

PFAS Fingerprint

- Findes der et unikt
fingeraftryk for brancher?

ATV Vintermøde 2024

Helena Hjørringgaard, Jette Kjøge Olsen, Nanna Thomsen og Monique Beyer

RAMBOLL



*Projektet er finansieret med puljemidler fra Den Syddanske Udviklingspulje for rent vand og jord
Offentliggørelse af rapportens resultater betyder nødvendigvis ikke, at Region Syddanmark er enig i rapportens indhold og konklusioner.*

Projektets formål

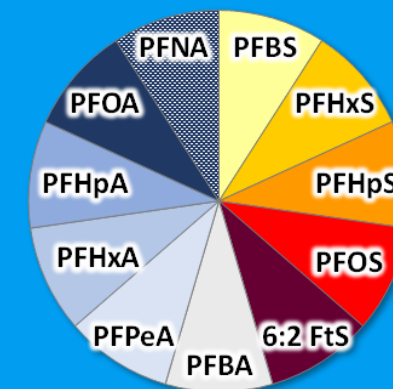
Formålet er at undersøge om der kan laves PFAS-fingerprints for brancher af terrænnært grundvand

Der findes mange tusinde PFAS-forbindelser som anvendes til forskellige formål

En branche kan have en unik sammensætning af PFAS-forbindelser → **Et fingerprint**

Et fingerprint kan anvendes til at stedfæste og tidsfæste en forurening til en bestemt kilde/branche

Projektet omfatter et litteraturstudie samt opstilling af fingerprint for brancher



Eksempel på fingerprint

Datagrundlag

- Et udtræk fra GeoGIS af regionernes PFAS-analyser indtil august/november 2023
- Et udtræk af aktiviteter tilknyttet lokaliteten i JAR
- Inddelt aktiviteter fra JAR i PFAS-relaterede brancher (14 stk.)
 - Inddelingerne er baseret på liste fra Region Midtjylland

Losse- og fyldpladser

Forkromningsindustri

Brancher relateret til brand

Træ- og møbelindustri

Kemisk industri

Jern- og metalvareindustri

Tæppeindustri

Malingsindustri

Gummi- og plastindustri

Pap- og papirindustri

Renserier

Elektronikvirksomheder

Trykkerier

Betonvarefabrikker

Ikke specificeret

Ikke PFAS-relevant

```
no_main_group_count = 0
for location in allePFAS.Lokalitetsnr.unique():
    activities = allePFAS[allePFAS.Lokalitetsnr==location].Aktiviteter
    activities_string = activities.iloc[0]
    try:
        activities_string_split = activities_string.split("#")
    except:
        # branche ikke specificeret
        continue

    main_groups = []
    for activity in activities_string_split:
        try:
            activity_identifier = activity.split("|")[2].strip()
            if activity_identifier in list(brancheGrupper.index):
                main_groups.append(brancheGrupper.loc[activity_identifier].values[0])
            elif activity_identifier in "ikke specificeret":
                main_groups.append("ikke specificeret")
            else:
                main_groups.append("ikke PFAS relateret")
                no_main_group_count += 1 #branche ikke i PFAS brancheliste
        except:
            continue
    if len(main_groups) > 0:
        for group in main_groups:
            temp_df = allePFAS[allePFAS.Lokalitetsnr==location]
            a = len(temp_df)
            temp_df.insert(1, "Hovedgruppe", [group for i in range(len(temp_df))])
            allePFAS_m_hovedgrupper_liste.append(temp_df)
```

Udklip fra vores Python script

Hvordan ser data ud?

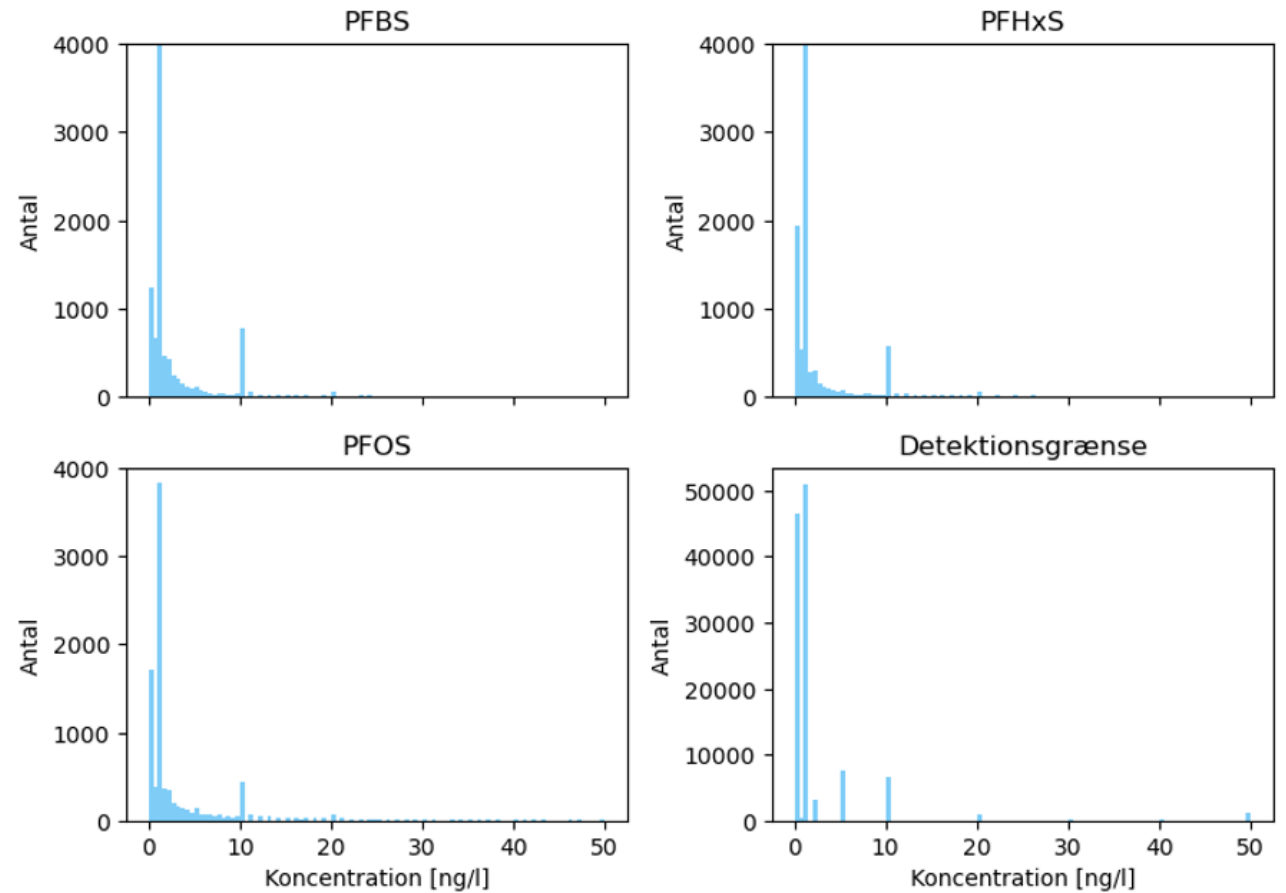
- Der er analyseret 158.235 gange for enkeltstoffer som indgår i PFAS Σ 22-kriteriet
- PFOA påvises oftest (58% á ca. 10.000 analyser)
- PFTTrDS er aldrig blevet påvist, men der er 170 analyser hvor detektionsgrænsen er hævet til mere end 100 ng/l
 - Der har været relativt mange analyser med forhøjede detektionsgrænser i datasættet
- Otte ud af de 22 stoffer påvises i mindre end 1% af analyserne
 - Ofte langkædede forbindelser

PFAS-stof	Analyserede stoffer	Påviste analyser [%]	Over kvalitetskriteriet inkl. ikke-påviste
22 PFAS stoffer	158.235	23,7	6.507
Øvrige	1.743	14,3	98
PFBA	10.076	45	404
PFPeA	9.728	29,9	567
PFHxA	9.915	38,7	595
PFHpA	9.951	40,9	412
PFOA	9.954	57,5	576/4.698*
PFNA	9.946	12,4	178/1.246*
PFDA	9.888	4,5	173
PFUnDA	4.033	0,9	170
PFDoDA	4.027	0,4	170
PFTTrDA	4.045	0,2	171
PFBS	9.737	43,2	328
PFPeS	3.850	22,4	206
PFHxS	9.967	34,9	454/2.530*
PFHpS	4.027	10,7	185
PFOS	9.957	44,4	596/3.511*
PFNS	3.850	0,2	170
PFDS	4.087	0,17	170
PFUnDS	3.842	0,03	170
PFDoDS	3.849	0,1	170
PFTTrDS	3.844	0,0	170
PFOSA	9.942	5,6	197
6:2 FTS	9.720	6,6	301

* Σ 22 / Σ 4, hvis der er et Σ 4-kriterie

Hvordan ser data ud?

- Data er ikke normalfordelt
 - Centreret omkring detektionsgrænsen
 - Højt antal af prøver ved få bestemte værdier (detektionsgrænser)
- Vi ønsker at finde sammenhænge og tendenser mellem stofferne for en branche, men det er svært at finde når der er mange parametre
- Hvordan kan man finde ud af hvilke parametre der er vigtige?
 - Reducere dimensioner i datasættet

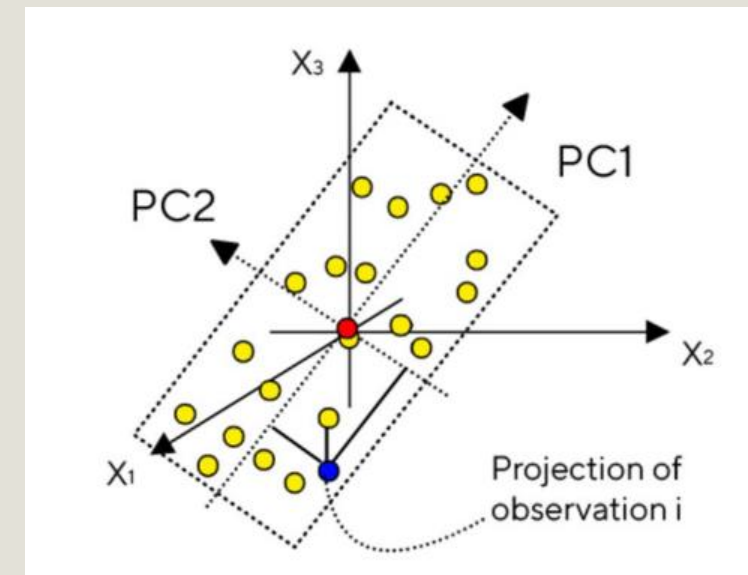
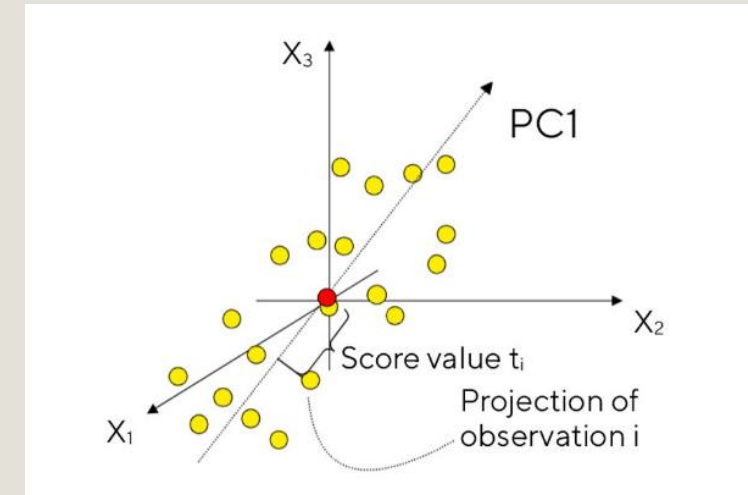


Figuren viser ikke maks koncentrationer

Principal Component Analysis (PCA)

Kort fortalt: Kan beskrive hvilke parametre der er medvirkende til at danne et mønster i data (forklarer *variansen* af data)

- Formålet er at reducere antallet af *dimensioner* (forklaringsgrader) til at forklare mest muligt af variansen i datasættet
 - Kan have mange dimensioner, men det er svært at tolke mange dimensioner
 - I vores tilfælde er parametrene PFAS-forbindelser
- Data behøver ikke være normalfordelt, men det er en fordel
- Danner en vektor i den retning der forklarer mest muligt af variansen, dimension 1, og efterfølgende en vektor ortogonal med dimension 1, som forklarer næst mest af variansen, dimension 2.



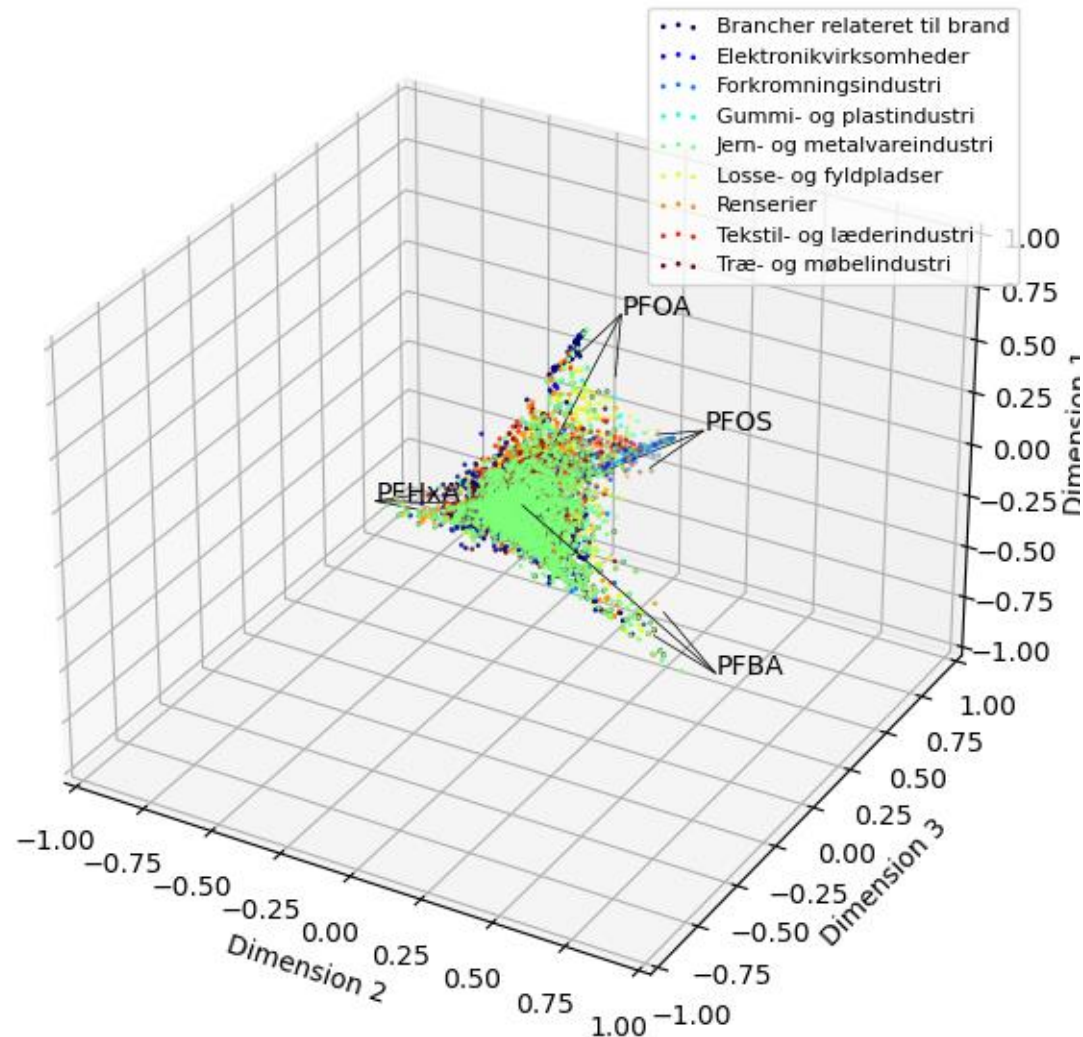
Satorius (2020)

Hvordan har vi lavet en PCA?

Der er foretaget en række antagelser for at kunne lave en PCA:

- Data er **normaliseret**. Vi regner ikke med koncentrationer, men med **procentfordelinger** af en prøves samlede indhold, da vi ønsker at lave et typisk fingerprint for en branche
- Hvis et stof ikke er påvist, er indholdet af stoffet **halvdelen af detektionsgrænsen**
- Hvis et stof ikke er analyseret, er indholdet af stoffet **halvdelen af medianen** af alle detektionsgrænser (0,5 ng/l)
- PCA'en udføres med **seks stoffer** som er påvist i mere end **30%** af alle analyser, og kun for brancher hvor der er undersøgt mere end **40 lokaliteter**
 - PFBA, PFHxA, PFHpA, PFOA, PFBS og PFOS
- Vi ved ikke hvilken branche der har forårsaget forureningen på en given lokalitet. Derfor **gentages analyser** for hver registreret branche

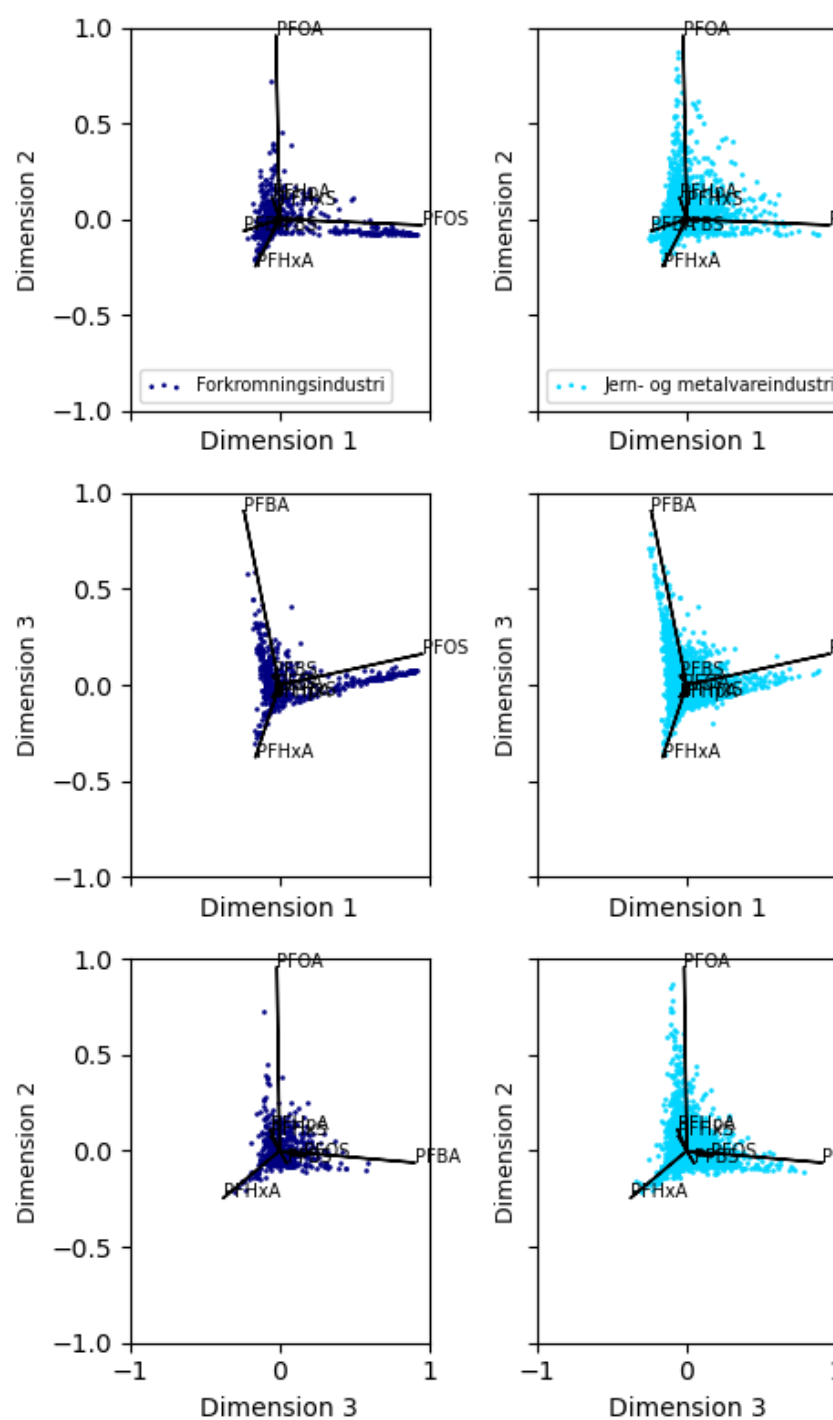
Resultater af PCA: 3D Plot



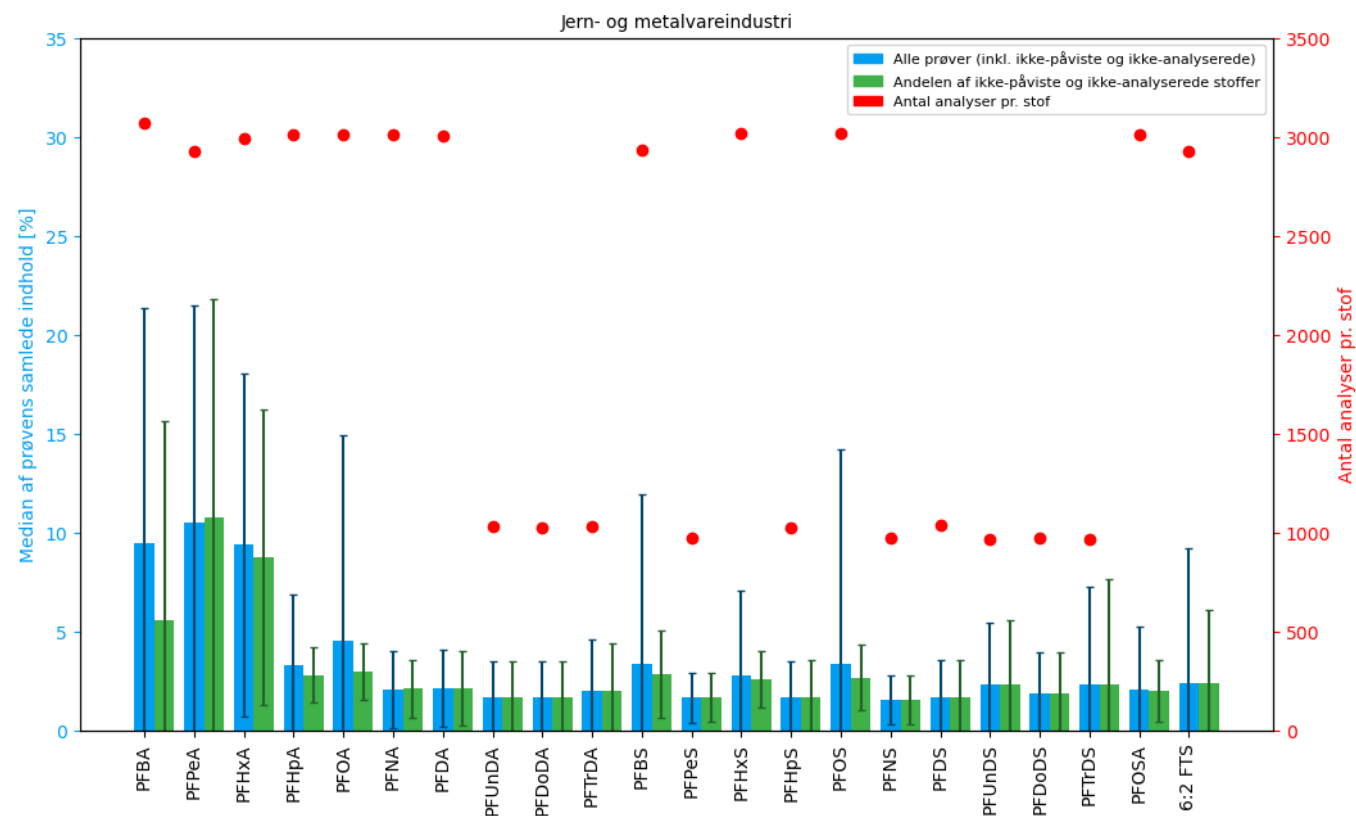
- Ingen tydelig tendens
- Brancherne grupperes ikke sammen
- Stor tæthed omkring nulpunktet
- Ingen tydelige korrelationer mellem stofferne
 - Pilene indikerer ingen korrelationer
- Alle seks stoffer er på plottet, men kan ikke ses tydeligt
- Nogle af punkterne er gentaget flere gange for forskellige brancher

Resultater af PCA

- De tre første dimensioner forklarer tilsammen 75% af variansen
 - Dimension 1 korrelerer i høj grad med **PFOS**
 - Dimension 2 korrelerer i høj grad med **PFOA**
 - Dimension 3 korrelerer med **PFBA**, og i nogen grad **PFHxA**
- PFOS korrelerer med den dimension der forklarer mest af variansen, og er påvist lige så ofte som PFBS, som kun har en lille indflydelse på de tre første dimensioner
- Der ses ingen tydelige tendenser i datasættet for de ni brancher. I projektet har vi skulle opstille fingerprints for to udvalgte brancher:
 - Jern- og metalvareindustri har mange spredte prøver
 - Forkromningsindustri har relativt mange prøver der korrelerer med dimension 1



Fingerprint: Jern- og metalvareindustri

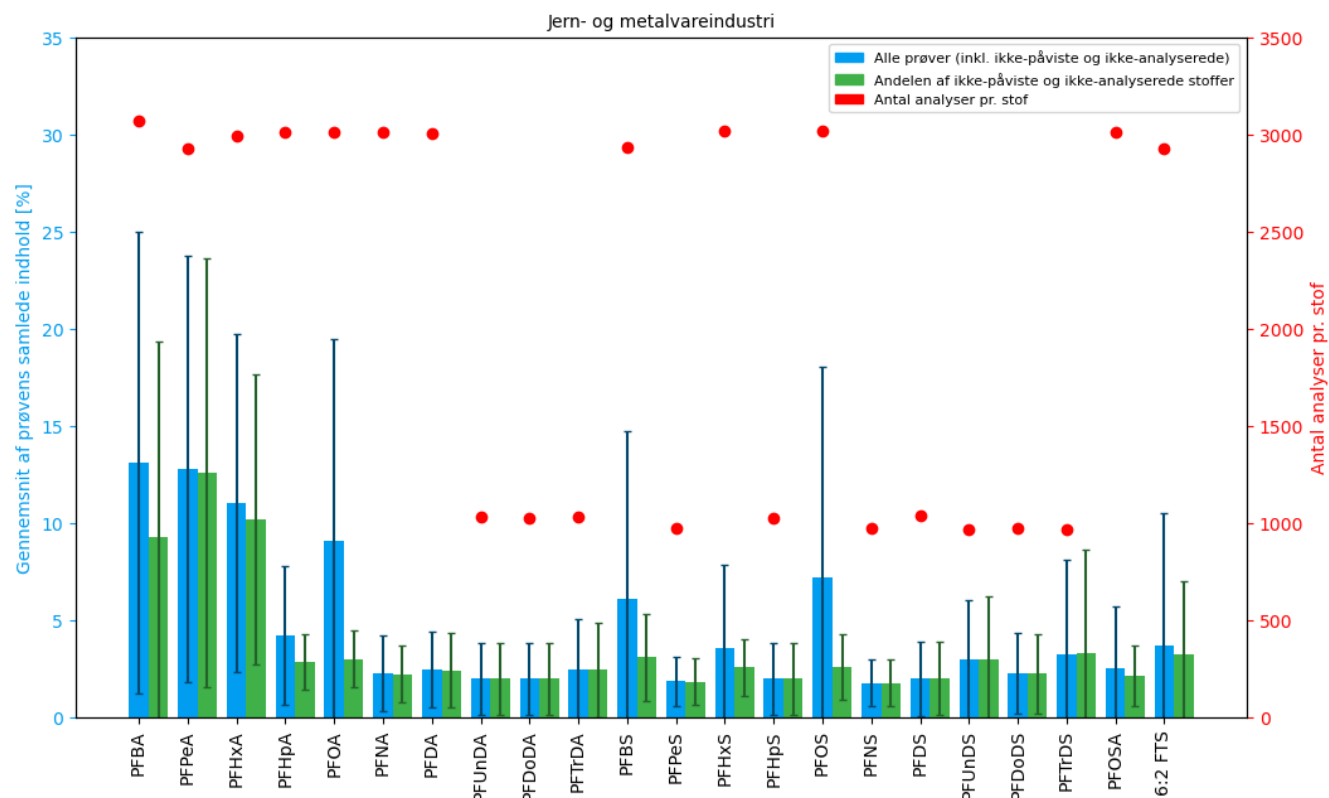


Figur 6-1 Fingerprint for jern- og metalvareindustrien. Fingerprintet viser medianen af et stofs procentmæssige indhold i en prøve samt \pm standardafvigelse. Bemærk at y-aksen går til 35% af prøvens samlede PFAS-indhold. Røde prikker og den anden y-akse er antal analyser pr. stof for branchen.

- Grønne søjler, standardafvigelser og røde prikker viser robustheden af resultaterne
- Ingen tydelig tendens for indholdet af et stof i en typisk prøve for jern- og metalvareindustrien
 - Høje standardafvigelser (under nul)
 - Lille forskel ml. alle prøver og ikke-påvist/ikke-analyserede (blå vs. grøn)
- De største medianer af procentfordelingen i en prøve er for kortkædede PFCA'er f.eks. PFBA, PFPeA og PFHxA
 - I en "typisk" prøve vil de udgøre ca. 10%

Median

Fingerprint: Jern- og metalvareindustri

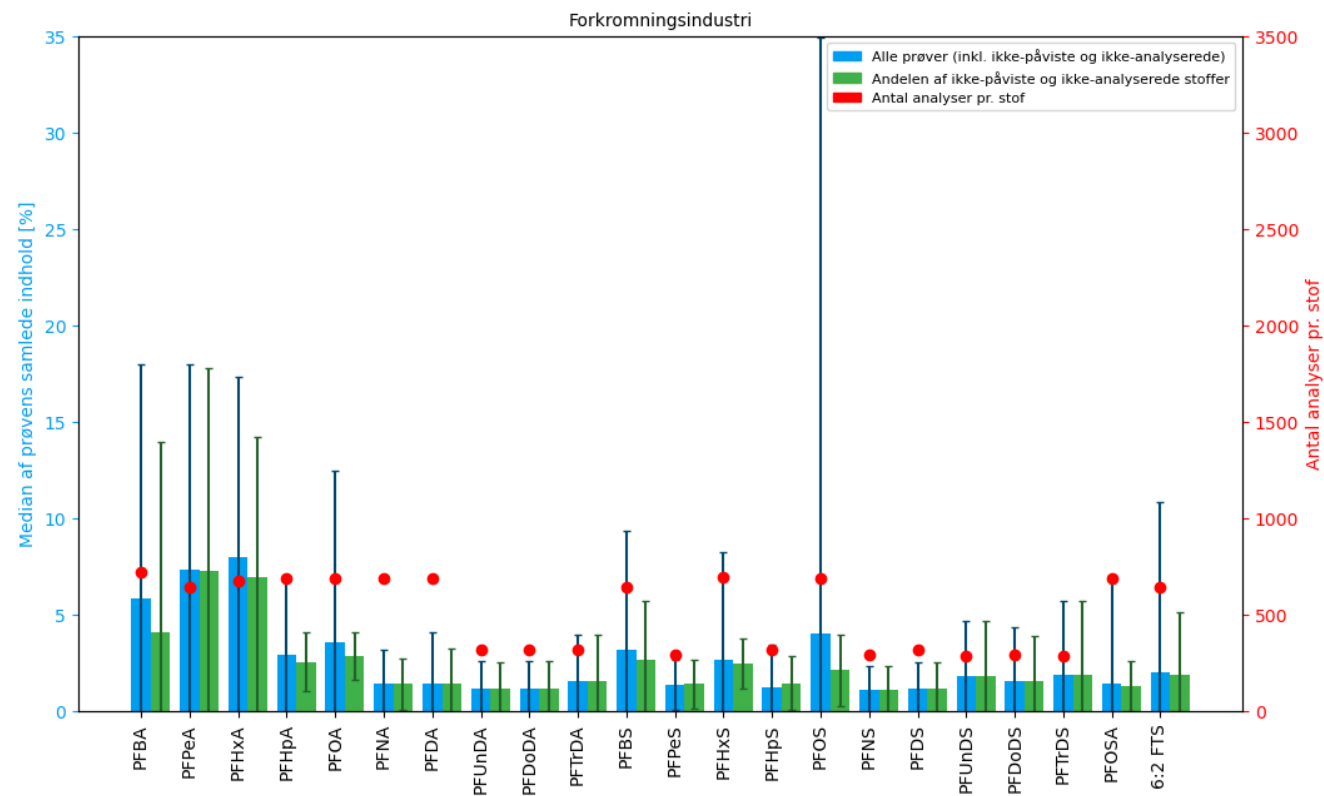


Figur 6-2 Fingerprint for jern- og metalvareindustrien. Fingerprintet viser gennemsnittet af et stofs procentmæssige indhold i en prøve samt \pm standardafvigelse. Bemærk at y-aksen går til 35% af prøvens samlede PFAS-indhold. Røde prikker og den anden y-akse er antal analyser pr. stof for branchen.

- Ingen tydelig tendens
- De største middelværdier af procentfordelingen i en prøve er for kortkædede PFCA'er f.eks. PFBA, PFPeA og PFHxA
 - I en "typisk" prøve vil de udgøre ca. 11-14%
- De største procentmæssige forskelle ml. middel og ikke-påvist/ikke-analyseret er for PFBA, PFOA, PFBS og PFOS
 - Når disse forbindelser påvises i en prøve, vil de typisk udgøre en større andel
- I litteraturstudiet fandt vi at PFOA, PFHxS, PFOS, PFOSA og 6:2 FTS blev anvendt/anvendes i jern- og metalvareindustrien

Middel

Fingerprint: Forkromningsindustri

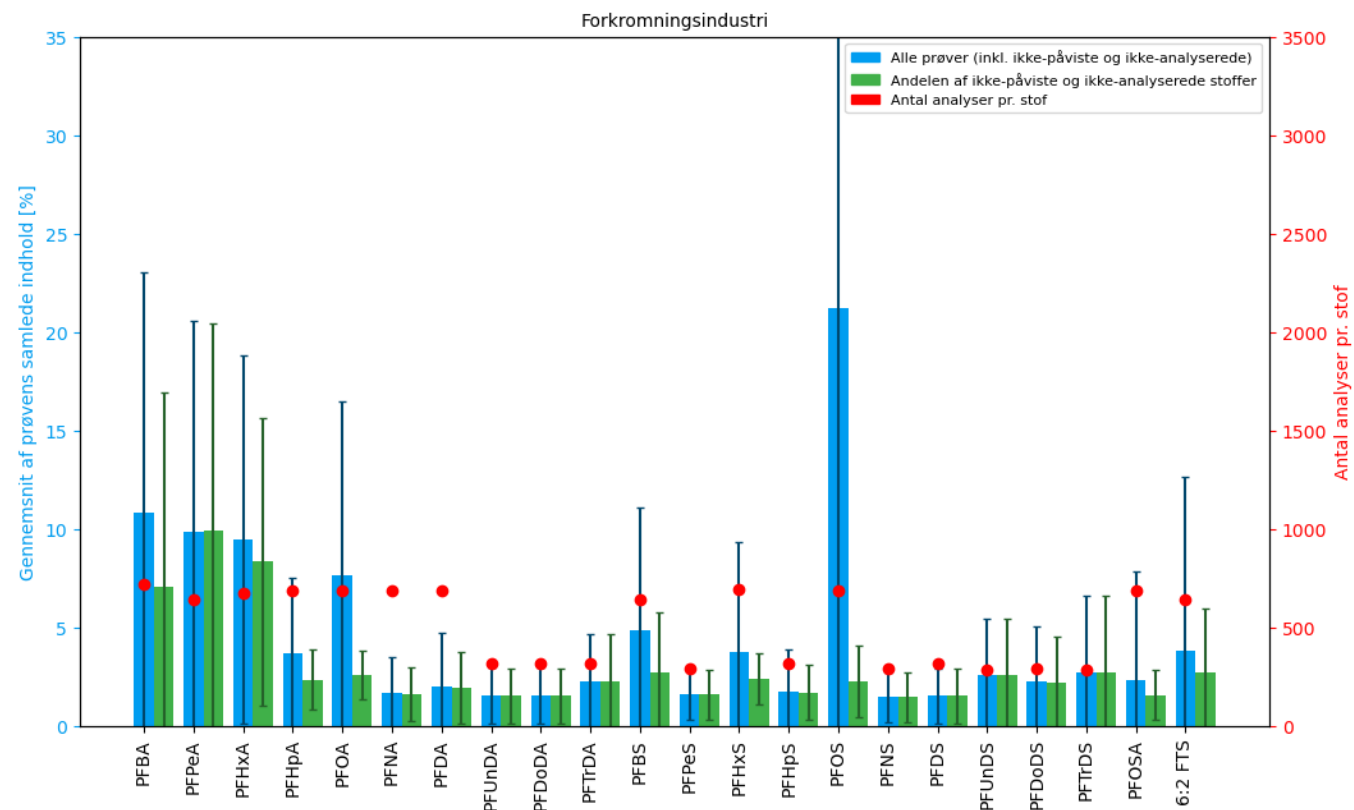


Figur 7-1 Fingerprint for forkromningsindustrien. Fingerprintet viser medianen af et stofs procentmæssige indhold i en prøve samt ±standardafvigelse. Bemærk at y-aksen går til 35% af prøvens samlede PFAS-indhold. Røde prikker og den anden y-akse er antal analyser pr. stof for branchen.

- Grønne søjler, standardafvigelser og røde prikker viser robustheden af resultaterne
- Ingen tydelig tendens for indholdet af et stof i en typisk prøve for forkromningsindustrien
 - Høje standardafvigelser (under nul)
 - Lille forskel ml. alle prøver og ikke-påvist/ikke-analyserede (blå vs. grøn)
- Høj standardafvigelse for PFOS som går til 35%
- I PCA'en så vi at der var en korrelation med PFOS
- De største medianer af procentfordelingen i en prøve er for kortkædede PFCA'er f.eks. PFBA, PFPeA og PFHxA
 - I en "typisk" prøve vil de udgøre ca. 6-8%

Median

Fingerprint: Forkromningsindustri



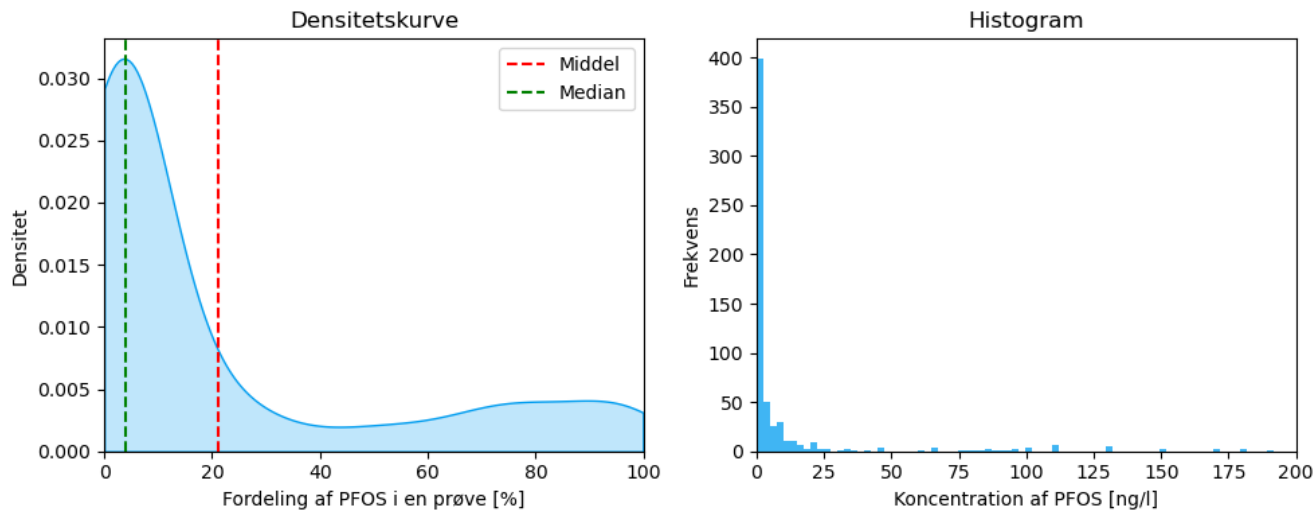
Figur 7-2 Fingerprint for forkromningsindustrien. Fingerprintet viser gennemsnittet medianen af et stofs procentmæssige indhold i en prøve samt \pm standardafvigelse. Bemærk at y-aksen går til 35% af prøvens samlede PFAS-indhold, men gennemsnittet + standardafvigelsen for PFOS går til 53%. Røde prikker og den anden y-akse er antal analyser pr. stof for branchen.

- Ingen tydelig tendens
- Høje standardafvigelser
 - PFOS går til 53%
- Der er stor forskel ml. median og middel for PFOS. Fra 4% til 21%
- De største middelværdier af procentfordelingen i en prøve er for PFOS og kortkædede PFCA'er
 - I en "typisk" prøve vil de udgøre hhv. 21% og ml. 9-11%
- I litteraturstudiet fandt vi at PFOS og 6:2 FTS blev anvendt/anvendes i forkromningsindustrien

Middel

Fingerprint: Forkromningsindustri

Hvorfor er der så stor forskel på PFOS?



Figur 7-3 Estimeret densitetskurve af procentfordelingen af PFOS i en prøve samt histogram af koncentrationer for PFOS. Histogrammet er afskåret til intervallet 0-200 ng/l.

- Data er binormal (to *modes*)
 - Major mode er ved detektionsgrænsen
 - Minor mode er ca. 75%-95%
- Det vil sige, at når PFOS påvises i en prøve, så vil den måske udgøre omkring 75%-95% af prøvens samlede PFAS-indhold
- Både middel og median underestimerer den "sande" værdi når PFOS påvises
 - Ikke-påviste analyser har en stor indflydelse på resultatet!

Kan vi lave fingerprints?

- På nuværende tidspunkt er der usikkerheder forbundet med det – særligt pga. en stor andel af ikke-påviste analyser
- Kan formentlig laves med større sikkerhed, men vil kræve noget mere arbejde f.eks.:
 - Andre analysemetoder og antagelser ift. ikke-påviste analyser (sensitivitetsanalyse)
 - Grundigere datavask – hvordan undgår vi gentagelser af prøver ved flere brancher?
 - Skal vi bruge indikatorstoffer for brancher? Skal PFOS f.eks. være påvist for forkromningsindustri?
 - Skal der være en angivelse af om boringen er placeret ved kilde eller er afgrænsende?
- Men der kan sagtens lave fingerprints på mindre skala og ved konkrete cases
- Rambøll har anvendt fingerprints til at tids- og stedfæste en forurening for en kunde i USA i forbindelse med påbudssag
- Fingerprints kan være brugbare, men svært at sige noget om et generelt billede af brancher med det anvendte datasæt og indledende metoder

Bright
ideas.
Sustainable
change.

RAMBOLL



Helena Hjørringgaard
hhd@ramboll.dk