

Optimaal gebruik van Citizen Science data: wat (niet) te doen?

Arjan Droste & Bas Walraven



Waar moet je op letten: Doelstelling

- Wat is je doel? Verschillende soorten data beantwoorden verschillende vragen. Is het jouw vraag, of is het de vraag van de burgers zelf?
 - Onderzoek, beleidsontwikkeling of inventaris, monitoring, participatie
 - Onderzoek: wetenschappelijk inzicht in een bepaald fenomeen, vaak op ingezoomde mate waardoor ‘traditionele’ metingen ontoereikend zijn. Baat bij intensievere of langdurige CS metingen van gecontroleerde kwaliteit
 - Beleid // inventaris: inventaris maken van een bepaald probleem of ontwikkeling (ten behoeve van toekomstig beleid). Lokaal, met participatie van lokale kennis – kan minder kwantitatief afhankelijk van de vraagstelling
 - Monitoring: Monitoring van een fenomeen of probleem (bijv grondwaterstand) niet ten behoeve van onderzoek, maar om een vinger aan de pols te houden bij bijv. overschrijdingen. Langdurig netwerk, mogelijk meer uitgespreid, passieve metingen (geautomatiseerd).
 - Participatie: betrokkenheid van burgers bij een project van wat voor aard dan ook (een van de bovenstaande of los daarvan). Metingen hebben als doel om terugkoppeling te veroorzaken naar de burgers en vooral hen te informeren over de stavaza in hun buurt of wijk, met slechts als secundair doel om de organiserende partij te informeren.

Doelstelling van CS data

<p>Onderzoek</p> <p>Vaak langdurig, gebaat bij veel metingen, op gebieden waar traditionele methoden ontoereikend zijn</p> <p>Kwaliteit van metingen is belangrijk</p>	<p>Beleid of Inventaris</p> <p>CS data ter behoeve van (toekomstig) beleid</p> <p>Lokaal, belang van plaatselijke kennis en ervaringen</p> <p>Kan meer kwalitatief, probleem in kaart brengen</p>
<p>Monitoring</p> <p>‘Vinger aan de pols’: monitoring van een bepaald fenomeen om eventueel in te grijpen</p> <p>Baat bij langdurige, automatische metingen</p>	<p>Participatie of raadplegen</p> <p>Doel is niet zozeer de data, maar de deelname van de burger</p> <p>De burger gaat eerder aan de slag met de data!</p> <p>Betrokkenheid bij besluiten of vernieuwing, inspraak in beslissingen</p>

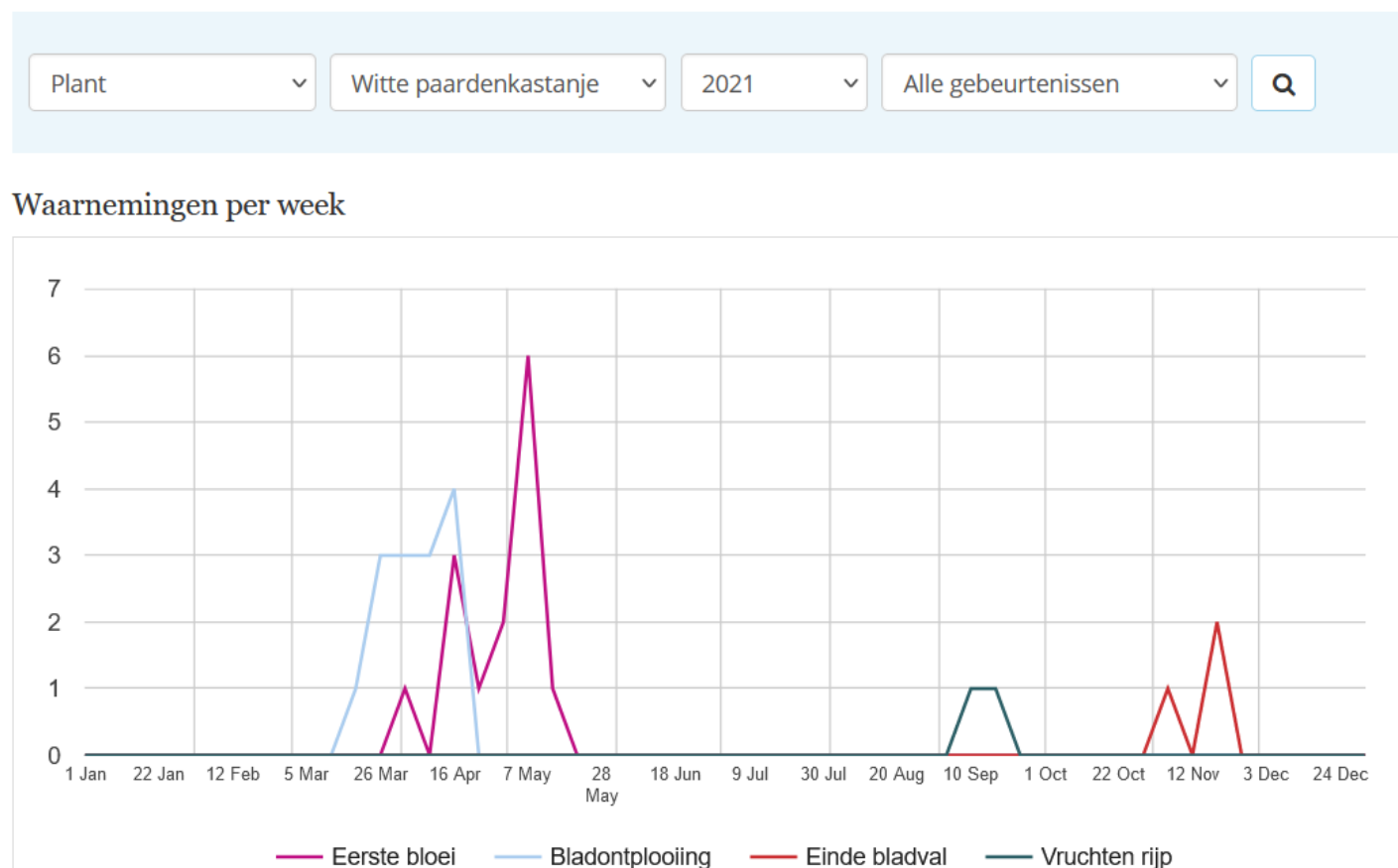
Voorbeeld: Delft Meet Regen

- Personal Weather Stations
- Langdurig project, doorlopende tijdsreeksen
- Kwantitatief: reeks neerslagwaarden
- Representatief?
- Opstelling?
- Hoe wordt data doorgegeven?



Voorbeeld: Natuurkalender

- Fenologisch waarnemingsprogramma
- Deelnemers geven moment van bloei, start vogeltrek, etc door
- Langdurig project (sinds 2001), geen tijdsreeks – momentopnames.
- (meer) Kwalitatief
- Individuele datapunten hebben meer gewicht
- Handmatige data aanvoer
- Misclassificaties?



Data types



A) Handmatige ingevuld,
kwalitatieve data

Campagnes en/of
onderbroken tijdreeksen



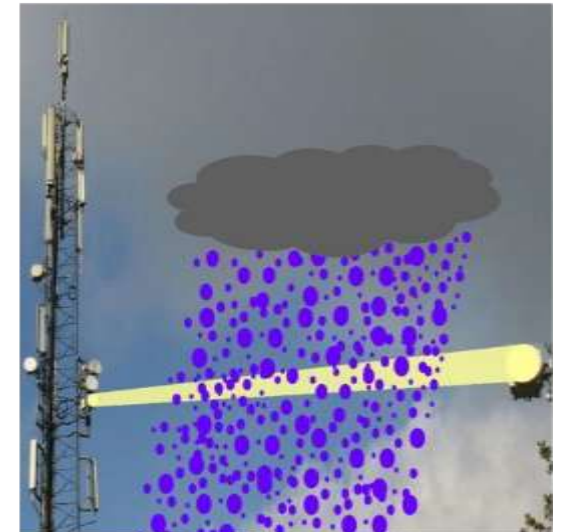
B) Handmatige ingevuld,
numerieke data

Campagnes en/of
onderbroken tijdreeksen



C) Automatisch gelogd,
numerieke data
(‘crowdsourced’)

Continue tijdreeksen



D) Afgeleide
‘opportunistic’ data;
numerieke data
(‘crowdsourced’)

Continue tijdreeksen

Data types



A) Handmatige ingevuld, kwalitatieve data

Campagnes en/of onderbroken tijdreeksen



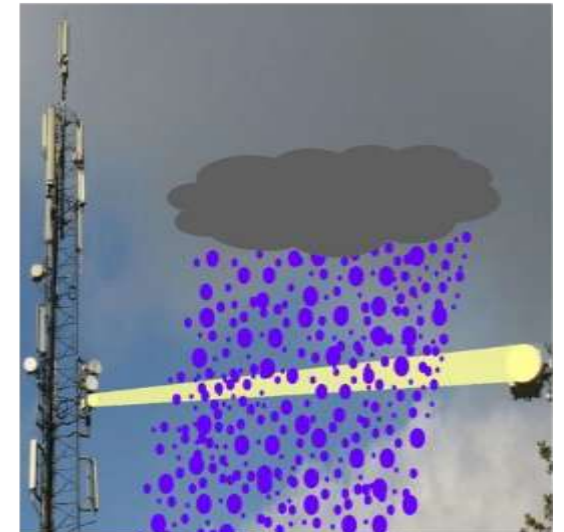
B) Handmatige ingevuld, numerieke data

Campagnes en/of onderbroken tijdreeksen



C) Automatisch gelogd, numerieke data ('crowdsourced')

Continue tijdreeksen



D) Afgeleide 'opportunistic' data; numerieke data ('crowdsourced')

Continue tijdreeksen

Foutbronnen en mogelijke oplossingen

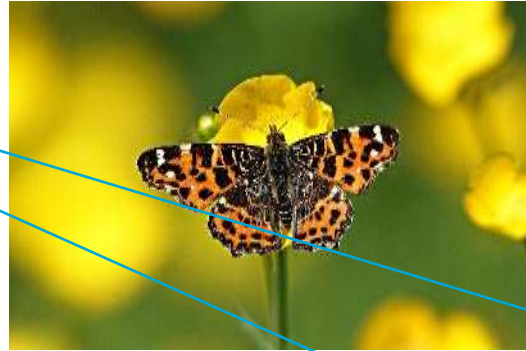
- Misclassificatie



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

- Misclassificatie



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

- Misclassificatie
- Meetopstelling



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

- Misclassificatie
- Meetopstelling



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

- Misclassificatie
- Meetopstelling
- Sensor drift



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

- Misclassificatie
- Meetopstelling
- Sensor drift



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

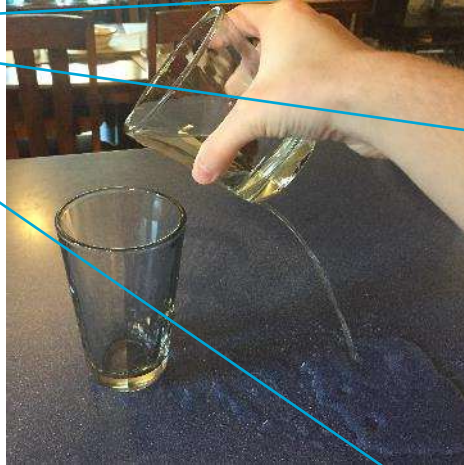
- Misclassificatie
- Meetopstelling
- Sensor drift
- Meetfout & ontbrekende data



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

- Misclassificatie
- Meetopstelling
- Sensor drift
- Meetfout & ontbrekende data



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

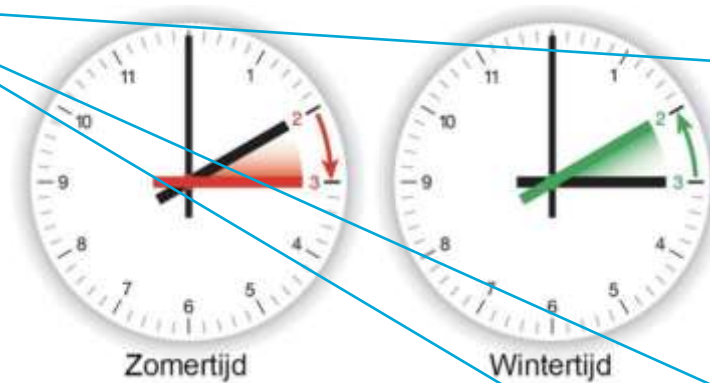
- Misclassificatie
- Meetopstelling
- Sensor drift
- Meetfout & ontbrekende data
- Tijd en/of locatie verschillen



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

- Misclassificatie
- Meetopstelling
- Sensor drift
- Meetfout & ontbrekende data
- Tijd en/of locatie verschillen



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

- Misclassificatie
- Meetopstelling
- Sensor drift
- Meetfout & ontbrekende data
- Tijd en/of locatie verschillen
- Subjectiviteit



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

- Misclassificatie
- Meetopstelling
- Sensor drift
- Meetfout & ontbrekende data
- Tijd en/of locatie verschillen
- Subjectiviteit



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

- Misclassificatie
- Meetopstelling
- Sensor drift
- Meetfout & ontbrekende data
- Tijd en/of locatie verschillen
- Subjectiviteit
- Representativiteit



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Foutbronnen en mogelijke oplossingen

- Misclassificatie
- Meetopstelling
- Sensor drift
- Meetfout & ontbrekende data
- Tijd en/of locatie verschillen
- Subjectiviteit
- Representativiteit



- 1. Data filteren op basis van één vaste waarde
- 2. Data corrigeren met officiële waarnemingen
- 3. Data corrigeren met omliggende CS bronnen
- 4. Data a priori filteren op basis van een classificatie
- 5. Metadata gebruiken om te filteren
- 6. Data corrigeren o.b.v. een trend in de data zelf

Wat moet je wel doen met CS data, en wat niet?

1. Beleid puur baseren op CS data?
 2. Blindelings vertrouwen op de CS data?
 3. Blindelings vertrouwen op de officiële data?
 4. Start je eigen CS campagne?
 5. Delen en publiceren van de data?
 6. Maakt het uit wie je metingen doet?
1. **Neem CS data en bevindingen mee, maar vertrouw ook op in-house kennis en meetnetwerken**
 2. **Nee, maar let op! Ga er ook niet vanuit dat de CS data foutief is. Het is foutengevoelig, niet foutief.**
 3. **Officiële data heeft ook een foutenmarge. Combineer zo veel mogelijk databronnen**
 4. **Maak gebruik van wat er al ligt. Het opbouwen en onderhouden van een platform is duur**
 5. **Data kunnen privacy-gevoelig zijn, zeker beeldmateriaal, maar ook coördinaten bijvoorbeeld.**
 6. **Moeten deelnemers een bepaalde skill hebben? Dan maakt het uit voor de data kwaliteit.**

Checklist

1. Wat is je doel? Zijn er al data of initiatieven om te raadplegen?
2. Wat voor data heb je of ga je krijgen? En wat voor gevoeligheden zitten daar waarschijnlijk in?
3. Waar komt je data terecht? Een (web)platform, handmatige inzameling, ...
4. Zijn er maatregelen die je vooraf kan nemen? Bijv. instructie of controleren van opstellingen?
5. Wat voor andere (niet-CS) data is er, ter complement?
6. Hoe zien je data er uit? Zijn er fouten die je gelijk opvallen?
7. Bekijk (mits mogelijk) de data statistieken, voor outliers, foutieve nulwaarden, hiaten, etc.
8. Maak gebruik van interactie met de burgerwetenschappers! Viel hen iets op?
9. Vergelijk CS data met elkaar: zijn afzonderlijke meetpunten in overeenstemming waar je dat verwacht?
10. Vergelijk CS data met data van andere netwerken, om fouten op te speuren

Bedankt voor uw aandacht

Voor vragen of opmerkingen: mail
A.M.Droste@tudelft.nl