

# Datadriven dammsäkerhet

SWEDCOLD, 20 okt 2020

Kerim Genel Waldenström, Vattenfall Vattenkraft AB

Torbjörn Johansson, Anders Björk, Håkan Fridén, Anton Jacobsson,  
Oscar Samuelsson, IVL Svenska Miljöinstitutet AB



# IVL Svenska Miljöinstitutet

- Ca 350 anställda i Sverige, Kina och Indien
- Oberoende, icke-vinstdrivande forskningsinstitut
- Vision: Ett Hållbart Samhälle



- Omfattande erfarenhet av statistisk och matematisk modellering.
- Implementationsvana
- Flera leder och har lett komplexa EU- och nationella projekt

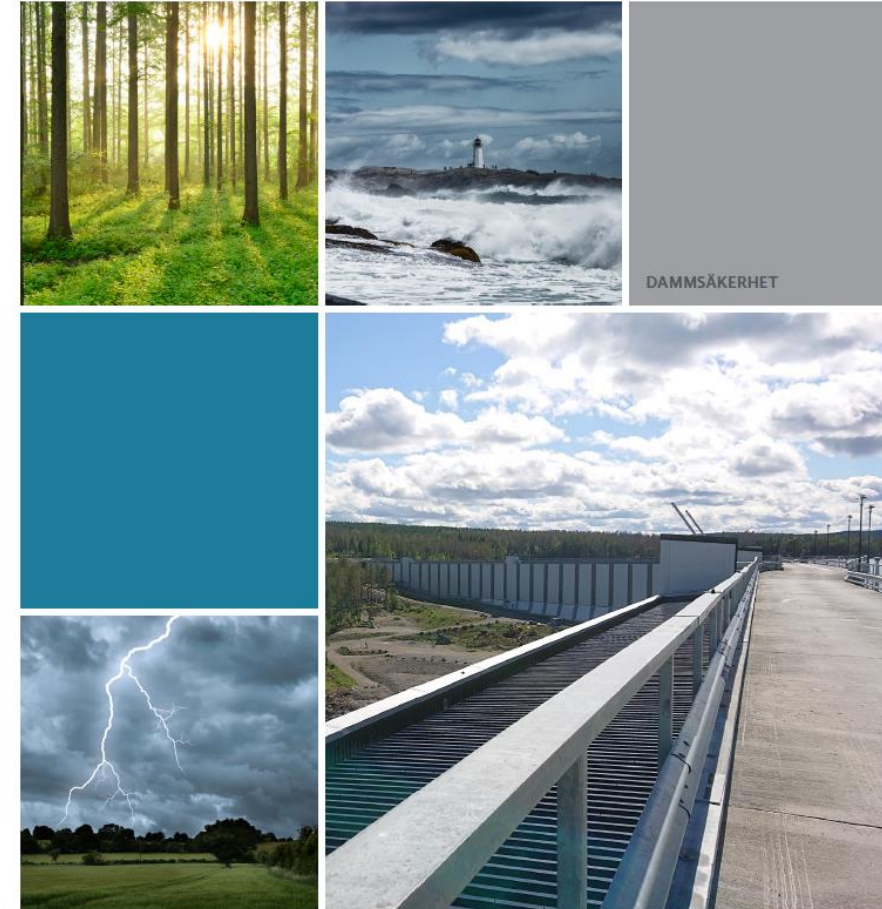
Mer tillämpade än universitet  
Stor bredd (många branscher och kompetenser)  
Många disputerade forskare



# Energiforsk rapport: datadrivna metoder för dammövervakning

Finansierat av Energiforsk och SIVL  
och utfört vintern och våren 2019

- Kunskapsläget
- Metoder
- Rekommendationer



# Vad är en datadriven metod?

- Sensorer
- Insamling av sensordata
- Analys



# Sex vanliga antaganden vid datadriven analys

- Data är korrekta och pålitliga
- Variationer i data beror på faktorer vi observerar
- Data kan analyseras med valfri metod som verkar ge rimliga resultat
- Alla data är lika relevanta
- Data är oberoende eller beror linjärt av varandra
- Mängden data är en bra värdemätare på dess nytta



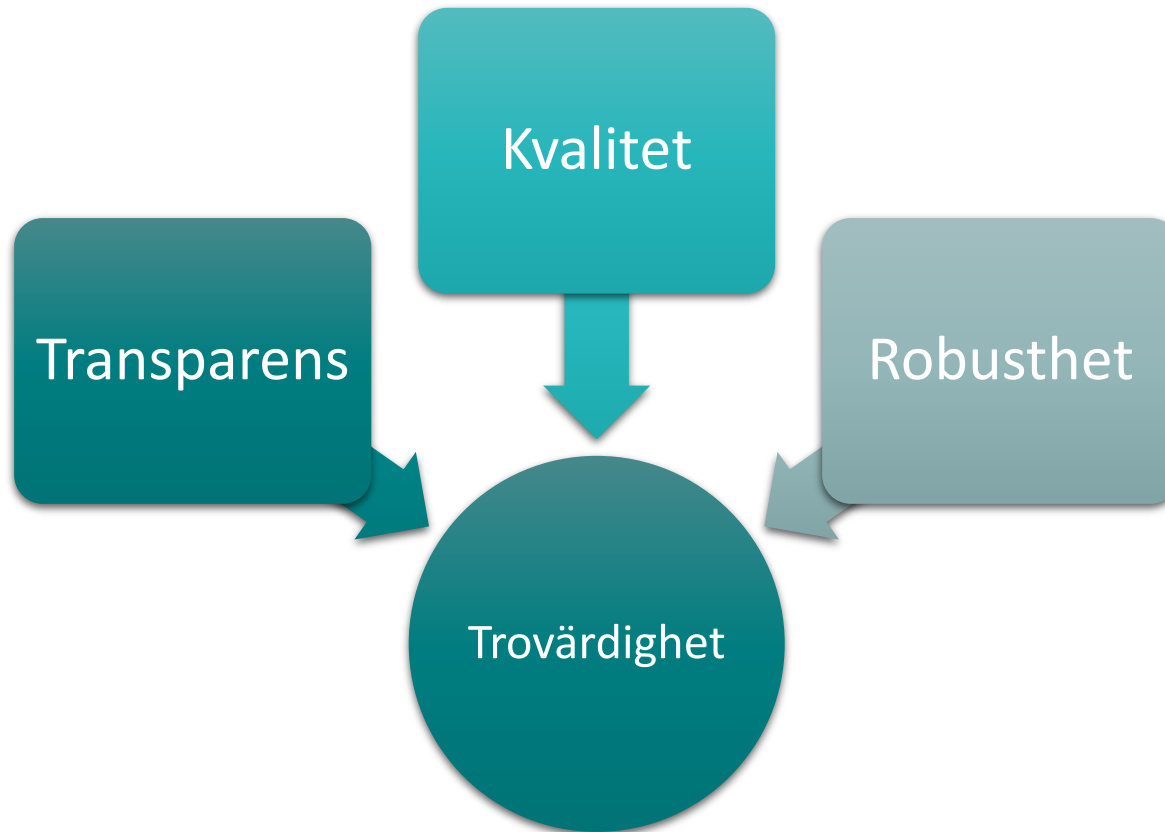


# Implementation av datadriven metod

- Utveckling i nära samarbete med användare
  - Domänkunskap ovärderlig
- Iterativ utveckling – erfarenhetsåterföring
- Förbehandling / "datatvätt" viktigt
- Utveckling på historiska data - provdrift på realtidsdata



# Datadrivna metoder måste vara trovärdiga





## Tidigare resultat om datadrivna metoder för dammövervakning

- Totalt 57 artiklar och böcker refererade i rapporten
  - 30 om dammsäkerhet
  - 23 om datadriven dammsäkerhet
  - 50 om datadrivna metoder

# Detektion av avvikande beteende med hjälp av prediktion

- Prediktera sensordata
- Mät skillnad mellan observerade och prediktera sensordata
- Ju större skillnad, ju troligare att beteendet avviker
- Överanpassning är vanligt i litteraturen
  - I prediktionen
  - Träning på endast "normala" data kan vara ok

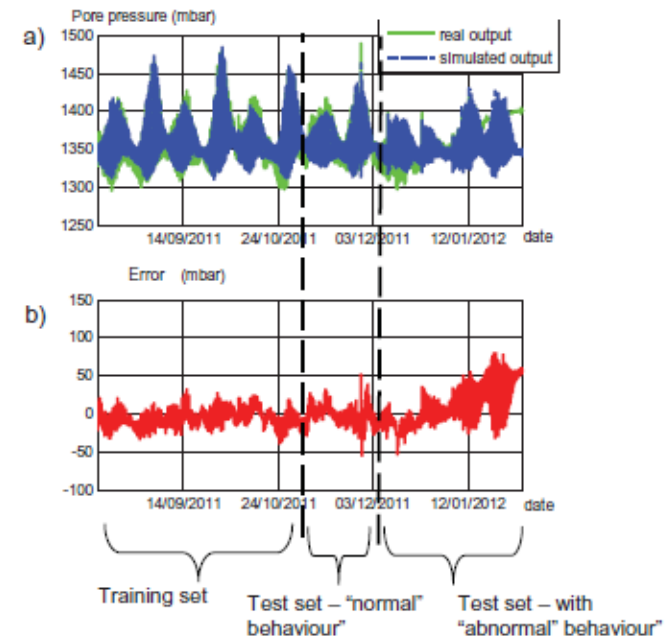
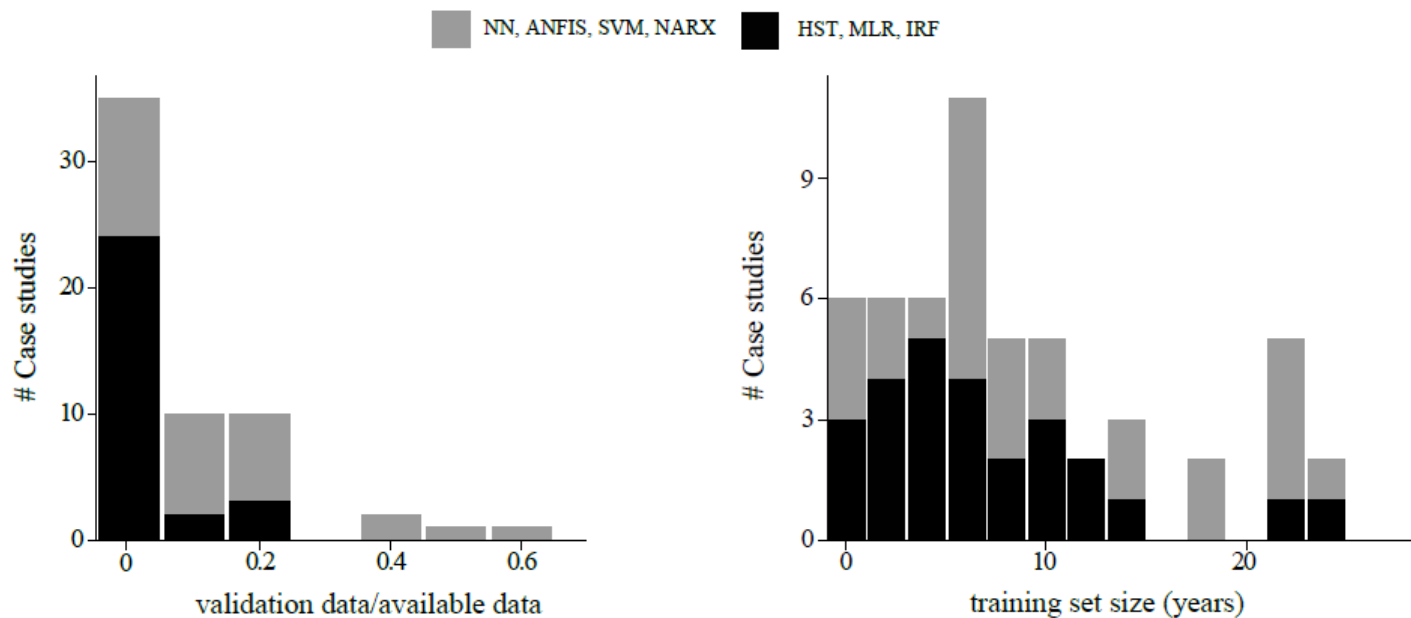


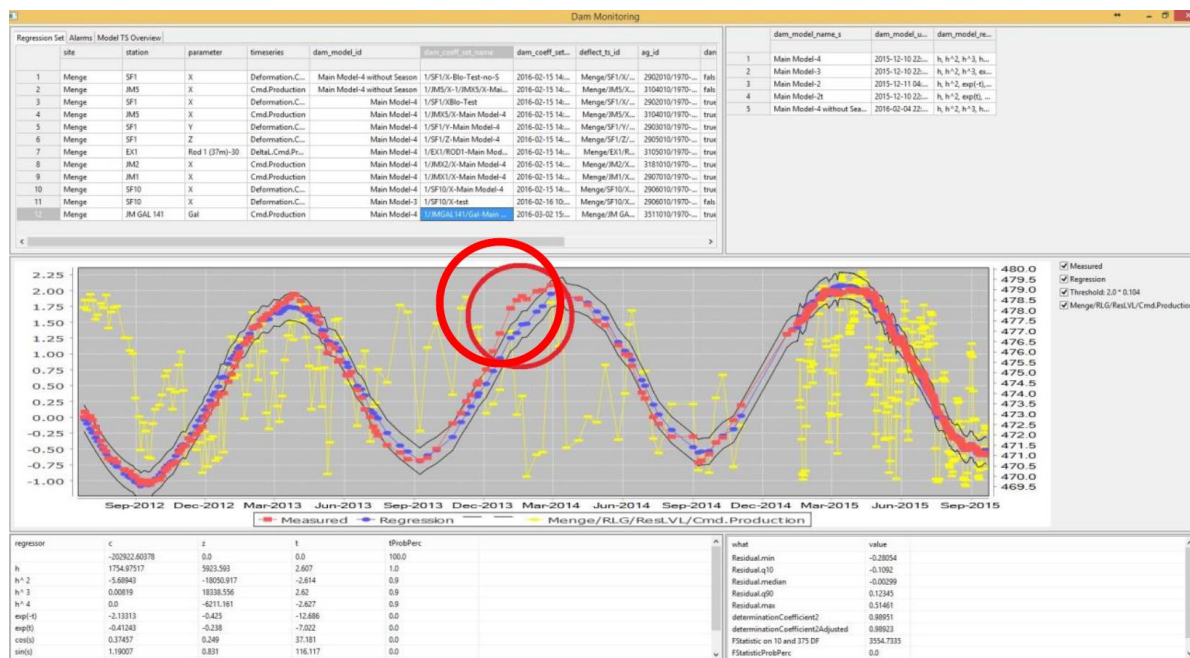
Figure 9. a) Results of one-step forecasting using the *second* linear model; b) Error of one-step forecasting.

## Träning av en datadriven metod kräver flera års data – och validering är viktigt



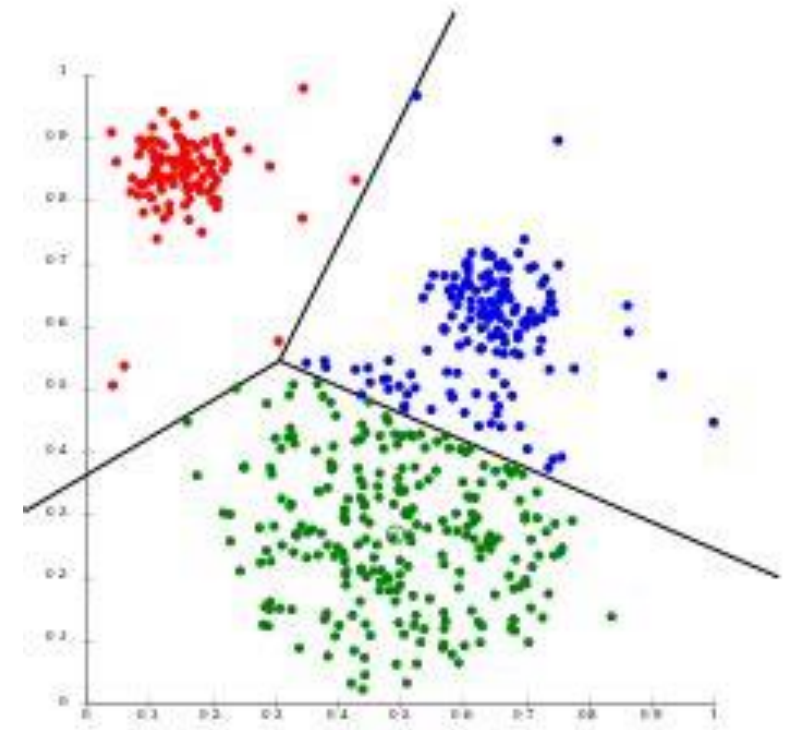


# Dammövervakningssystem för 9 dammar i Turkiet



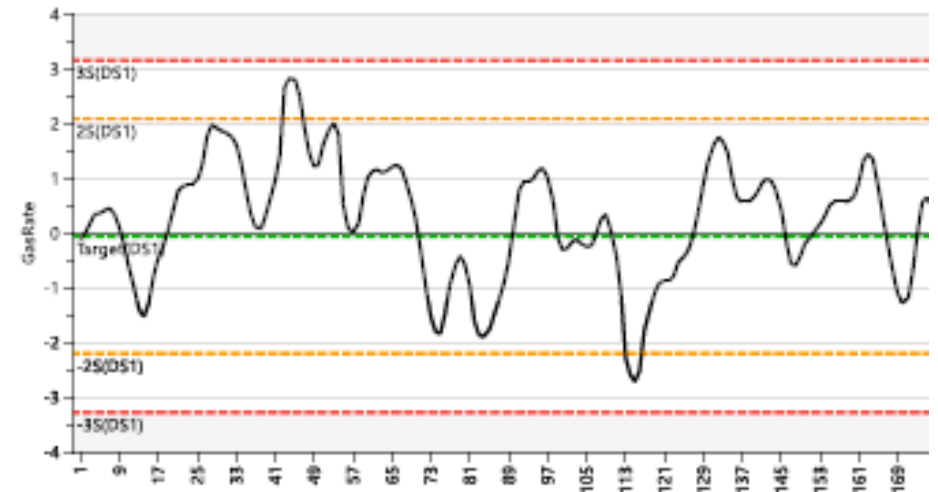
# Prediktion av data

- Hydrostatic Season Time (HST) och varianter
  - Pseudolinjär regressionsmodell
  - Fysikaliskt relevant för betongdammar
- Neurala nätverk och andra moderna maskininlärningsmetoder
  - Icke-linjära
  - Generella och kraftfulla
- Statistiska metoder
  - Principalkomponentanalys (PCA)
  - Pseudolinjär Least Squares (PLS)



# Avvikelsedektion

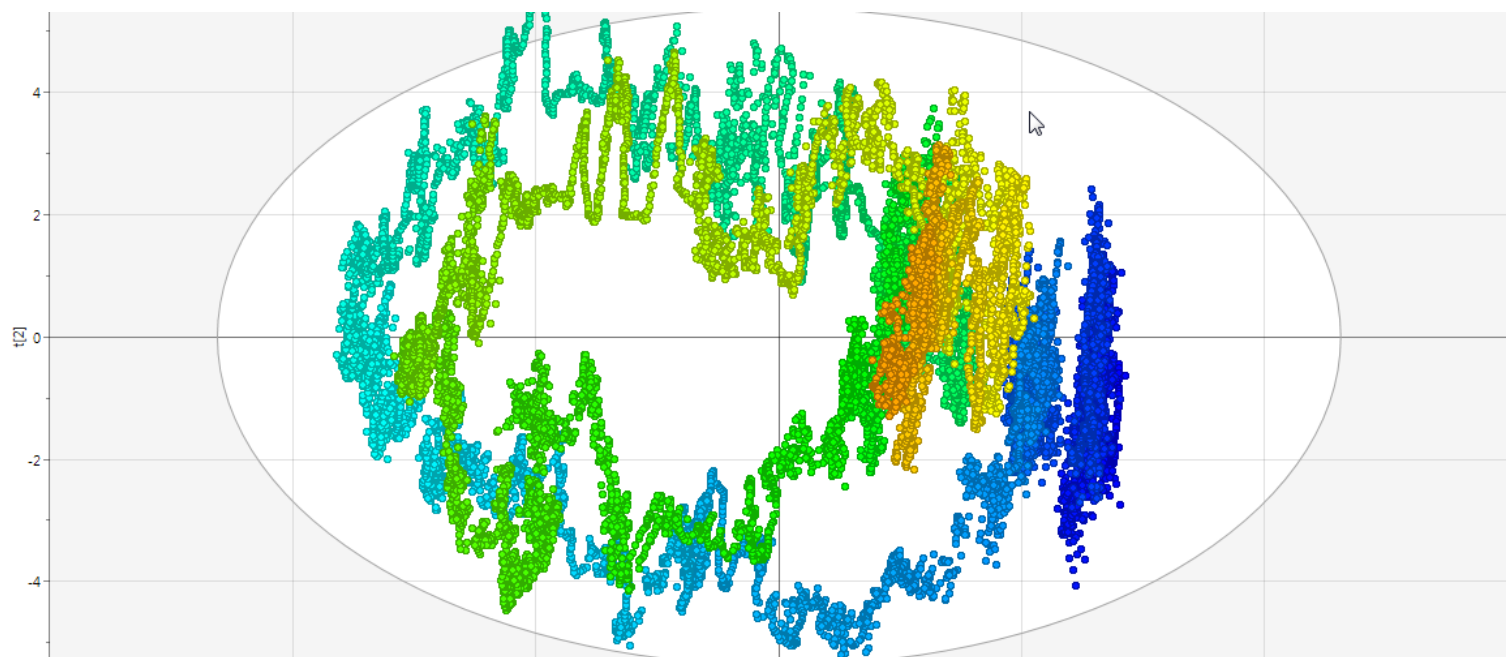
- (Multivariate) Statistical Process Control
  - Baserat på medel och standardavvikelser
- Finite and infinite moving average control charts
- Metoder från mönsterigenkänning
  - Support Vector Machines
  - Linjär diskriminant
  - Beslutsträd



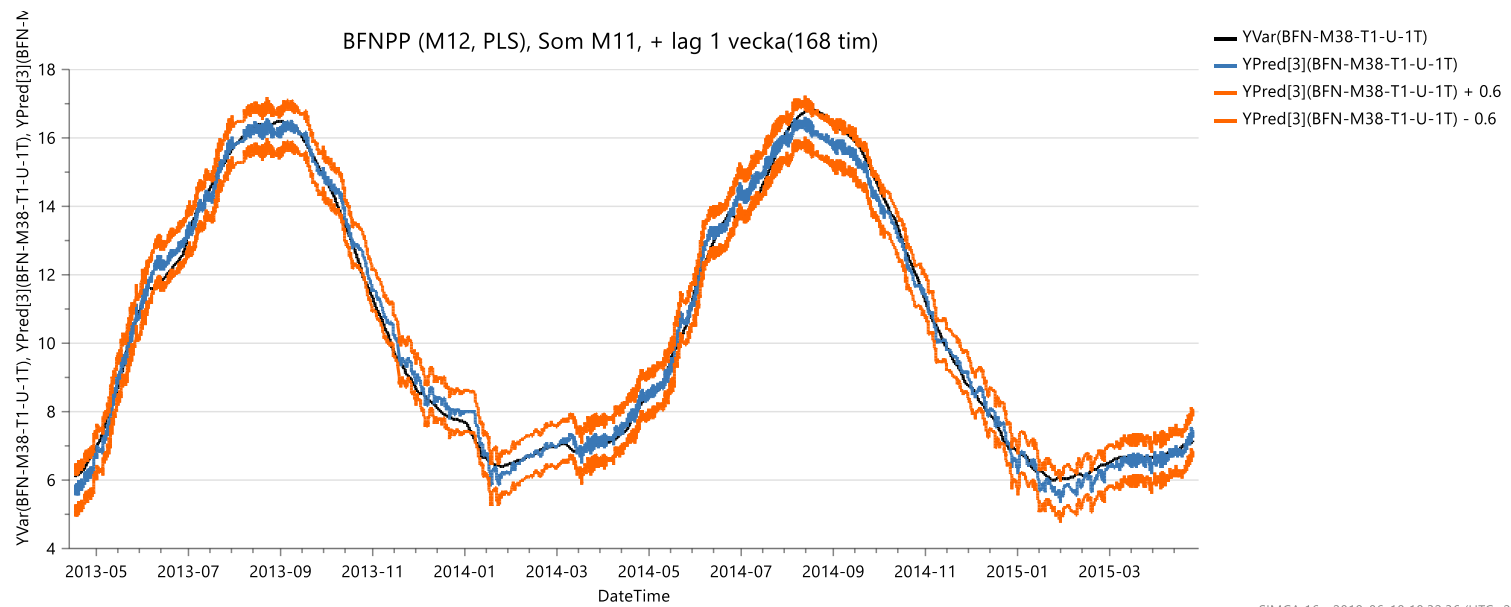


PCA, 2 komponenter från 36 variabler  
förklarar 75 % av data

# Exempel: övervakning av Bålforsens damm



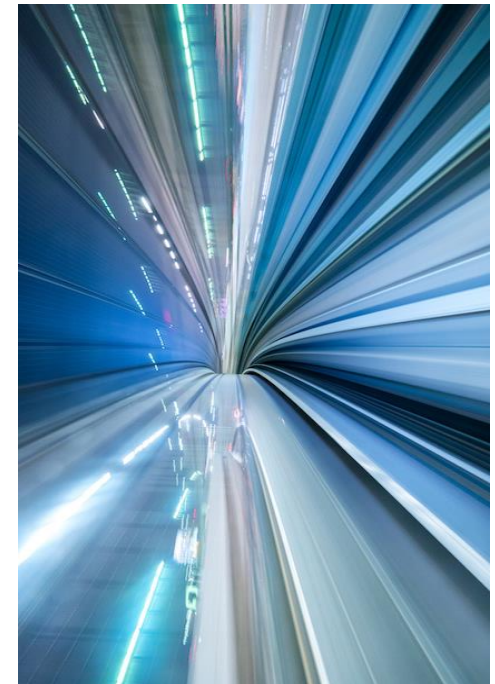
# Exempel: prediktion av data från Bålforsen



SIMCA 16 - 2019-06-18 18:32:36 (UTC+2)

# Rekommendationer för vidare arbete

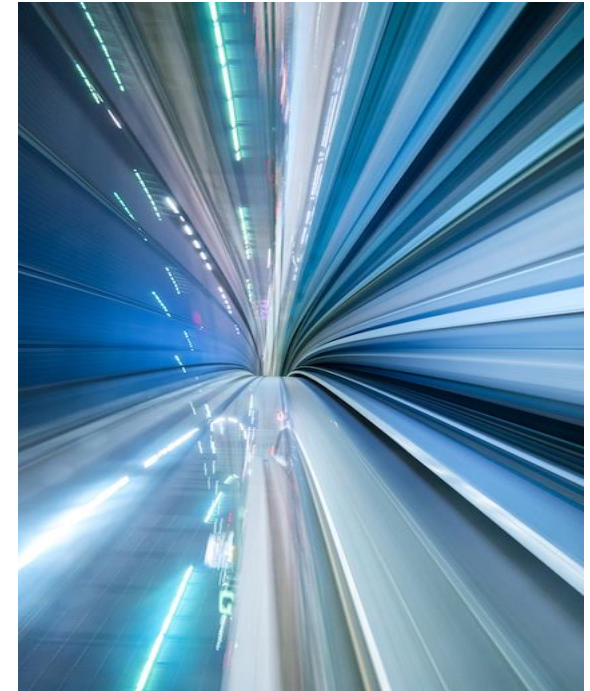
- Storskalig utvärdering på data från flera dammar
  - Detektion av sensorfel
  - Avvikelsesdetektion
- Utvärdera data
- Samla in data om regn och snö





# Rekommendationer för vidare arbete

- Samarbete!
  - Dataanalytiker
  - Dammexperter
  - Sensoreexperter
- Studera avvikardetektion
  - Metoder för att skilja på sensorfel och anomalier i dammen
  - Metoder för att ställa in larmgränser automatiskt
  - Statistiska metoder för avvikardetektion
- Utbilda mottagande organisationer



# Fortsättningsprojekt 2020-2021

- Finansierat av Energiforsk och SIVL
- Automatisk detektion av sensorfel
- Implementation av demonstrator i Vattenfalls system