**A close up of a logo

Description automatically generated**

**Beslutningstræ for grundvandsforekomsters påvirkning af overfladevand**

**Risiko- og tilstandsvurdering notat**

Oliver B. Lund

6. Oktober 2025

Indholdsfortegnelse

[Introduktion 3](#_Toc210643386)

[Projektbaggrund 3](#_Toc210643387)

[Formål og Vision 3](#_Toc210643388)

[Metodisk Tilgang 4](#_Toc210643389)

[Risikovurdering (trin 1-5): 4](#_Toc210643390)

[Tilstandsvurdering (fremtidigt arbejde): 4](#_Toc210643391)

[Denne Rapports Fokus 4](#_Toc210643392)

[Datagrundlag 4](#_Toc210643393)

[Shape-filer 4](#_Toc210643394)

[CSV-filer 4](#_Toc210643395)

[Hvordan laves CSV-filerne? 4](#_Toc210643396)

[Risikovurdering 5](#_Toc210643397)

[Trin 1: Optælling af Grundvandsforekomster 5](#_Toc210643398)

[Trin 2: Grundvandsforekomster med Vandløbskontakt 5](#_Toc210643399)

[Trin 3: V1/V2-lokaliteter med Aktive Forureninger i GVFK med Vandløbskontakt 6](#_Toc210643400)

[Trin 4: Afstandsanalyse til Vandløb 8](#_Toc210643401)

[Trin 5: Tærskel-vurdering og Kategorisering 10](#_Toc210643402)

[1. Generel risikovurdering (500 m universal tærskel) 11](#_Toc210643403)

[2. Stofspecifik risikovurdering med losseplads-override (to-fase tilgang) 11](#_Toc210643404)

[Tærskel-sammenligning (Normal vs. Losseplads-specifik) 12](#_Toc210643405)

[Litteraturbaserede stofgrupper (refined\_compound\_analysis.py) 13](#_Toc210643406)

[Tilstandsvurdering 15](#_Toc210643407)

[Planlagt Metodisk Tilgang 15](#_Toc210643408)

[Kvantitativ Fluxberegning 15](#_Toc210643409)

[Koncentrationsvurdering i Vandløb 15](#_Toc210643410)

[Prioritering og Kvantificering 16](#_Toc210643411)

[Samarbejde med GEUS 16](#_Toc210643412)

[Samlet Overblik 16](#_Toc210643413)

[Konklusion 16](#_Toc210643414)

# Introduktion

## **Projektbaggrund**

Denne rapport dokumenterer **risikovurderingsdelen** af projektet “Beslutningstræ for grundvandsforekomsters påvirkning af overfladevand” - et samarbejde mellem DTU Sustain, GEUS, **SGAV??** og Miljøstyrelsen. Projektet udvikler en systematisk metodik til vurdering af, hvorvidt forurenede lokaliteter i grundvandsforekomster udgør en risiko for påvirkning af overfladevand.

## **Formål**

Projektets overordnede mål er at etablere et **automatiseret beslutningstræ** til risikovurdering og tilstandsvurdering af forureningsforekomster i grundvandsforekomster og deres potentielle påvirkning af vandløb. Metoden skal kunne:

* **Screene alle danske grundvandsforekomster** systematisk for risiko
* **Identificere GVF i risiko samt ”højrisiko” lokaliteter** der kræver nærmere undersøgelse
* **Automatiseres** så den kan håndtere landsdækkende data effektivt
* **Integreres med eksisterende modeller** (DK-model, DK-jord data)

## **Metodisk Tilgang**

Tilgangen følger de etablerede principper for screening af jordforurening mod overfladevand fra Miljøstyrelsen, men anvender en struktureret beslutningstræ-approach med to hovedfaser:

### Fase 1: Risikovurdering (trin 1-5):

1. **Baseline-etablering**: Optælling af det totale antal grundvandsforekomster (GVFK) i Danmark
2. **Kontakt-identifikation**: Identifikation af GVFK med kontakt til vandløbssegmenter
3. **Kildelokalisation**: Identifikation af GVFK der indeholder V1/V2-lokaliteter med aktive forureninger
4. **Afstandsanalyse**: Beregning af afstande fra V1/V2-lokaliteter til vandløbssegmenter inden for samme GVFK
5. **Tærskel-vurdering**: Kategorisering af lokaliteter baseret på afstandstærskler og stofspecifikke spredningsafstande

### Fase 2: Tilstandsvurdering (fremtidigt arbejde):

* Kvantitativ fluxberegning og koncentrationsvurdering i vandløb
* Sammenligning med miljøkvalitetskrav
* Prioritering af indsatsområder

## **Denne Rapports Fokus**

Nærværende metodebeskrivelse dokumenterer **den komplette risikovurdering (trin 1-5)** af beslutningstræet. Dette udgør den systematiske identifikation og karakterisering af alle potentielt problematiske lokaliteter baseret på afstand og forureningstyper, samt antallet og hvilken GVFK der er i risiko.

**Tilstandsvurderingen** - den kvantitative vurdering af faktiske koncentrationer og overskridelser i vandløb - vil blive gennemført som næste projektfase efter finalisering af risikovurderingsmetodikken.

Metoden præsenteret her identificerer **1.740 højrisiko-lokaliteter** gennem stofspecifik vurdering (ud af 40.870 analyserede lokaliteter), som danner grundlag for den kommende tilstandsvurdering og prioritering af miljøindsats.

# Datagrundlag

Risikovurderingen er baseret på følgende datafiler:

## **Shape-filer**

* VP3Genbesøg\_grundvand\_geometri.shp: Grundvandsforekomster (GVFK) - 2.043 unikke forekomster
* Rivers\_gvf\_rev20230825\_kontakt.shp: Vandløbsstrækninger med tilknyttet GVFK og kontaktflag (14.454 segmenter, 7.496 med kontakt)
* V1FLADER.shp: V1-lokaliteter som polygoner (28.717 polygoner, 23.209 unikke lokaliteter)
* V2FLADER.shp: V2-lokaliteter som polygoner (33.040 polygoner, 21.269 unikke lokaliteter)

## **CSV-filer**

Følgende CSV-filer er genereret via V1V2.py scriptet og “Fremgangsmåde til klassifikationer af forurenede grunde.docx” notatet (lavet af: Luc Taliesin Eisenbrückner, september 2024), som behandler og kombinerer lokalitetsdata med grundvandsforekomster:

* v1\_gvfk\_forurening.csv: V1-lokaliteter med GVFK-tilknytning, forureningsdata og brancheoplysninger (84,601 rækker, 23,209 unikke lokaliteter)
* v2\_gvfk\_forurening.csv: V2-lokaliteter med GVFK-tilknytning, forureningsdata og brancheoplysninger (134,636 rækker, 21,269 unikke lokaliteter)

### Hvordan laves CSV-filerne?

* **DK-jord udtræk (27-09-2024):** Fra Danmarks Miljøportal med .shp filer (alle V1 og V2 kortlagte grunde) og .csv filer med lokation, forurening, branche, aktivitet og forureningsstatus
* **Geometri-forbehandling:** ArcGIS Dissolve værktøj anvendt på V1 og V2 .shp filer med Lokalitetsnummer som dissolve field og “create multipart feature” aktiveret, så hver unik lokalitet blev til én multipart feature i stedet for opdelte polygoner
* **Kodeliste-join:** Simpelt join mellem forskellige koder fra DK-jord data

#### Kobling til grundvandsforekomster (ArcGIS spatial analyse)

* **Overlapsanalyse:** Spatial join mellem de grundvandstruende V1 og V2 lokaliteter og .shp fil med 2,043 grundvandsforekomster (VP3)
* **Join-operation:** “One to many” join med “keep all target features” aktiveret og match option sat til “intersect”
* **Resultat:** Hver V1/V2 lokalitet får tilknyttet de grundvandsforekomster, den geografisk overlapper med

#### Databehandlingsproces (V1\_V2.py)

1. Indlæsning af dkjord-View\_Lokaliteter med lokation, forurening, branche og aktivitetsdata
2. Ekspansion af stoffdata (opdeling af Lokalitetensstoffer ved semikolon til separate rækker)
3. Fjernelse af dubletter baseret på alle kolonner
4. Join med de ArcGIS-forbehandlede V1/V2 GVFK-data på lokalitetsnummer
5. Fjernelse af GIS-relaterede kolonner og oprydning af datasæt

#### Overlap mellem datasæt

3,608 lokaliteter findes i både V1 og V2 data.

#### Vigtige kolonner i V1/V2 .csv filerne til risikovurdering

* Lokalitetensbranche: Branche/industri-information
* Lokalitetensaktivitet: Aktivitetstype
* Lokalitetensstoffer: Forureningsstoffer
* Navn: GVFK-tilknytning fra ArcGIS spatial join

Datastrukturen af .csv filerne er således at Lokalitetensbranche og Lokalitetensaktivitet er ”lister” af ”;” separaede branche eller aktivitets værdier, da der kan være mange forskellige af disse assosiceret med samme V1/V2 lokalitet. Dette liste format blev anvendt for at undgå at ”eksplodere” antallet af rækker i datasættet. Dog kan samme V1/V2 lokalitet fremgå op til flere gange grundet kolonnen ”Lokalitetensstoffer”, som ikke er i liste format, men for hvert stof for en lokalitet laves der en ny række.

# **Risikovurdering**

## **Trin 1: Optælling af Grundvandsforekomster**

**Formål**  
Etablere baseline for det totale antal unikke grundvandsforekomster (GVFK) i Danmark, så det kan sammenlignes med filtrerede undersæt.

**Input Data (fra Datagrundlag)**  
• Fil: VP3Genbesøg\_grundvand\_geometri.shp  
• Anvendte kolonner:  
 – Navn: Unik tekst-identifikator for hver grundvandsforekomst

**Proceslogik** (step1\_all\_gvfk.py)

1. Indlæsning – Læser shapefil med geopandas.read\_file()
2. Validering – Kontrollerer eksistensen af kolonnen *Navn*
3. Optælling – Beregner antal unikke værdier med Navn.nunique()

**Output**  
Ingen filer gemt (kun i hukommelsen)  
Returnerer: (GeoDataFrame, antal\_unikke\_GVFK)

**Aktuelle Resultater**

2.043 unikke grundvandsforekomster identificeret

## **Trin 2: Grundvandsforekomster med Vandløbskontakt**

**Formål**  
Identificere grundvandsforekomster, hvor grundvand–overfladevand-interaktion forekommer, hvilket er afgørende for vurdering af forureningsspredning til vandløb.

**Input Data (fra Datagrundlag)**

• **Fil 1:** Rivers\_gvf\_rev20230825\_kontakt.shp (14 454 vandløbssegmenter)  
 – **Anvendte kolonner:**  
  • GVForekom – GVFK-navn tilknyttet hvert vandløbssegment  
  • Kontakt – Numerisk flag (1 = har kontakt, 0 = ingen kontakt)  
 – **Datatype:** GeoDataFrame med vandløbslinjer

• **Fil 2:** VP3Genbesøg\_grundvand\_geometri.shp (genbrugt fra Trin 1)  
 – **Anvendte kolonner:**  
  • Navn – GVFK-identifikator for matching med vandløbsdata  
 – **Datatype:** GeoDataFrame med GVFK-polygoner

**Proceslogik** (step2\_river\_contact.py)

1. **Vandløbsfiltrering**  
    – Indlæser vandløbsdata med geopandas.read\_file()  
    – Filtrerer til kun segmenter hvor Kontakt == 1 (7.496 af 14.454)  
    – Årsag: Kun segmenter med aktuel grundvand–overfladevand-interaktion er relevante
2. **GVFK-ekstraktion**  
    – Udtrækker unikke værdier fra GVForekom-kolonnen  
    – Fjerner None-værdier og ikke-teksttyper  
    – Opretter liste med 593 GVFK-navne
3. **Geometri-kobling**  
    – Indlæser GVFK-geometri fra Trin 1  
    – Filtrerer hvor Navn findes i vandløbskontakt-listen  
    – Årsag: Bevar kun GVFK-geometrier med dokumenteret vandløbskontakt
4. **Output-lagring**  
    – Gemmer filtrerede GVFK-geometrier til step2\_river\_gvfk.shp

**Output**  
• Fil: step2\_river\_gvfk.shp (588 GVFK-geometrier)  
• Returnerer: (liste\_med\_593\_GVFK\_navne, antal\_unikke\_GVFK, GeoDataFrame)

**Aktuelle Resultater**  
– 593 GVFK har kontakt med vandløb (29,0 % af alle GVFK)  
– 588 GVFK-geometrier gemt med vandløbskontakt

**Forskel mellem 593 og 588 skyldes:**

1. Fejl i navne mellem VP3Genbesøg\_grundvand\_geometri.shp og Rivers\_gvf\_rev20230825\_kontakt.shp.
2. 5 GVFK-navne i vandløbsdata findes ikke i geometrifilen.

## Trin 3: V1/V2-lokaliteter med Aktive Forureninger i GVFK med Vandløbskontakt

**Formål**  
Identificere V1/V2-lokaliteter med aktive forureninger i grundvandsforekomster med vandløbskontakt. Bevare én-til-mange lokalitet–GVFK-relationer, som er kritiske for korrekte afstandsberegninger i Trin 4.

**Eksempel på én-til-mange-relation**  
Lokalitet *“12345”* kan overlappe flere GVFK-polygoner (*“GVFK\_A”* og *“GVFK\_B”*), hvilket resulterer i to kombinationer: (12345, GVFK\_A) og (12345, GVFK\_B). Hver kombination kræver separate afstandsberegninger.

**Input Data (fra Datagrundlag)**

1. **CSV-filer (pre-processeret via ArcGIS spatial join)**  
    • v1\_gvfk\_forurening.csv (84.601 rækker)  
     – **Anvendte kolonner:**  
      • Lokalitetsnr – Lokalitetsidentifikator  
      • Navn – GVFK-navn (fra ArcGIS spatial join)  
      • Lokalitetensstoffer – Forureningsstoffer (kritisk filterkolonne)  
      • Lokalitetensbranche, Lokalitetensaktivitet – Metadata til anvendelse i Trin 5

 • v2\_gvfk\_forurening.csv (134.636 rækker)  
  – Samme kolonnestruktur som V1

1. **Shapefiler (geometrisk data)**  
    • V1FLADER.shp (28.717 polygoner → 23.209 unikke lokaliteter)  
     – **Anvendte kolonner:**  
      • geometry – Polygongeometrier for forurenede lokaliteter  
      • Lokalitet\_ – Lokalitetsidentifikator (matcher Lokalitetsnr fra CSV)

 • V2FLADER.shp (33.040 polygoner → 21.269 unikke lokaliteter)  
  – Samme struktur som V1

1. **Fra Trin 2**  
    Liste med 593 GVFK-navne med vandløbskontakt

**Proceslogik (step3\_v1v2\_sites.py)**

1. **Filtrering**  
    – Filtrer hvor *Lokalitetensstoffer* eller *Lokalitetensbranche / Lokalitetensaktivitet* ikke er tomme  
    – V1: 84.601 → 84.401 rækker (99,8 % retained)  
    – V2: 134.636 → 134.491 rækker (99,9 % retained)  
    – Årsag: Inkluderer lokaliteter med dokumenterede aktive forureninger **eller** branche/aktivitet oplysninger, da dette kan indikere potentielle risici også uden stofdata.
2. **Geometri-processering**  
    – Indlæs shapefiler med geopandas.read\_file()  
    – *Dissolve* geometrier efter *Lokalitet\_* for at håndtere multipart-polygoner  
    – Årsag: Enkelte lokaliteter kan bestå af flere separate polygoner
3. **Vandløbskontakt-filtrering**  
    – Filtrer CSV-data, hvor *Navn* findes i Trin 2’s GVFK-liste  
    – Årsag: Bevar kun lokaliteter i GVFK med dokumenteret vandløbskontakt
4. **Data-kobling**  
    – Join CSV-attributter med *dissolved* geometrier via *Lokalitetsnr*  
    – Resultat: Komplet spatial + attribut-datasæt
5. **Deduplikering og datakonsolidering (to-trins proces)**  
    – **Trin 1:** Aggreger lokalitet–GVFK-kombinationer inden for V1/V2  
     • Stoffer: Sammensæt alle unikke stoffer med semikolon-adskillelse  
     • Andre felter: Bevar første værdi (identisk indenfor samme kombination)  
     • Årsag: Bevarer *alle* forureningsstoffer per lokalitet–GVFK-kombination  
    – **Trin 2:** Håndter lokaliteter i både V1 og V2 (marker som *“V1 og V2”*)  
     • Stoffer: Sammensæt stoffer fra både V1- og V2-registreringer  
     • Årsag: Sikrer komplet stoffortegnelse for lokaliteter med dobbelt klassificering

**Output**  
• Filer:  
 – step3\_v1v2\_sites.shp – Alle lokalitet–GVFK-kombinationer med geometri  
 – step3\_gvfk\_with\_v1v2.shp – GVFK-polygoner med V1/V2-lokaliteter  
 – step3\_site\_gvfk\_relationships.csv – Detaljerede relationsdata  
• Returnerer: (sæt\_med\_GVFK\_navne, v1v2\_kombineret\_GeoDataFrame)

**Aktuelle Resultater**  
– 35728 unikke V1/V2-lokaliteter med aktive forureninger eller brancheoplysninger  
– 69627 totale lokalitet–GVFK-kombinationer efter deduplikering  
– 490 GVFK har V1/V2-lokaliteter (82,8 % af vandløbs-GVFK fra Trin 2)

**Lokalitet-fordeling efter type**  
– V2: 15 610 lokaliteter (43,7 %)  
– V1 og V2: 3 099 lokaliteter (8,7 %)  
– V1: 17 019 lokaliteter (47,6 %)

**Videre dataflow**  
Lokalitet–GVFK-kombinationerne fra dette trin danner grundlag for afstandsberegninger i **Trin 4**.

## Trin 4: Afstandsanalyse til Vandløb

**Formål**  
Beregne minimumsafstanden fra hver V1/V2-lokalitet til vandløbssegmenter med grundvandskontakt inden for samme GVFK.  
Dette kvantificerer forureningsspredningsrisikoen baseret på fysisk afstand.

**Input Data**  
• Fra Trin 3: 35 728 unikke V1/V2-lokaliteter med 69 627 lokalitet–GVFK-kombinationer  
• Fra datagrundlag: Vandløbsstrækninger med kontakt til grundvand (*Rivers\_gvf\_rev20230825\_kontakt.shp*)

**Proceslogik (step4\_distances.py)**

Trin 4 håndterer én-til-mange-relationer mellem lokaliteter og GVFK ved at beregne afstande for hver kombination separat.

1. **Afstandsberegning per lokalitet–GVFK-kombination**  
    – For hver af de 69 627 kombinationer:  
     • Hent lokalitetens geometri og tilknyttet GVFK-navn  
     • Find matchende vandløbssegmenter, hvor GVForekom = lokalitetens GVFK og Kontakt = 1  
     • Beregn minimumsafstand mellem lokalitetspolygon og alle matchende vandløbssegmenter  
     • Gem resultatet for denne specifikke kombination
2. **Identifikation af final minimumsafstand per lokalitet**  
    – Gruppér resultater efter lokalitets-ID  
    – Find den GVFK-kombination med absolut korteste afstand til vandløb  
    – Markér denne som lokalitetens primære risikosti (Is\_Min\_Distance = True)

**Attributmatchingslogik**  
Korrekt afstandsberegning kræver præcis matching mellem lokalitets-GVFK-tilknytninger og vandløbssegmenter:  
• Lokalitet har prædefinerede GVFK-tilknytninger fra CSV-filer (kolonnen *Navn*), skabt i tidligere spatial analyse  
• Trin 3 tilføjer kun geometrier til disse eksisterende relationer  
• Vandløbssegmenter har GVFK-tilknytning i kolonnen *GVForekom* (fra DK-Model v. Lars Troldborg)  
• Kun når *Navn = GVForekom* og *Kontakt = 1* beregnes afstand  
• Dette sikrer, at forurening kun kan nå vandløb gennem faktisk grundvand–vandløb-kontakt

**Koordinatsystem og Afstandsmåling**  
• Alle beregninger udføres i *UTM32 / EUREF89* (meter-baseret)  
• Afstande beregnes med geometry.distance() (minimum euklidisk afstand mellem geometrier)  
• Korteste afstand måles mellem ethvert punkt på lokalitetspolygonen kanten og ethvert punkt på vandløbslinjen

**Aktuelle Resultater**  
Algoritmen behandlede 69.627 lokalitet–GVFK-kombinationer med følgende resultater:  
– 35.728 unikke lokaliteter har beregnet afstand til vandløb (100 % success rate)  
– Gennemsnitlig afstand: 3.116 m til nærmeste vandløb  
– Median afstand: 1.368 m  
– Minimum afstand: 0,0 m (lokaliteter direkte ved vandløb)  
– Maksimum afstand: Varierer afhængigt af GVFK-størrelse

**Output-filer**

1. *step4\_final\_distances\_for\_risk\_assessment.csv*  
     – Én række per lokalitet med minimum afstand  
     – Kolonner: *Lokalitet\_ID*, *Final\_Distance\_m*, *Closest\_GVFK* samt forureningsmetadata  
     – Bruges direkte af Trin 5 til risikovurdering
2. *step4\_valid\_distances.csv*  
     – Indeholder alle lokalitet–GVFK-kombinationer med gyldige afstande  
     – Bruges til visualisering og detaljeret analyse
3. *unique\_lokalitet\_distances.shp*  
     – Indeholder lokalitetsgeometri og minimumsafstande  
     – Bruges til GIS-baserede visualiseringer

**Særlige Overvejelser**  
– Lokaliteter uden matchende vandløbssegmenter i deres GVFK filtreres fra  
– Trinnet bevarer alle forureningsattributter (branche, aktivitet, stoffer) til videre analyse  
– GVFK-information bevares for sporbarhed af kritiske forureningsstier

**Eksempel – Lokalitet 12345 med multiple GVFK**

──────────────────────────────────────────────  
**Inddata fra Trin 3**  
Lokalitet 12345 findes i tre lokalitet–GVFK-kombinationer:  
 • 12345 → GVFK\_A (*Navn = "DK\_GVF\_001"*)  
 • 12345 → GVFK\_B (*Navn = "DK\_GVF\_002"*)  
 • 12345 → GVFK\_C (*Navn = "DK\_GVF\_003"*)

**Niveau 1 – Beregning per kombination**  
 • *GVFK\_A:* 3 matchende vandløbssegmenter → afstande = 450 m, 720 m, 890 m → **Minimum = 450 m**  
 • *GVFK\_B:* 2 matchende segmenter → afstande = 320 m, 580 m → **Minimum = 320 m**  
 • *GVFK\_C:* 1 segment → afstand = 1 200 m → **Minimum = 1 200 m**

**Niveau 2 – Final minimum per lokalitet**  
 • Sammenlign: 450 m (GVFK\_A), 320 m (GVFK\_B), 1 200 m (GVFK\_C)   
 • **Endeligt resultat:** Lokalitet 12345 = 320 m afstand via GVFK\_B ("DK\_GVF\_002")

**Output (i step4\_final\_distances\_for\_risk\_assessment.csv)**  
 Lokalitet\_ID: 12345  
 Final\_Distance\_m: 320  
 Closest\_GVFK: DK\_GVF\_002  
 [+ metadata-kolonner]  
──────────────────────────────────────────────

## Trin 5: Tærskel-vurdering og Kategorisering

**Formål**

Identificere højrisiko V1/V2-lokaliteter baseret på afstand til vandløb og stofspecifikke mobilitetsegenskaber. Implementerer to-lags risikovurdering med både generelle og stofspecifikke tærskler.

**Inputdata**

**Fra Trin 4**: step4\_final\_distances\_for\_risk\_assessment.csv med 35.728 lokaliteter og deres minimumsafstande  
**Kategoriseringsmodul**: risikovurdering/compound\_matching.py anvender de litteraturbaserede grupper. Disse er baseret på følgende kilder:

*- Jordforureningens påvirkning af overfladevand, delprojekt 2: Afstandskriterier og fanebredder. Miljøprojekt nr. 1565, 2014. Tabel 2 og 3.*

*- Risikovurdering af lossepladsers påvirkning af overfladevand. Bjerg, Poul Løgstrup; Sonne, Anne Thobo; Tuxen, Nina; Skov Nielsen, Sanne; Roost, Sandra. 2014. Tabel 7.1.*

**Kvalificering og opdeling af input data**

1. Trin 5 modtager alle lokaliteter fra trin 4 (minimumsafstand per lokalitet).
2. Funktionen separate\_sites\_by\_substance\_data opdeler data i to grupper:
   * *Kvalificerende sites*: Mindst én registreret substans eller losseplads nøgleord i branche/aktivitet. Disse fortsætter til de efterfølgende vurderinger.
   * *”Parkede” sites*: hverken substansdata eller relevante losseplads nøgleord. De gemmes i step5\_unknown\_substance\_sites.csv til dokumentation og manuel opfølgning. **Se notatet: ”risikovurdering\_endelig\_analyse” for analysen af disse.**
3. De kvalificerende sites udgør grundlaget for både den generelle 500 m screening og den stofspecifikke analyse.

**Proceslogik (**step5\_risk\_assessment.py**)**:

### 1. Generel risikovurdering (500 m universal tærskel)

* Filtrer lokaliteter hvor Final\_Distance\_m ≤ 500 m
* Konservativ screening uafhængig af forureningstype
* Output: step5\_high\_risk\_sites.csv og GVFK-shapefiler

### 2. Stofspecifik risikovurdering med losseplads-override (to-fase tilgang)

#### Fase 1 – Kategorisering og initial screening

*Kilde: Jordforureningens påvirkning af overfladevand, delprojekt 2: Afstandskriterier og fanebredder. Miljøprojekt nr. 1565, 2014. Tabel 2 og 3.*

* Parse semikolon-separerede stoffer per lokalitet og hent kategori samt stofspecifik tærskel fra Kategoriseringsmodulet.
* Kategorierne spejler de aktive mobilitetsklasser (fx **PAH\_FORBINDELSER** 30 m, **BTXER** 50 m, **PHENOLER** 100 m, **UORGANISKE\_FORBINDELSER** 150 m, **POLARE\_FORBINDELSER** 300 m, **KLOREREDE\_OPLØSNINGSMIDLER** 500 m, **KLOREDE\_KULBRINTER** 200 m, **KLOREREDE\_PHENOLER** 200 m, **PESTICIDER** 500 m, **PFAS** 500 m, **LOSSEPLADS** 100 m, **ANDRE** 500 m) og respekterer stofspecifikke overrides som Benzen 200 m.
* Registrer allerede her om lokaliteten har losseplads-karakteristika via Lokalitetensbranche eller Lokalitetensaktivitet
* Hvis lokaliteten matcher en kategori i LANDFILL\_THRESHOLDS, sættes den effektive tærskel til max(kategori-tærskel, losseplads-tærskel) for ikke at frasortere kombinationer der kun består på grund af en lempelig lossepladsgrænse; mere restriktive losseplads-tærskler håndteres efterfølgende. Se afsnittet ”Eksempler” på næste side for mere forståelse af dette.
* Lokaliteter uden stofdata men med losseplads-brancher/aktiviteter klassificeres som **LOSSEPLADS** med en 100 m screenings-tærskel via categorize\_by\_branch\_activity
* Evaluer hver stof-lokalitet (eller branch-only) kombination mod den effektive tærskel og gem kvalificerende rækker

#### Fase 2 – Losseplads-override (post-processering)

*Kilde: Risikovurdering af lossepladsers påvirkning af overfladevand. Bjerg, Poul Løgstrup; Sonne, Anne Thobo; Tuxen, Nina; Skov Nielsen, Sanne; Roost, Sandra. 2014. Tabel 7.1.*

* Arbejder kun videre med kombinationer fra Fase 1 og fokuserer på lokaliteter med losseplads-karakteristika
* For kategorier der findes i LANDFILL\_THRESHOLDS anvendes nu den faktiske losseplads-tærskel; kombinationer over grænsen fjernes, mens de resterende reklassificeres

Reklassificering sætter Qualifying\_Category = "LOSSEPLADS", udfylderLosseplads\_Subcategory = "LOSSEPLADS\_<originalkategori>"**,** markererLandfill\_Override\_Applied = Trueog logger oprindelsen iQualifying\_Substance = "Landfill Override: <originalkategori>"

* Kombinationsrækker uden losseplads-flag eller uden relevant tærskel bevares uændret

#### Eksempler

* *BTXER ved 60 m med losseplads-flag*: passerer Fase 1 med 70 m (max(50,70)) og forbliver i datasættet efter override, nu mærket som LOSSEPLADS\_BTXER
* *Phenoler ved 80 m med losseplads-flag*: passerer Fase 1 (100 m), men fjernes i Fase 2, fordi losseplads-tærsklen er 35 m
* *Branch-only losseplads ved 120 m*: falder allerede i Fase 1, fordi den branch-baserede 100 m tærskel overskrides

### Tærskel-sammenligning (Normal vs. Losseplads-specifik)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Kategori | Normal Tærskel | Losseplads Tærskel | Repræsentativ Forbindelse | Override Status |
| **BTXER** | 50 m | **70 m** | Benzen | ✓ Override (lempligere) |
| **KLOREREDE\_OPLØSNINGSMIDLER** | 500 m | **100 m** | Tetrachlorethylen (TCE) | ✓ Override (strengere) |
| **PHENOLER** | 100 m | **35 m** | Phenol | ✓ Override (strengere) |
| **PESTICIDER** | 500 m | **180 m** | MCPP | ✓ Override (strengere) |
| **UORGANISKE\_FORBINDELSER** | 150 m | **50 m** | Arsen | ✓ Override (strengere) |
| **PAH\_FORBINDELSER** | 30 m | - | - | Ingen override |
| **POLARE\_FORBINDELSER** | 300 m | - | - | Ingen override |
| **LOSSEPLADS** | 100 m | - | - | Ingen override (allerede losseplads) |
| **ANDRE** | 500 m | - | - | Ingen override |

## Litteraturbaserede stofgrupper (refined\_compound\_analysis.py)

* refined\_compound\_analysis.py grupperer stoffer i 11 kategorier baseret på *Jordforureningens påvirkning af overfladevand, delprojekt 2* (Tabel 2 og 3).
* Kortlægningen danner grundlag for ovenstående tærskler og bruges både i foranalysen og i Trin 5.
* Kategorierne og deres standardtærskler er:
  + **BTXER** – 50 m (BTEX + oliefraktioner; tabel 2).
  + **KLOREREDE\_OPLØSNINGSMIDLER** – 500 m (klorerede opløsningsmidler; tabel 2).
  + **POLARE\_FORBINDELSER** – 300 m (MTBE m.fl.; tabel 2).
  + **PHENOLER** – 100 m (fenoler; tabel 2).
  + **KLOREDE\_KULBRINTER** – 200 m (klorerede/bromerede kulbrinter; tabel 3).
  + **KLOREREDE\_PHENOLER** – 200 m (klorerede fenoler; tabel 3).
  + **PAH\_FORBINDELSER** – 30 m (PAH; tabel 3).
  + **PESTICIDER** – 500 m (pesticider; tabel 2).
  + **PFAS** – 500 m (per- og polyfluorerede forbindelser; **IKKE NOGEN TABEL SOM KILDE**).
  + **UORGANISKE\_FORBINDELSER** – 150 m (metaller og uorganika; tabel 3).
  + **LOSSEPLADS** – 100 m (landfill-relaterede kilder).
* Stoffer der ikke matcher nogen kategori lander i **ANDRE** og får default 500 m (efter manuel opfølgning).
* **Output**: step5\_compound\_detailed\_combinations.csv med alle kvalificerende kombinationer efter override

# Endelige resultater

I alt **35.728 lokaliteter** blev modtaget fra Trin 4 og videreført til risikovurdering i Trin 5. Af disse blev **18.862 lokaliteter (52,8 %)** kvalificeret til videre behandling, idet de indeholdt enten stofdata eller lossepladsrelaterede nøgleord i branche- eller aktivitetsbeskrivelserne. De resterende **16.866 lokaliteter (47,2 %)** blev parkeret som ikke-kvalificerede, da de manglede relevante oplysninger. Blandt de kvalificerede var **13.231 stof-baserede lokaliteter**, **1.928 losseplads-baserede lokaliteter**, og **3.703 lokaliteter**, der indeholdt både stof- og lossepladsdata.

I den **generelle risikovurdering** blev der identificeret **4.156 lokaliteter** inden for en afstand på 500 meter til nærmeste vandløb med grundvandskontakt, svarende til **11,6 %** af de kvalificerede lokaliteter. I alt blev **300 unikke GVFK’er** berørt, og der blev oprettet et shapefile (*step5\_gvfk\_high\_risk*) omfattende **361 GVFK-geometrier**. De mest fremtrædende brancher blandt højrisikolokaliteterne var *servicestationer (657)*, *autoreparationsværksteder (638)* og *drift af affaldsbehandlingsanlæg (600)*. De hyppigst registrerede aktiviteter var *Andet (935)*, *Benzin og olie, salg af (666)* samt *Benzin og olie, erhvervsmæssig oplag af (461)*. På stofniveau dominerede *tungmetaller (675)*, *bly (661)* og *olieprodukter (561)*.

I den efterfølgende **stofspecifikke risikovurdering** blev **landfill-override** implementeret for udvalgte stofkategorier med definerede losseplads-tærskler: *BTXER, klorerede opløsningsmidler, phenoler, pesticider* og *uorganiske forbindelser*. I alt blev **4.355 stof-lokalitetskombinationer** evalueret, hvoraf **690 kombinationer** blev reklassificeret til *LOSSEPLADS* på baggrund af strengere tærskler. De største bidrag kom fra *uorganiske forbindelser (299 kombinationer)*, *BTXER (207)*, *pesticider (89)*, *klorerede opløsningsmidler (63)* og *phenoler (32)*. Efter justering blev antallet af registrerede *LOSSEPLADS*-kombinationer øget fra 234 til 924, mens især *uorganiske forbindelser* og *BTXER* viste markante reduktioner som følge af strengere kriterier.

Den **stofspecifikke vurdering** resulterede samlet i **1.743 højrisiko-lokaliteter (4,9 %)** ud af de 18.862 kvalificerede, svarende til **3.995 stof-lokalitetskombinationer**. I gennemsnit blev der identificeret **2,3 kvalificerende stoffer pr. lokalitet**. De berørte lokaliteter fordelte sig på **218 unikke GVFK’er**, som repræsenterer **10,7 %** af de danske grundvandsforekomster.

Fordelingen på tværs af stofkategorier viste, at *LOSSEPLADS* nu udgjorde den største kategori med **924 forekomster (415 lokaliteter)**, efterfulgt af *klorerede opløsningsmidler (807 forekomster, 412 lokaliteter)*, *BTXER (584 forekomster, 333 lokaliteter)*, *uorganiske forbindelser (542 forekomster, 300 lokaliteter)* og *pesticider (375 forekomster, 133 lokaliteter)*. Mindre udbredte grupper omfattede *andre stoffer (383 forekomster, 336 lokaliteter)*, *PAH-forbindelser (137 forekomster, 77 lokaliteter)* og *PFAS (130 forekomster, 89 lokaliteter)*.

En analyse af multi-stof-lokaliteter viste, at **1.012 lokaliteter (58 %)** indeholdt ét kvalificerende stof, **271 lokaliteter (16 %)** havde to, **176 lokaliteter (10 %)** havde tre, og **284 lokaliteter (16 %)** indeholdt fire eller flere. Den mest komplekse lokalitet omfattede **32 registrerede stoffer (lokalitet 751-00018)**.

Den samlede **GVFK-filtreringskaskade** viser den gradvise indsnævring gennem workflowet:  
Af **2.043 GVFK’er** i Danmark havde **593 (29,0 %)** vandløbskontakt (Trin 2), **432 (21,1 %)** indeholdt V1/V2-lokaliteter (Trin 3), **300 (14,7 %)** havde mindst ét site inden for 500 meter (generel vurdering), og **218 (10,7 %)** viste stofspecifik risiko i Trin 5. Den største reduktion mellem den generelle og den stofspecifikke vurdering skyldtes eksklusion af **2.413 lokaliteter**, primært med lavmobilitetsstoffer som *PAH (30 m)*, *BTEX (50 m)* og beslægtede forbindelser.

# Tilstandsvurdering

**Status**: Fremtidigt arbejde - planlagt som næste projektfase efter finalisering af risikovurderingsmetodikken.

## Planlagt Metodisk Tilgang

Tilstandsvurderingen vil bygge videre på de 2.013 højrisiko-lokaliteter identificeret gennem stofspecifik risikovurdering og omfatte:

### Kvantitativ Fluxberegning

* Beregning af forureningsflux fra de **2.013 højrisiko-lokaliteter** identificeret i risikovurderingen
* Anvendelse af infiltrationsdata fra DK-modellen: **Flux = Areal × Koncentration × Infiltration**
* Transport af flux langs strømlinjer til relevante kontaktstrækninger inden for de **232 påvirkede GVFK**

### Koncentrationsvurdering i Vandløb

* Beregning af blandingskoncentration: **C\_mix = Forureningsflux / Vandføring**
* Sammenligning med miljøkvalitetskrav (MKK) for specifike stoffer
* Identifikation af overskridelser på stofniveau

### Prioritering og Kvantificering

* Kategorisering af overskridelser efter alvorlighedsgrad
* Vurdering af mindre overskridelser (1-10 gange MKK)
* Udarbejdelse af prioriterede indsatslister

## Samarbejde med GEUS

Tilstandsvurderingen kræver tæt samarbejde med GEUS vedrørende: - Kontaktzoner og strømningsveje fra DK-modellen - Vandløbsstrækninger og vandføringsdata - Automatiserede udtræk til landsdækkende anvendelse

# Samlet Overblik

# Konklusion