecompileerd)
- EPANET-MSX 1.1 (linux 64 bit gecompileerd)

## Externe Python-bibliotheken

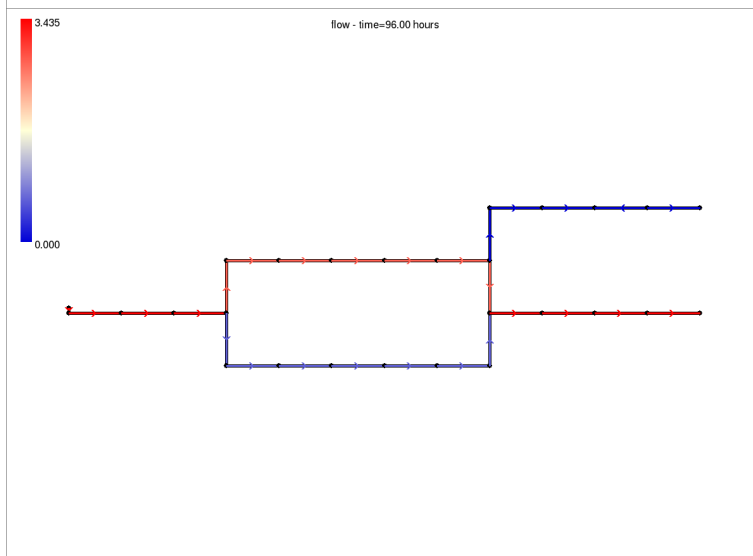
- pygame (<http://www.pygame.org>)
- numpy (<http://www.numpy.org/>)
- scipy (<http://www.scipy.org/>)
- matplotlib (<http://matplotlib.org/>)
- wxpython (<http://www.wxpython.org/>)



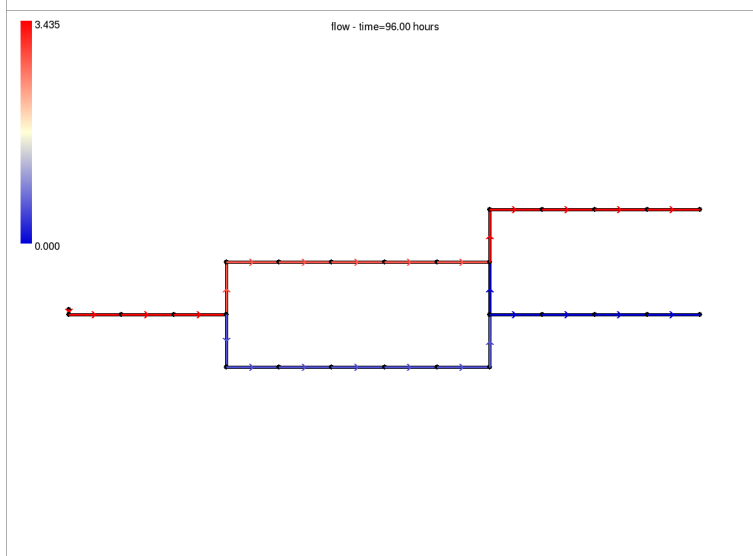
# VI Test van implementatie



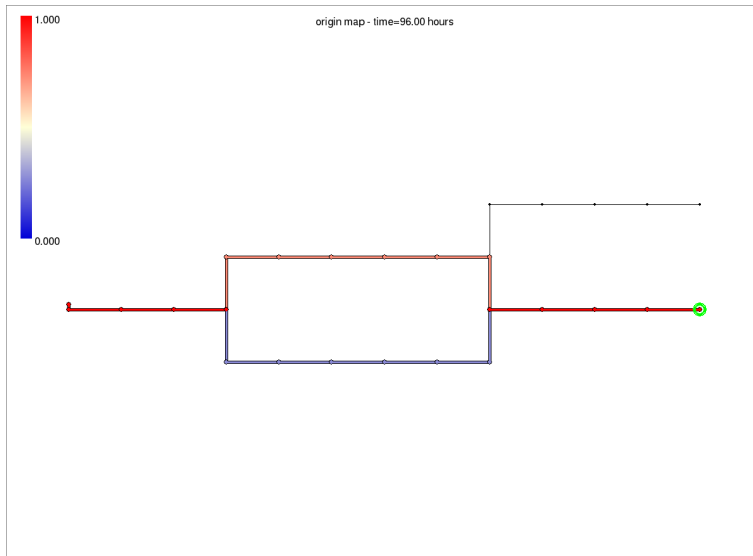
Realisatie 1 van het snelheidsveld.



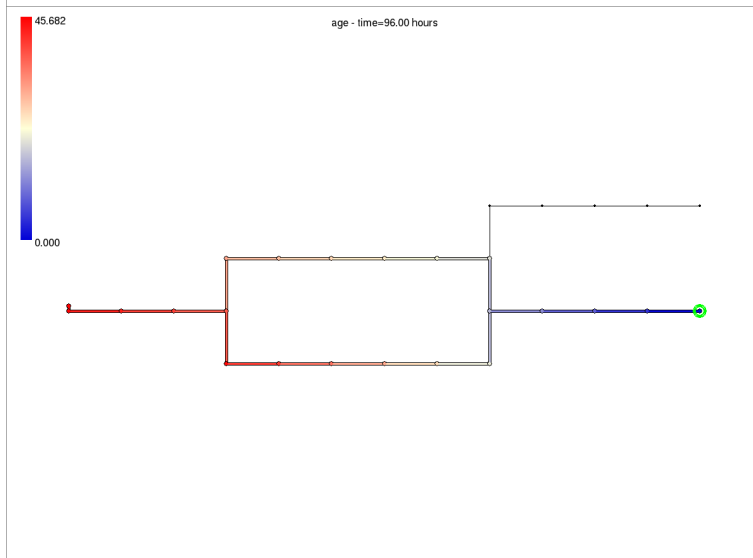
Realisatie 2 van het snelheidsveld.



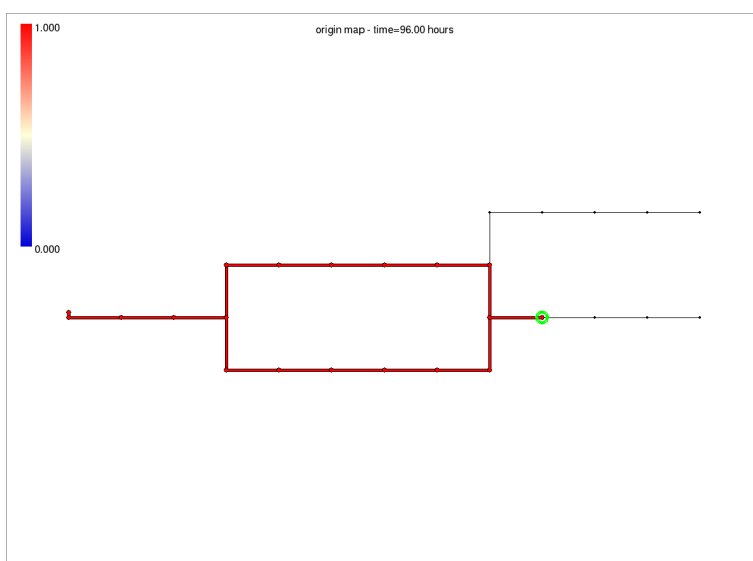
Realisatie 3 van het snelheidsveld.



Herkomstkaart voor realisatie 1 voor het groen gemarkeerde knooppunt.

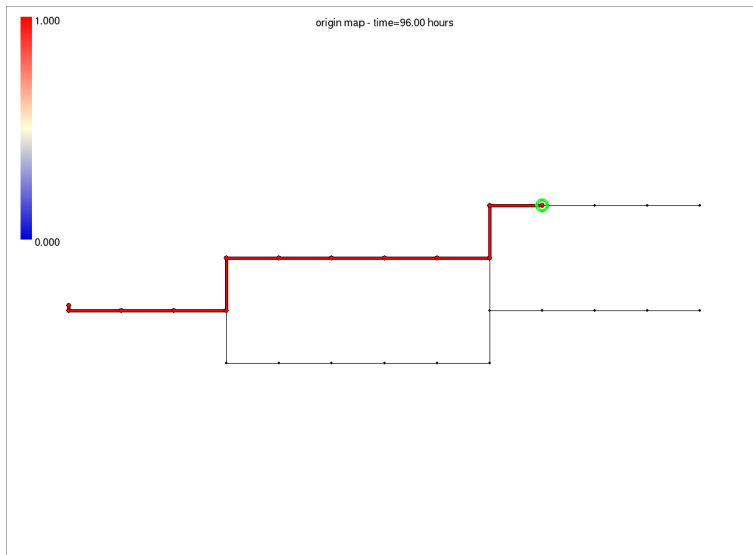


Bijbehorende waterleeftijdskaart voor het groen gemarkeerde knooppunt. Let op dat de weergegeven leeftijd in het eerste gedeelte van het traject het naar de volumestroom van beide takken gewogen gemiddelde van de leeftijden voor de afzonderlijke paden is.

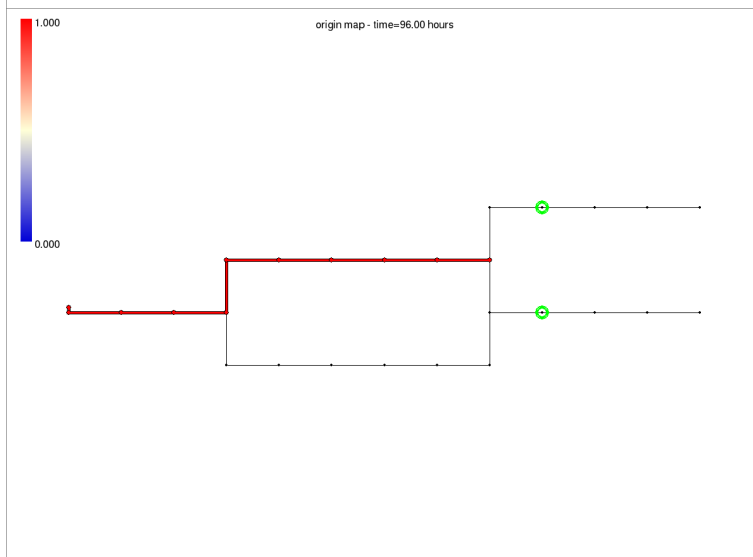


Herkomstkaart voor realisatie 1 voor het groen gemarkeerde knooppunt.

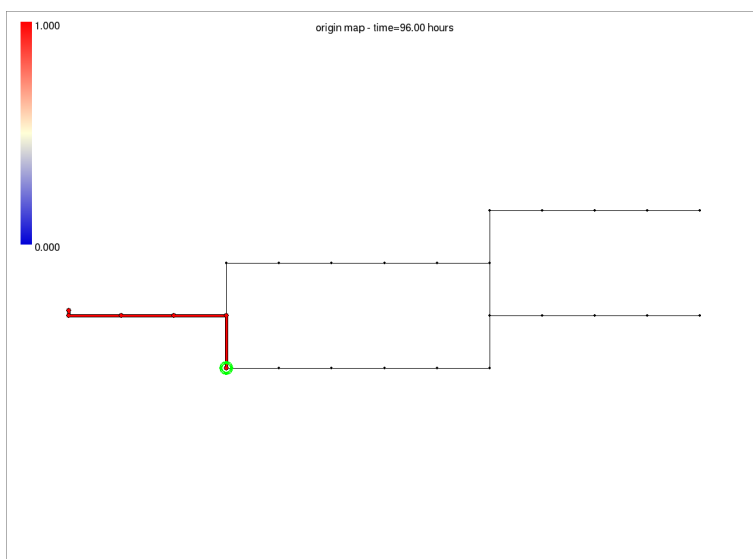




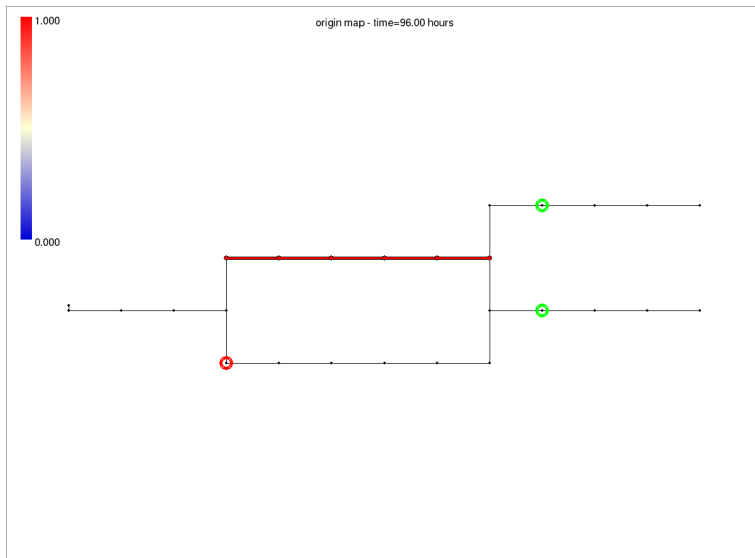
Herkomstkaart voor realisatie 1 voor het groen gemarkeerde knooppunt.



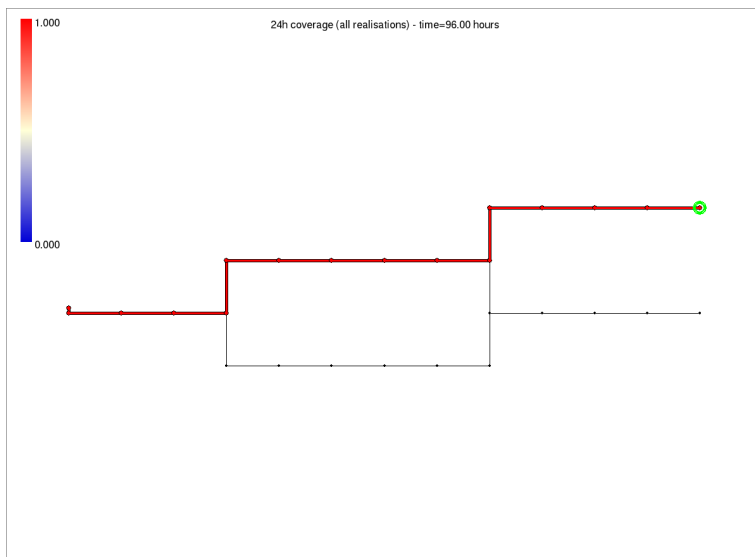
Gecombineerde herkomstkaart voor realisatie 1 voor de groen gemarkeerde knooppunten.



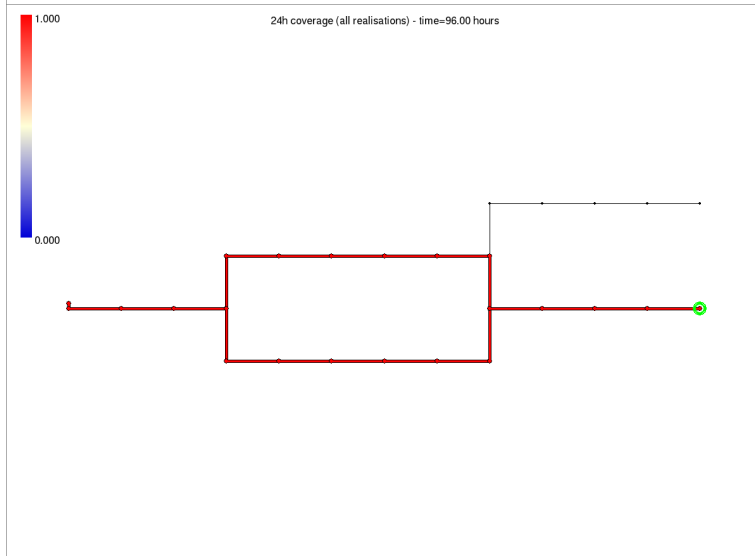
Herkomstkaart voor realisatie 1 voor het groen gemarkeerde knooppunt.



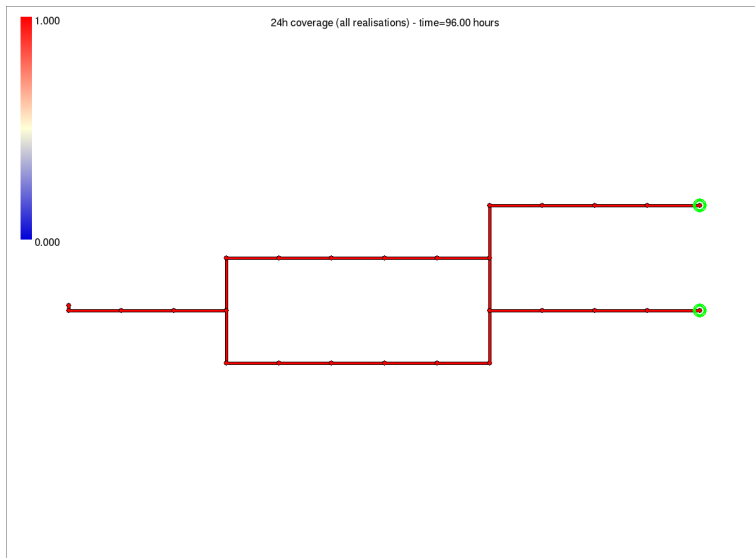
Gecombineerde herkomstkaart voor realisatie 1 voor een positieve sensormeting op de groen gemarkeerde knooppunten en een negatieve op het rood gemarkeerde punt.



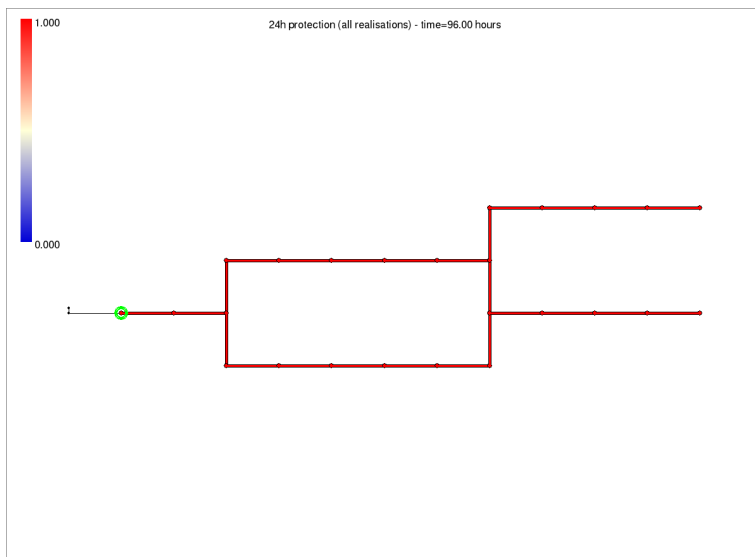
Dekking van groen gemarkeerde sensor.



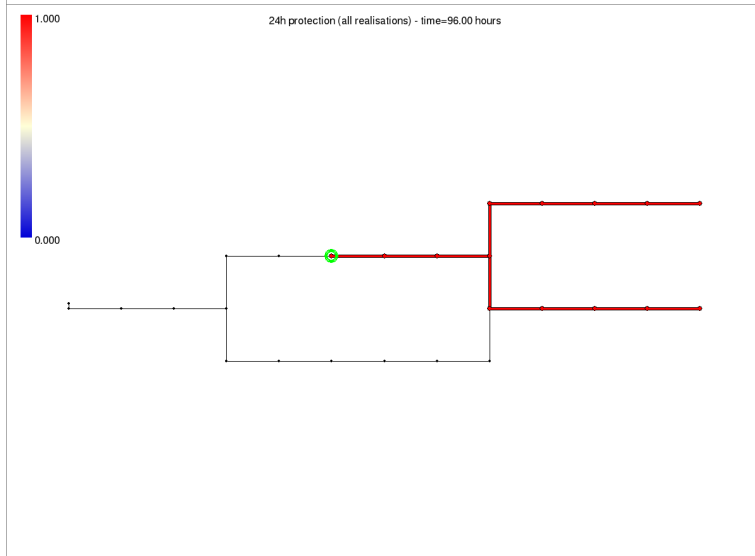
Dekking van groen gemarkeerde sensor.



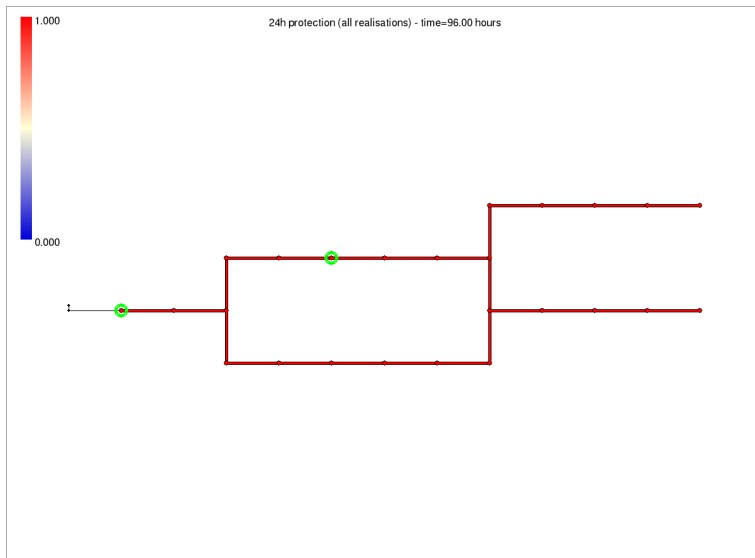
Dekking van groen gemarkeerde sensoren.



Beschermingsgebied van groen gemerkte sensor.



Beschermingsgebied van groen gemerkte sensor.



Gezamenlijk beschermingsgebied van groen gemerkte sensoren.

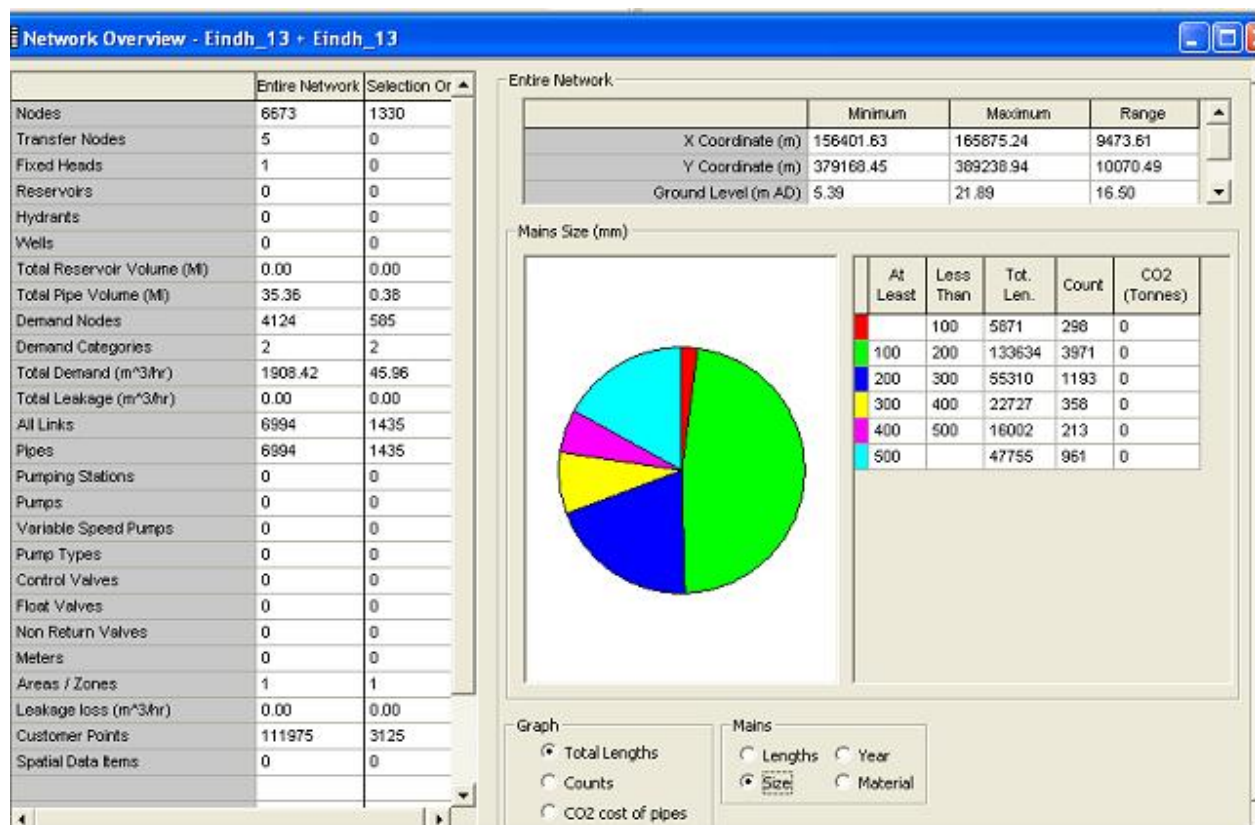
# VII Voorbereiding van de hydraulische modellen van Eindhoven en Almere

## Eindhoven

De volgende bewerkingen zijn uitgevoerd op het oorspronkelijk verkregen model van Eindhoven:

- x-coördinaat van afsluiters aangepast; van -0.1m ; naar +0.1m
- leidingen van Geldrop, Aalst, Waalre en Valkenswaard verwijderd
- Verdund behalve geselecteerde wijk
- 3 verbruikspunten Urkhovensweg toegevoegd
- verbruik vanuit alle CP opnieuw toegewezen
- dubbele leidingen (suffix=2) verwijderd
- afsluiters gewijzigd in leidingen
- hydranten omgezet naar knopen
- Verdund met behoud van uitlopers met diam > 230 mm en samenvoegen van leidingen met diam < 500 mm
- Verdunnen herhaald met diverse k-waarden en materiaalwaarden voor leidingen met lengte <1 m
- verbruik opnieuw toegewezen; deel van toewijzing handmatig gecorrigeerd
- reset van total connections
- run gecontroleerd op druk >20 m+mv op gemdag
- drie leidingen toegevoegd aanlegjaar 2013 ivm verwijderd bij verdunnen
- toewijzing van SIMDEUM-verbruiken aan knopen in onverdunde wijk

Onderstaande figuur geeft een overzicht van de karakteristieken van het uiteindelijke netwerkmodel.

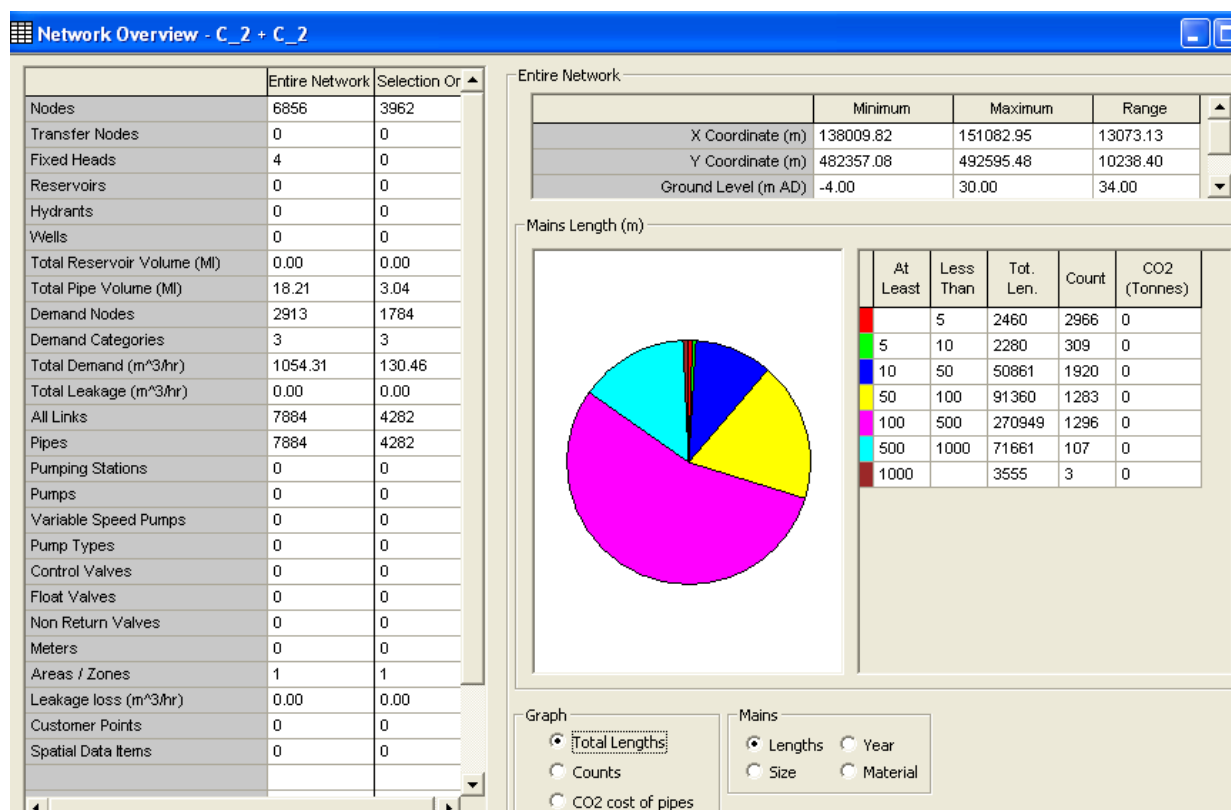


## Almere

De volgende bewerkingen zijn uitgevoerd op het oorspronkelijk verkregen model van Almere:

- dubbele leidingen (suffix=2) verwijderd
- lengte leiding < 0,1 m wordt 0,1 m
- Verdund met behoud van uitlopers met diam > 230 mm en samenvoegen van leidingen met diam < 500 mm
- lengte met local loss=1,5 wordt local loss=0
- diameter <59 mm gewijzigd in 59 mm in gebied buiten geselecteerde wijk
- diameter <97,4 mm gewijzigd in 103 mm in gebied buiten geselecteerde wijk
- Opnieuw verdund met behoud van uitlopers met diam > 230 mm en samenvoegen van leidingen met diam < 500 mm
- Diameter van leidingen met drukgradient >1 m/km gecontroleerd en daarvan 4 diameters gewijzigd in diameter van naburige leiding
- toewijzing van SIMDEUM-verbruiken aan knopen in onverdunde wijk

Onderstaande figuur geeft een overzicht van de karakteristieken van het uiteindelijke netwerkmodel.



## Tijdstappen en patronen

Bij de berekening is uitgegaan van de volgende tijdstappen:

hydraulische tijdstap	5 minuten
waterkwaliteitstijdstap	5 minuten
vraagpatroontijdstap	5 minuten

Vraagpatronen zijn gedefinieerd voor een periode van 24 uur. In de berekeningen van 72 uur zijn deze dus 3 keer herhaald.



