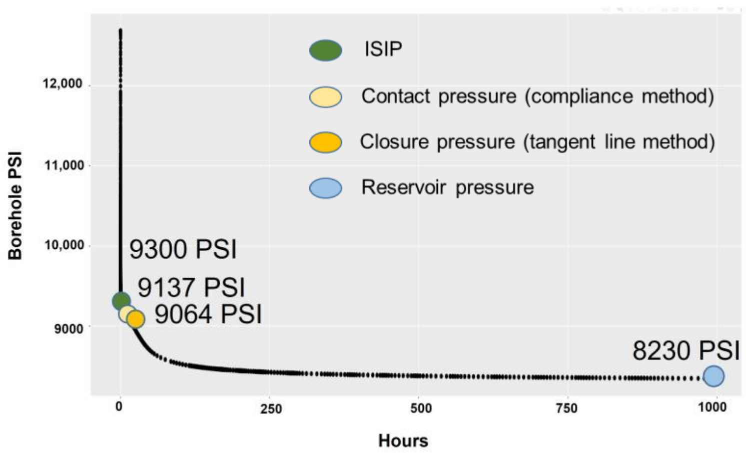
Shiny web applications for optimization of DFIT event detection with unsupervised learning

Lukas Sadownyk, Marcelo Guarido, Danial Zeinabady, Erfan Sarvaramini, Zhenzihao Zhang, Farshad Tabasinejad, Hashem Salari, Kristopher A. H. Innanen, and Christopher R. Clarkson

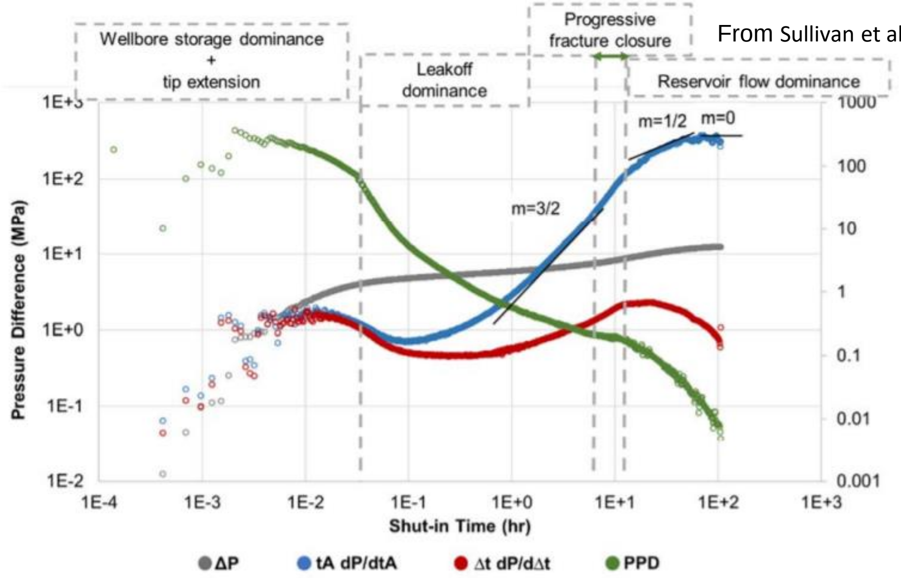
December 3, 2021, 11:40am

Presentation Overview

31- Layer Duvernay history match DFIT



From Sullivan et al. (2019)

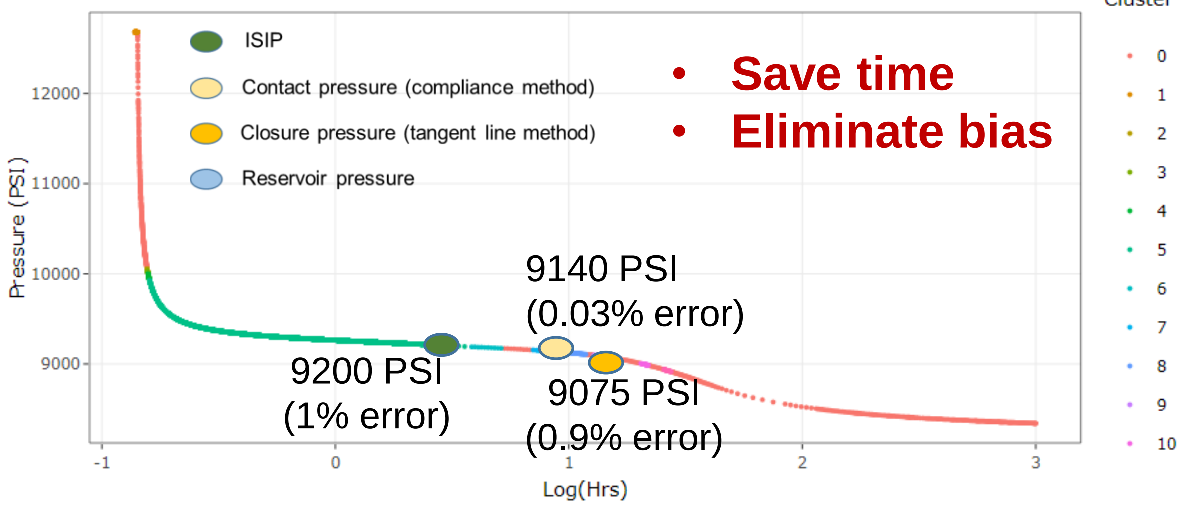


Old method: analytic interpretation, curve fitting derivatives

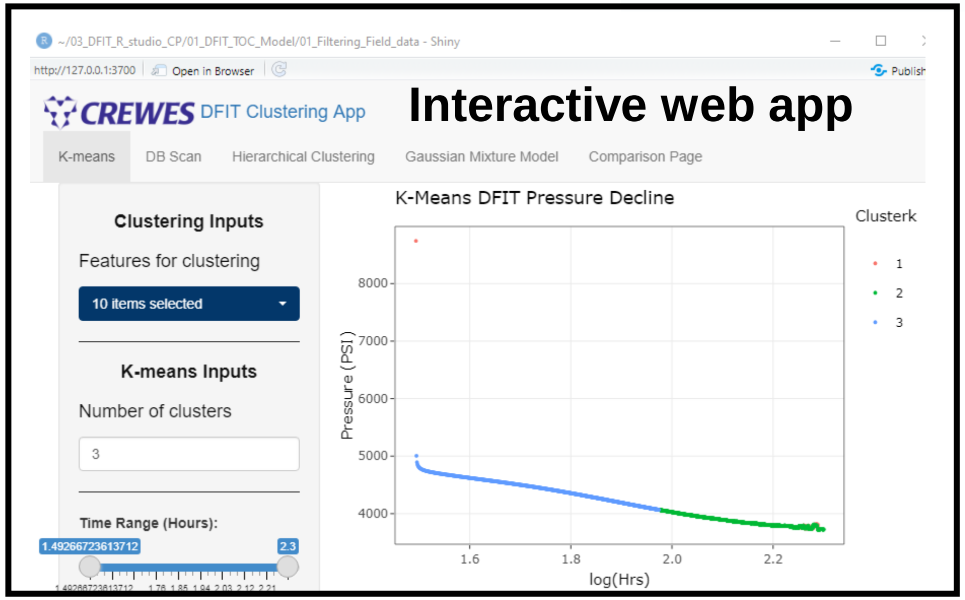
New unsupervised clustering method

Shiny from R Studio

DB Scan DFIT Pressure Decline



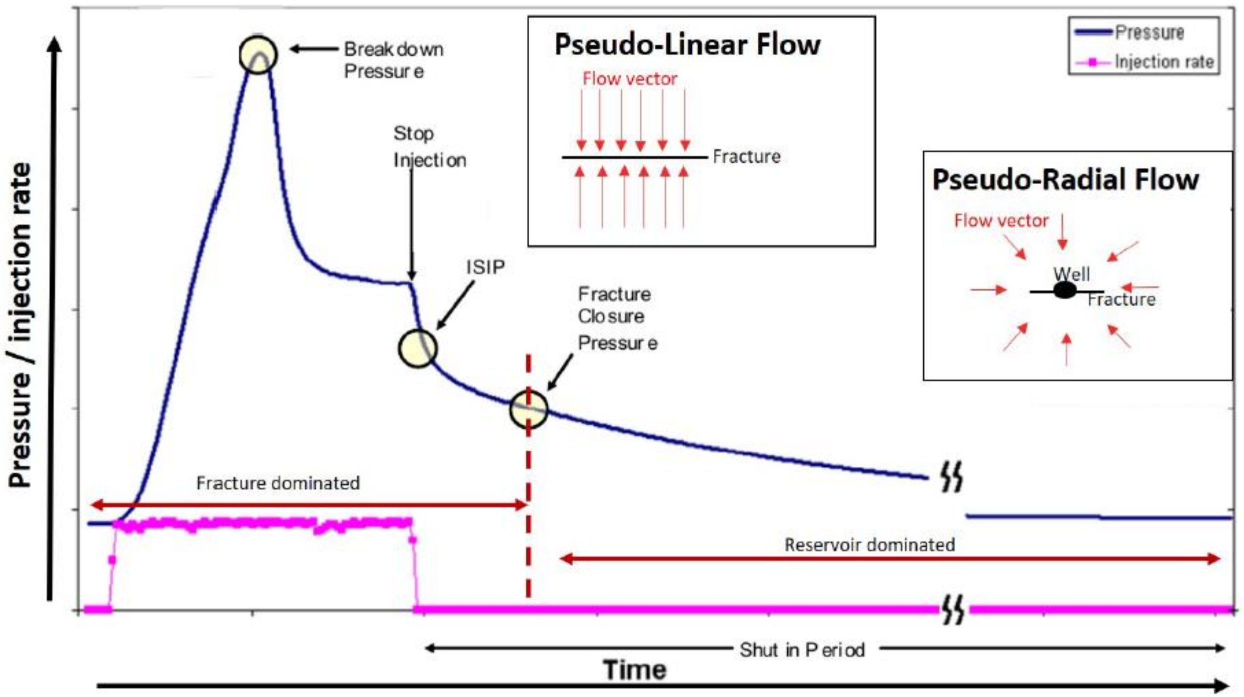
- **Save time**
- **Eliminate bias**



Outline

- ❑Introduction: Why DFIT?
- ❑Current techniques and proposed techniques of interpretation
- ❑CREWES DFIT interpretation app
- ❑App calibration
- ❑Synthetic testing + field testing
- ❑Mathematic inferences
- ❑Conclusions

What is a DFIT? Why is it Important?



Adapted from Cramer and Nguyen (2013)

- **DFIT (Diagnostic fracture injection test)**
- Information in unconventional reservoirs about:
  - Instantaneous shut-in pressure (ISIP)
  - Shmin
  - Reservoir pressure
- Leads to inferences about “fracability” of reservoir, stress regime, economics.

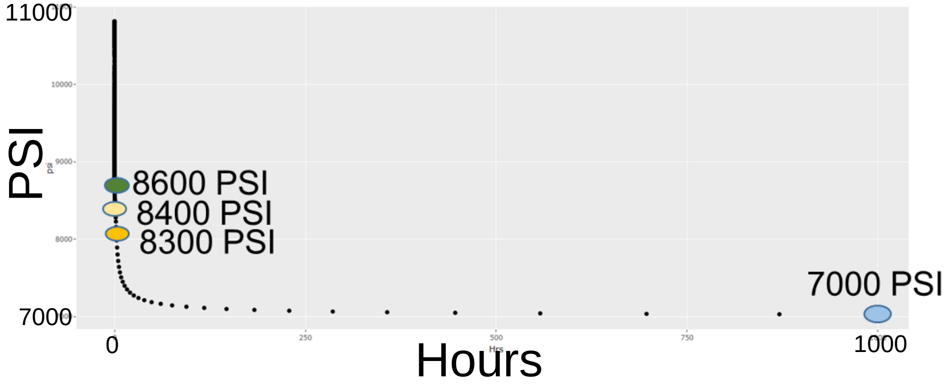
Important geologic information in DFIT, how can we efficiently interpret the data?

Outline

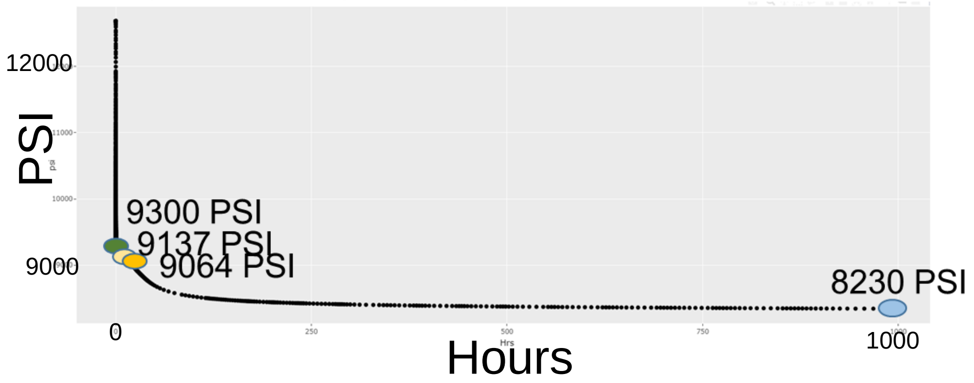
- ❑Introduction: Why DFIT?
- ❑Current techniques and proposed techniques of interpretation
- ❑CREWES DFIT interpretation app
- ❑App calibration
- ❑Synthetic testing + field testing
- ❑Mathematic inferences
- ❑Conclusions

DFIT's : Calibration Datasets From the TOC Group

3- Layer Res frac demo DFIT



31- Layer Duvernay history match DFIT



ISIP

Contact pressure (compliance method)

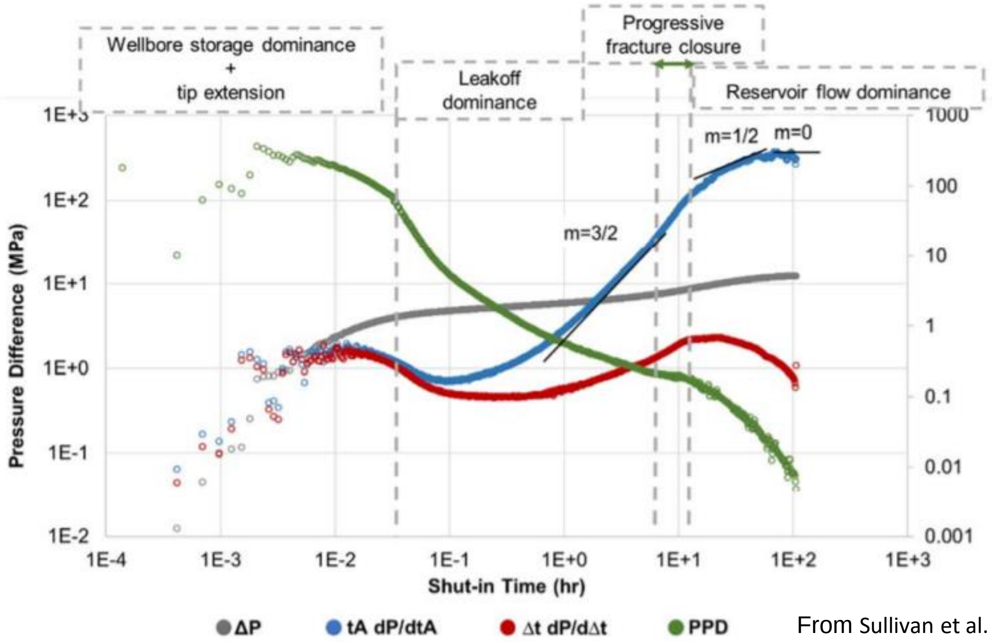
Closure pressure (tangent line method)

Reservoir pressure

Subtle variations hold important geologic information

How are these events picked on a curve with little apparent variation?

Analytic engineering method: time consuming and human bias

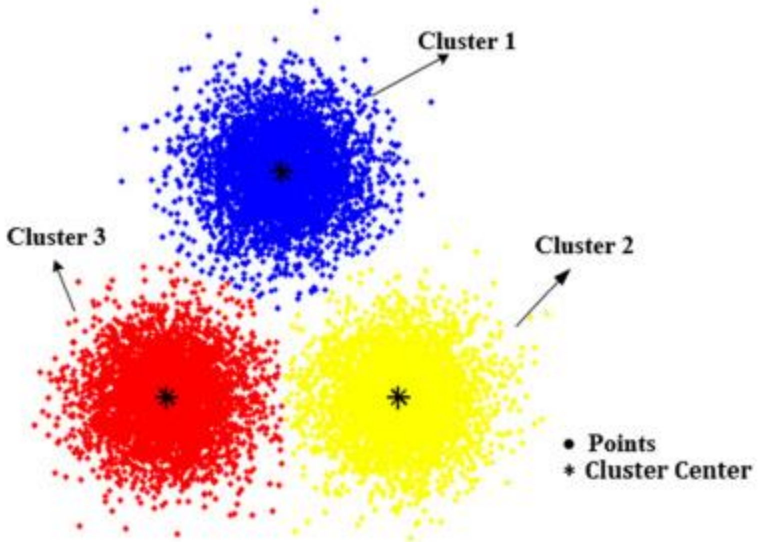


From Sullivan et al. (2019)

Idea: Could a unsupervised clustering method pick up on these variations **save time and eliminate bias?**

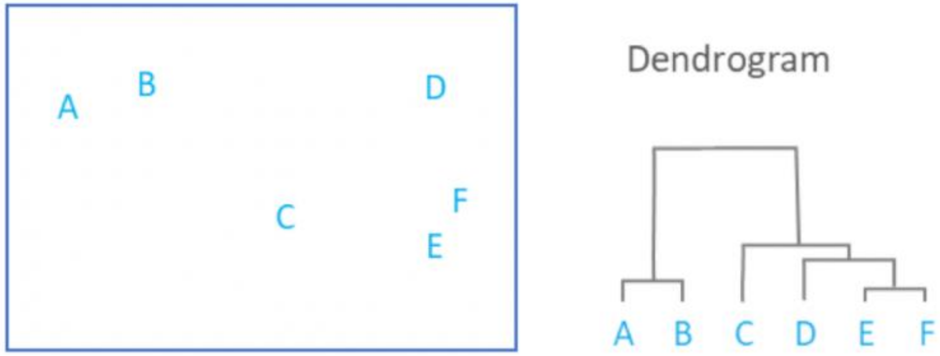
Clustering Methods Evaluated

K-means: centroid based, general method



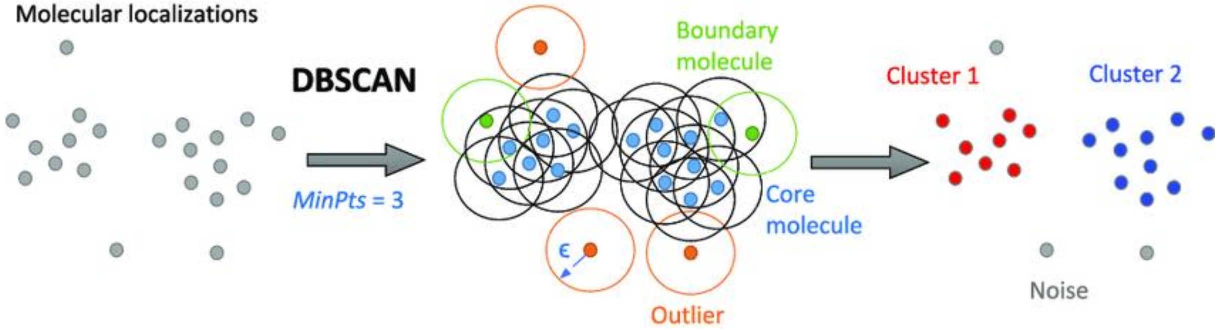
https://www.osapublishing.org/oe/fulltext.cfm?uri=oe-25-22-27570&id=375887

Hierarchical: clustering dendrogram makes easy to determine number of clusters



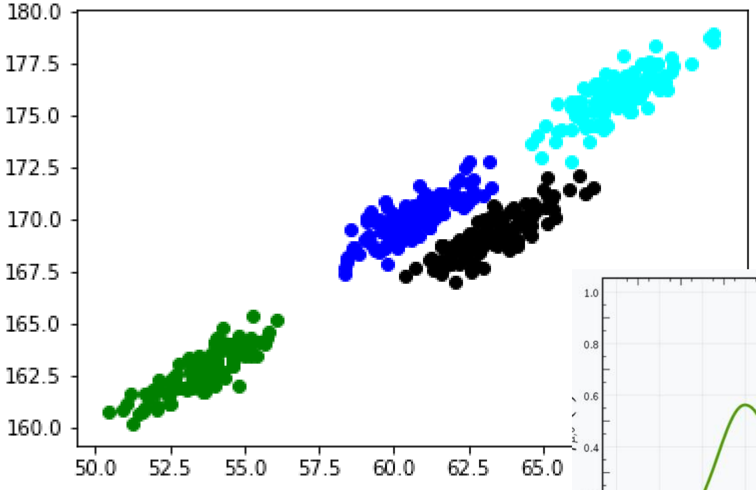
https://www.displayr.com/what-is-hierarchical-clustering/#:~:text=Hierarchical%20clustering%2C%20also%20known%20as,broadly%20similar%20to%20each%20other.

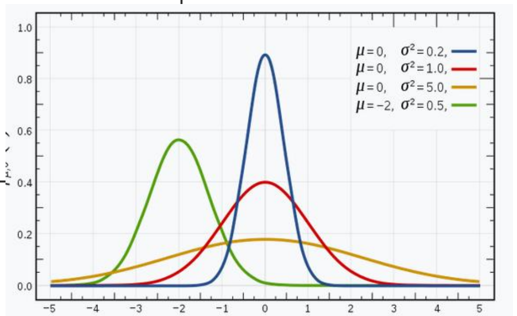
DB scan: density based- handles odd distributions + noise



https://www.researchgate.net/figure/An-Example-Illustrating-the-Density-Based-DBSCAN-Clustering-Method-Applied-to-SMLM-Data_fig4_342141592

Gaussian: probabilistic, set model distribution





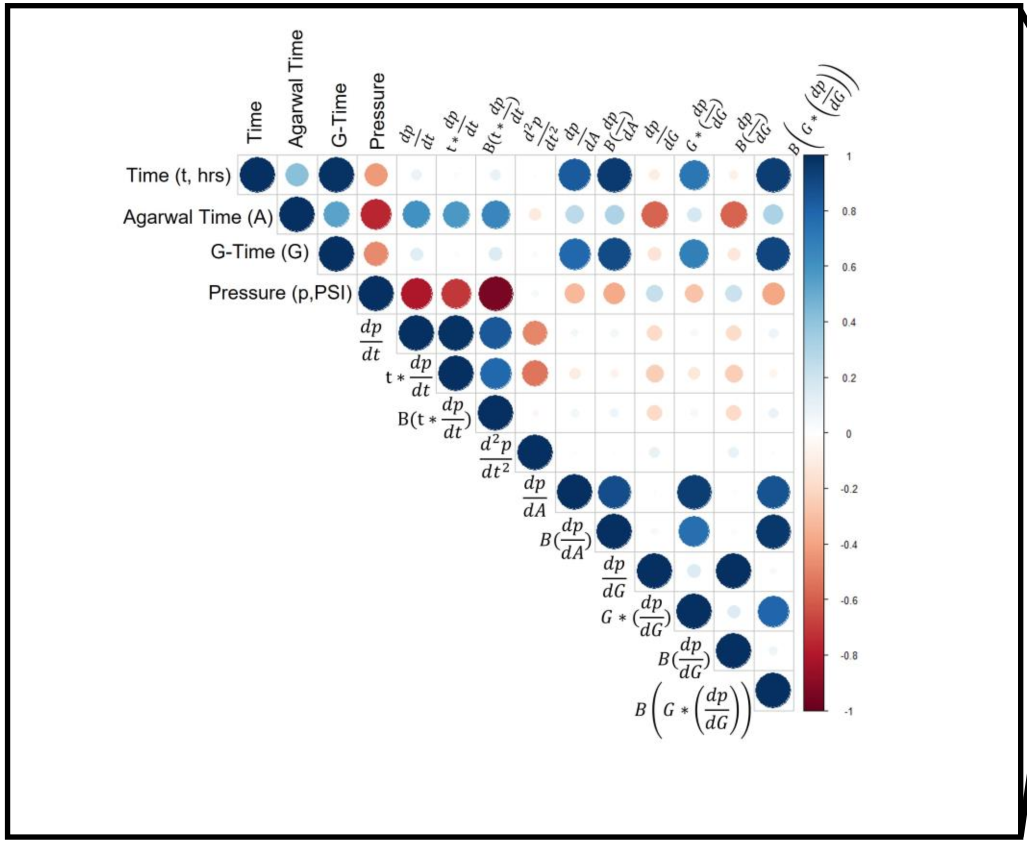
https://www.analyticsvidhya.com/blog/2019/10/gaussian-mixture-models-clustering/

Outline

- ❑Introduction: Why DFIT?
- ❑Current techniques and proposed techniques of interpretation
- ❑CREWES DFIT interpretation app
- ❑App calibration
- ❑Synthetic testing + field testing
- ❑Mathematic inferences
- ❑Conclusions

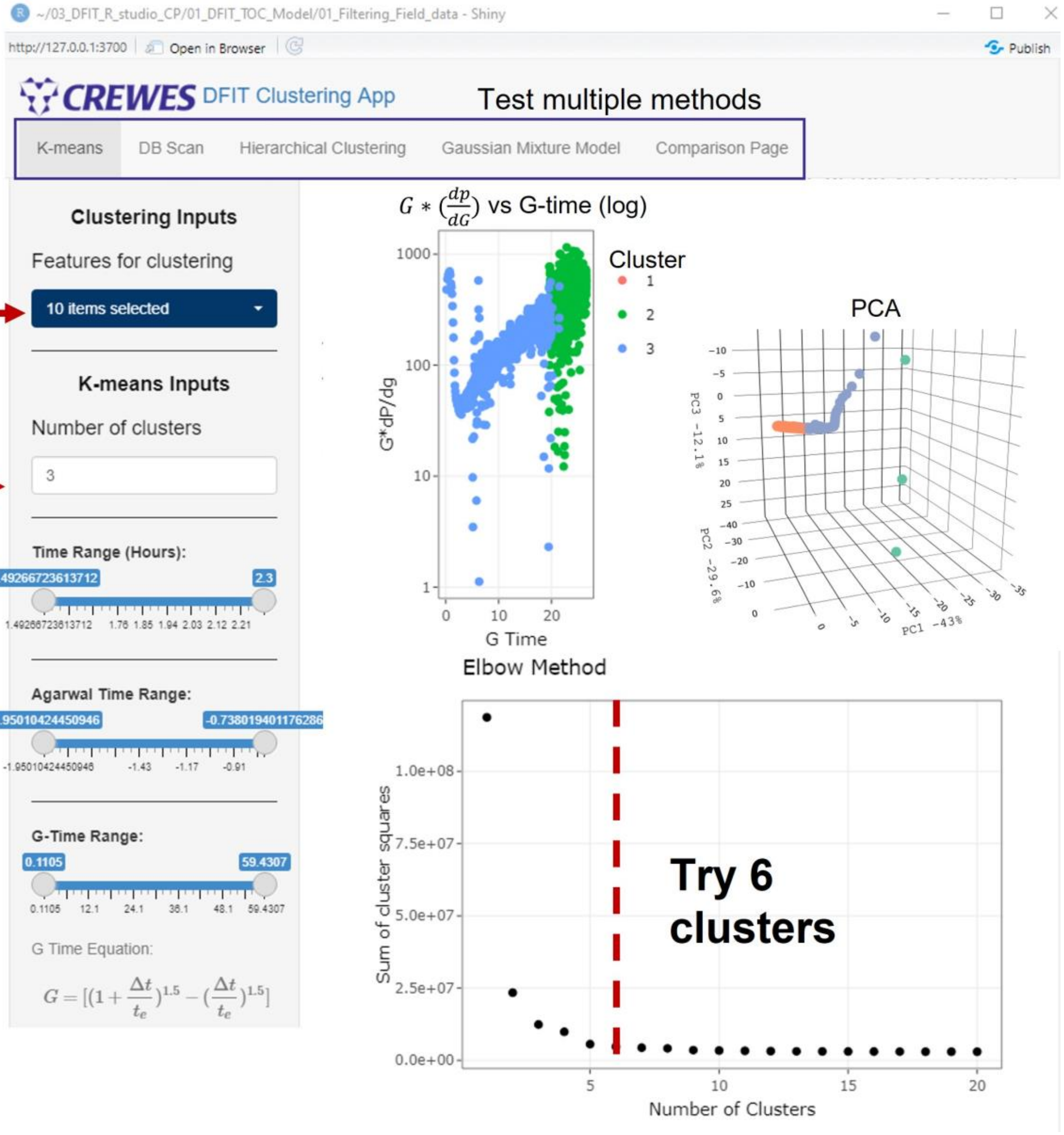
CREWES DFIT Interpretation APP

Complex variable relations: correlation matrix



Selected variables

Manipulate hyper-parameters



High-Level Overview – Unsupervised DFIT Event Detection

DFIT App Inputs

Clustering Methods Evaluated:

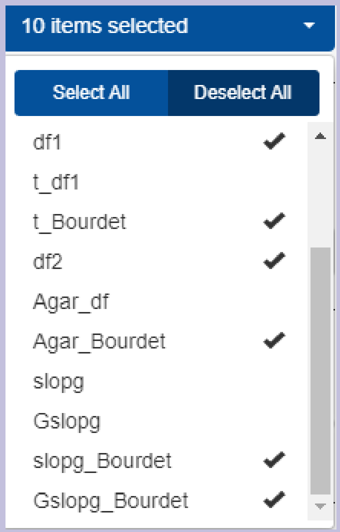
- K-means
- DB-scan
- Hierarchical
- Gaussian mixture

Calibration Datasets
(Known ISIP, Contact, Closure, Reservoir pressures)

- Duvernay single layer model (TOC Group)
- Duvernay multilayer (TOC group)

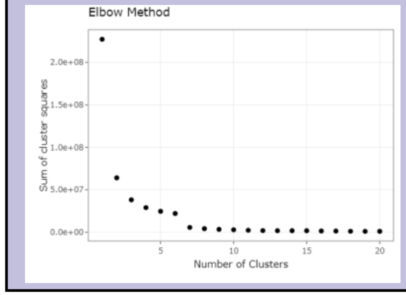
1) Clustering variable Optimization

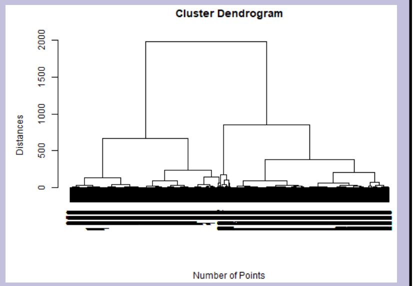
- Scaling vs no scaling
- Additional derivative data (G-time, Agarwal time, first and second order derivatives)



2) Hyperparameter optimization (# of cluster, search radius, etc.)

- Elbow method
- Dendrogram

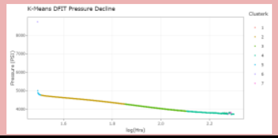




3) Metrics of Evaluation

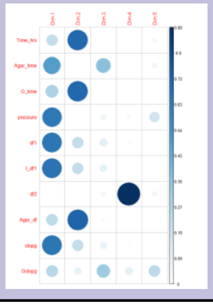
- Percentage of events identified
- Unidentifiable points (no interpretation)
- Repeatability of point

4) Field Data Test



5) Investigation of cluster picking

- Visual inspection
- Principle component analysis (PCA)



Outline

- ❑Introduction: Why DFIT?
- ❑Current techniques and proposed techniques of interpretation
- ❑CREWES DFIT interpretation app
- ❑App calibration
- ❑Synthetic testing + field testing
- ❑Mathematic inferences
- ❑Conclusions

Clustering Variable Optimization

"Simple" pressure and time data creates multiple derivative curves

Hrs	psi
31.16417	8742.817
31.23528	5005.373
31.30639	4894.596
31.37750	4855.438
31.44861	4833.481
31.51972	4818.560
31.59083	4807.476
31.66194	4798.455

\+ Mathematical operation =

Shiny from R Studio

10 items selected

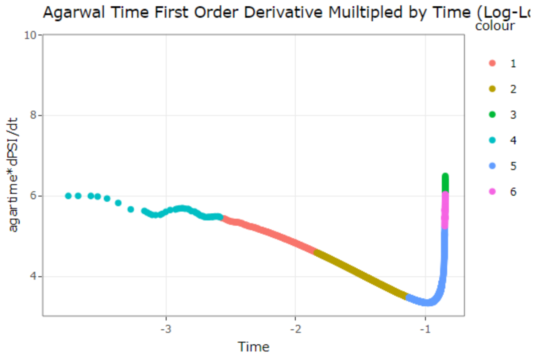
Select All | Deselect All

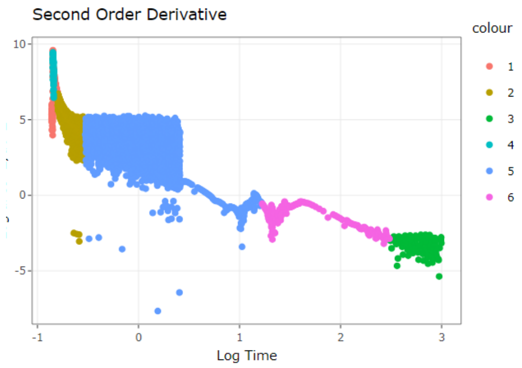
- df1 ✔
- t_df1
- t_Bourdet ✔
- df2 ✔
- Agar_df
- Agar_Bourdet ✔
- slopg
- Gslopg
- slopg_Bourdet ✔
- Gslopg_Bourdet ✔

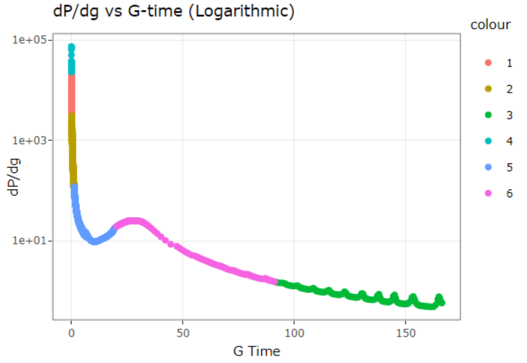
App allows for testing different variable combinations

Scaled app data clustering input

Quick interactive visualization



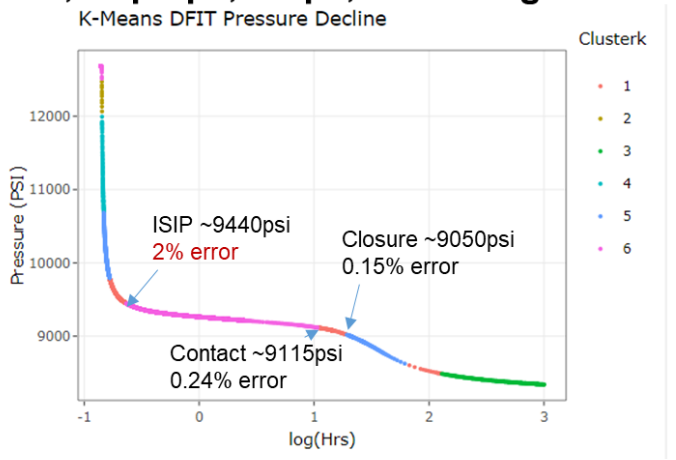




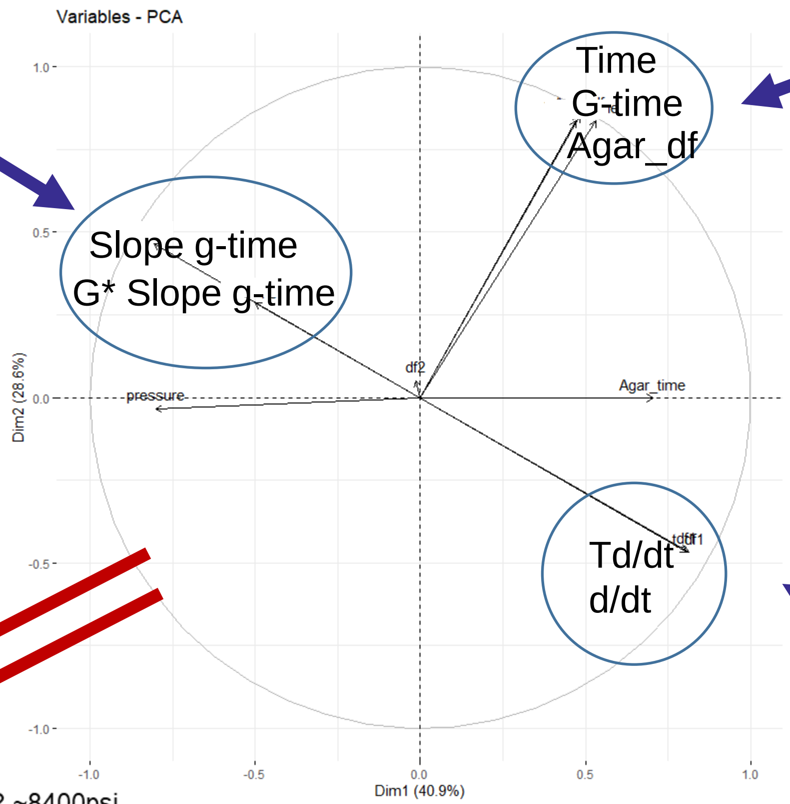
Variable Optimization: PCA Correlation Circle

Early time information variables

Gtime, Gsplope, Slope, clustering -6clusters

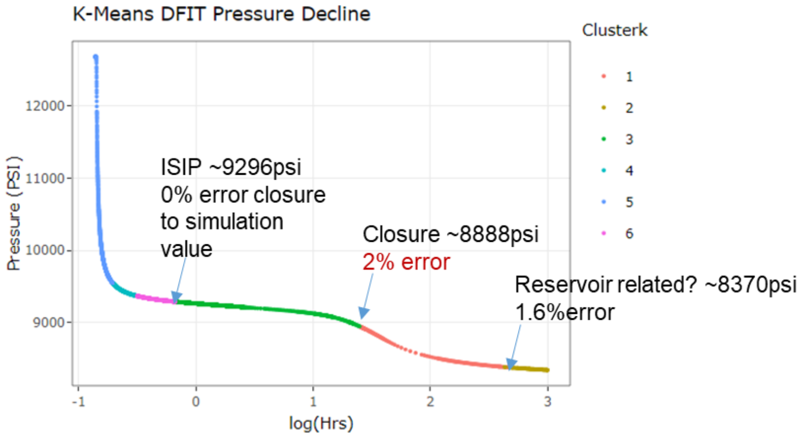


Clusters of correlated variables

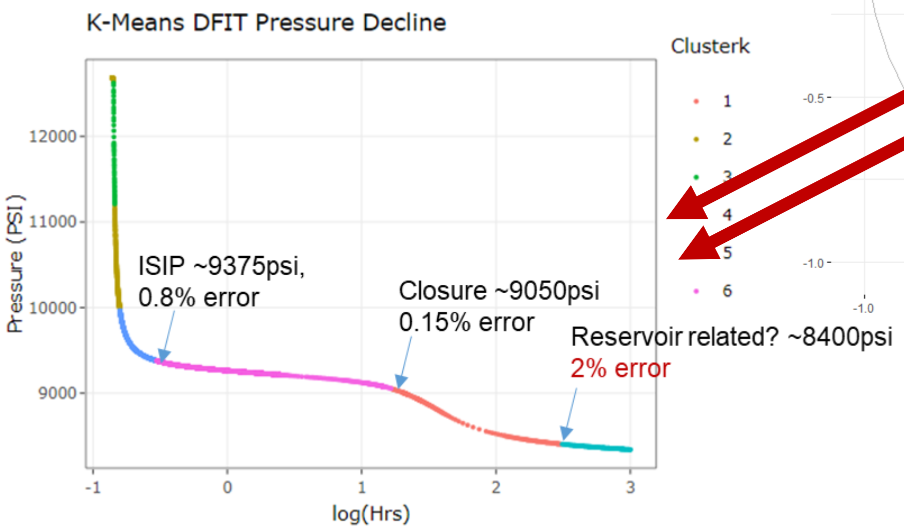


Late/early time information variables

Agar time + derivative – 6 clusters

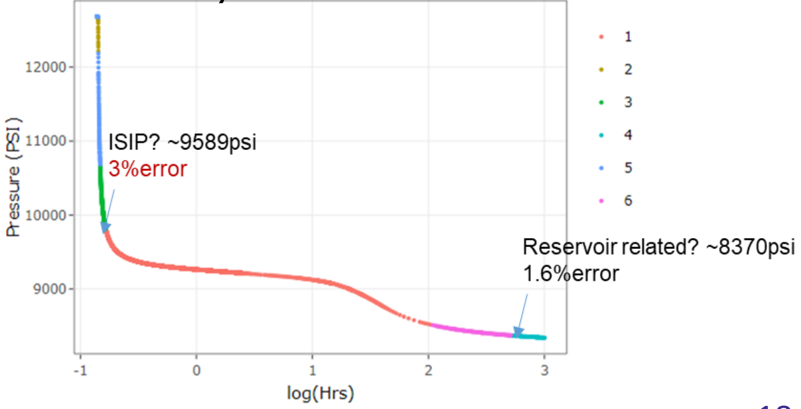


All parameter clustering -6clusters



Late time information variables

Time, pressure, first derivative, t*df1, second df – 6clusters



Collective clustering of all variables gives generalized version of all events

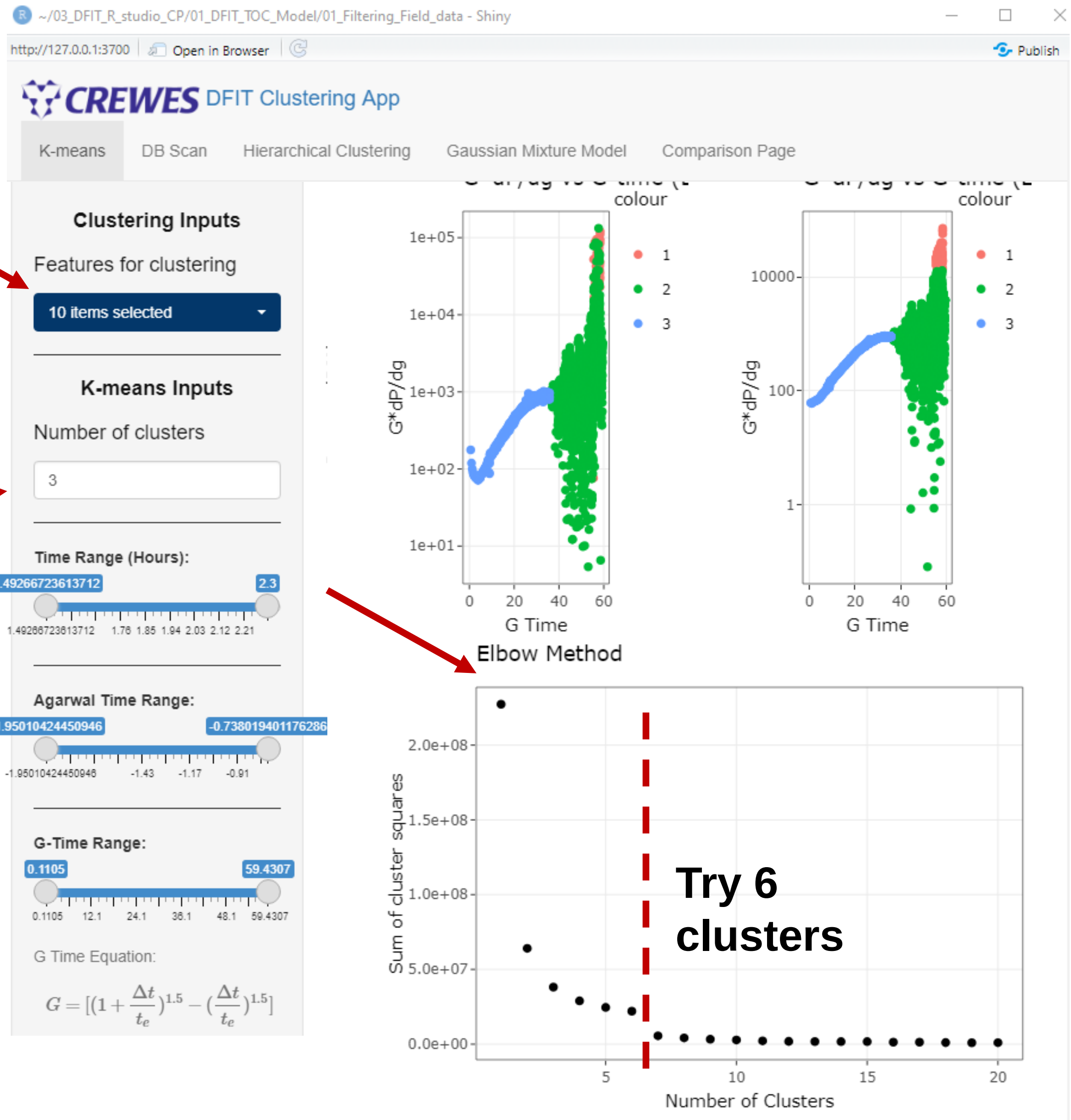
Hyperparameter Optimization

- Implementation of elbow method, iterations through parameters to find best convergence

Cluster Method
DB-Scan 3min, 10radi
DB-Scan 3min, 20radi
DB-Scan 3min, 30radi
DB-Scan 3min, 40radi
DB-Scan 3min, 50radi
DB-Scan 3min, 60radi
DB-Scan 3min, 100radi
DB-Scan 3min, 200radi
DB-Scan 4min, 10radi
DB-Scan 4min, 20radi
DB-Scan 4min, 30radi
DB-Scan 4min, 40radi
DB-Scan 4min, 50radi
DB-Scan 4min, 60radi
DB-Scan 4min, 100radi
DB-Scan 6min, 200radi
**** Optimal performance
DB-Scan 5min, 10radi
DB-Scan 5min, 20radi
DB-Scan 5min, 30radi
DB-Scan 5min, 40radi
DB-Scan 5min, 50radi
DB-Scan 5min, 60radi
DB-Scan 5min, 100radi
DB-Scan 5min, 200radi

Selected variables

Manipulate hyper parameters and investigate

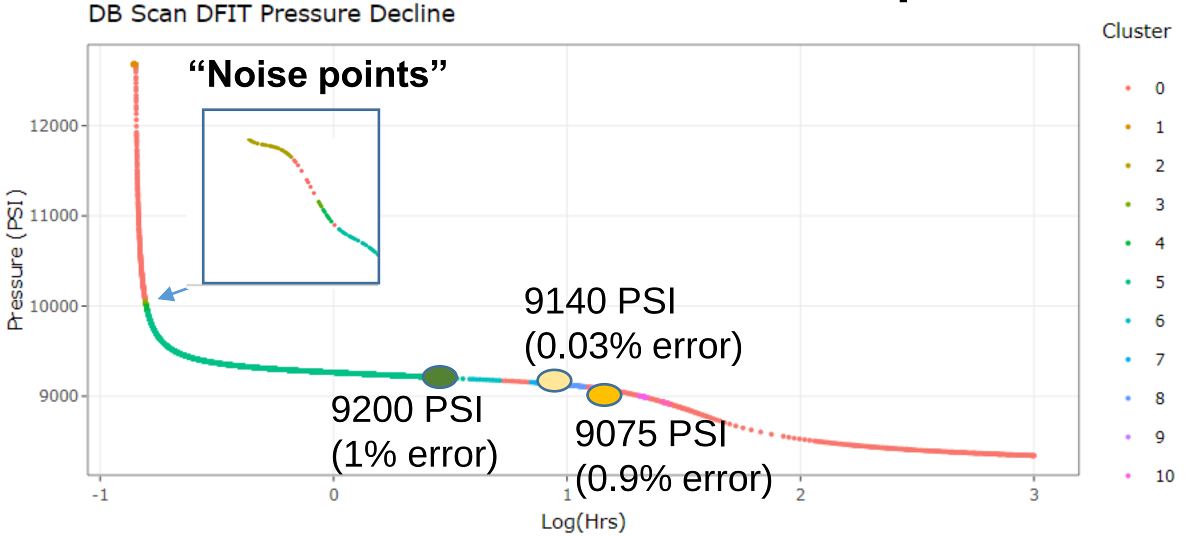


Outline

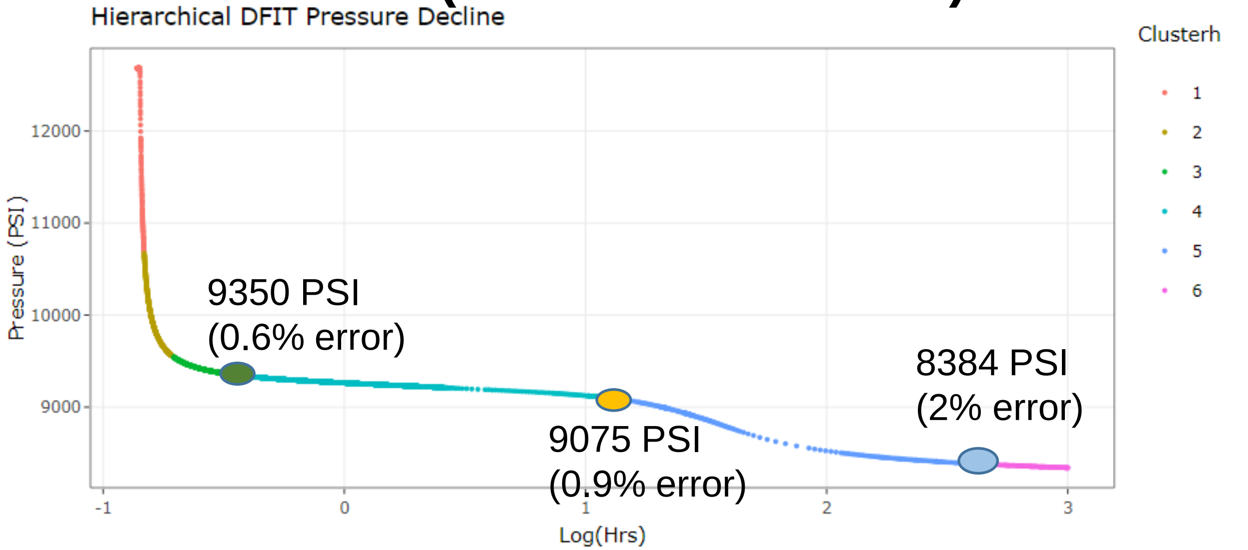
- ❑Introduction: Why DFIT?
- ❑Current techniques and proposed techniques of interpretation
- ❑CREWES DFIT interpretation app
- ❑App calibration
- ❑Synthetic testing + field testing
- ❑Mathematic inferences
- ❑Conclusions

Metrics of Evaluation: Best Clustering Method Ranked (Models)

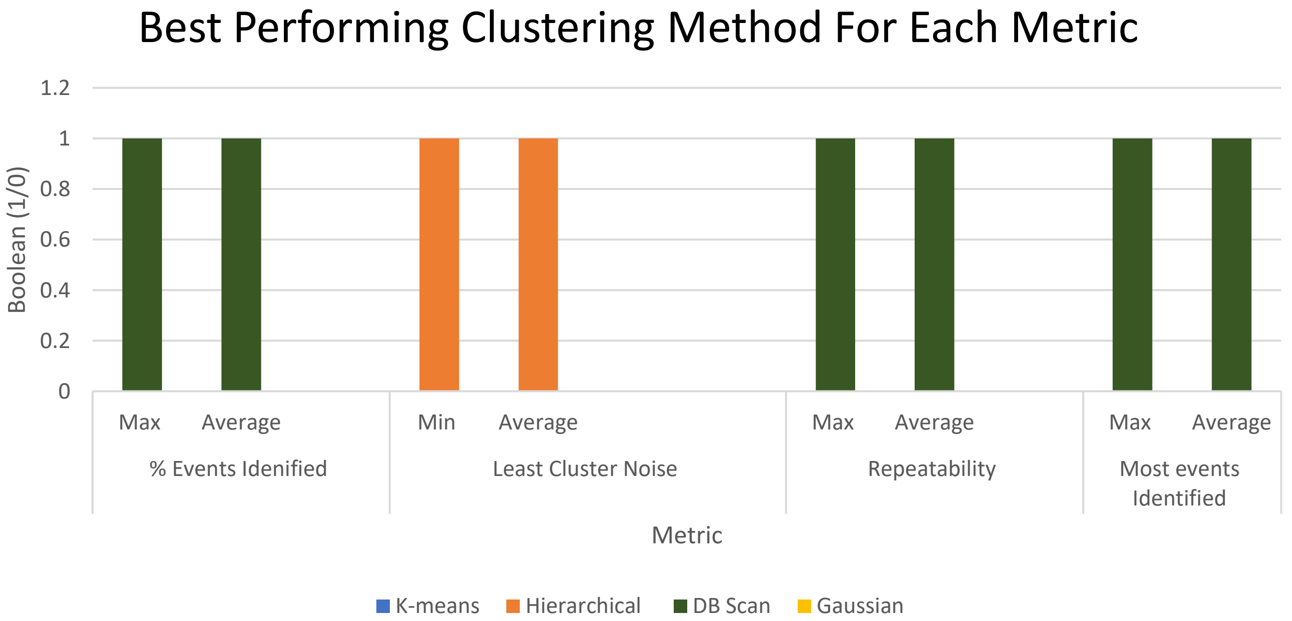
Dbscan: 0.04 sear radii, 6 min points



Hierarchical (6 clusters – Ward D)



- **DB scan appears to be an optimal performer**
- **Hierarchical Clustering produces least noise**



ISIP (9296 PSI)

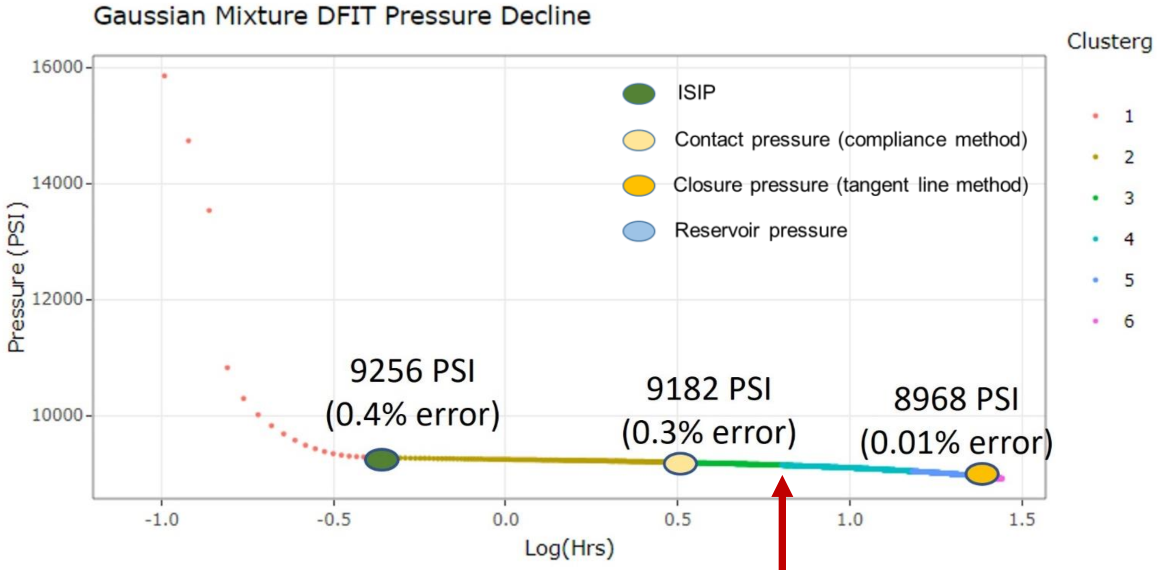
Contact (9137 PSI)

Closure (9064 PSI)

Reservoir (8230 PSI)

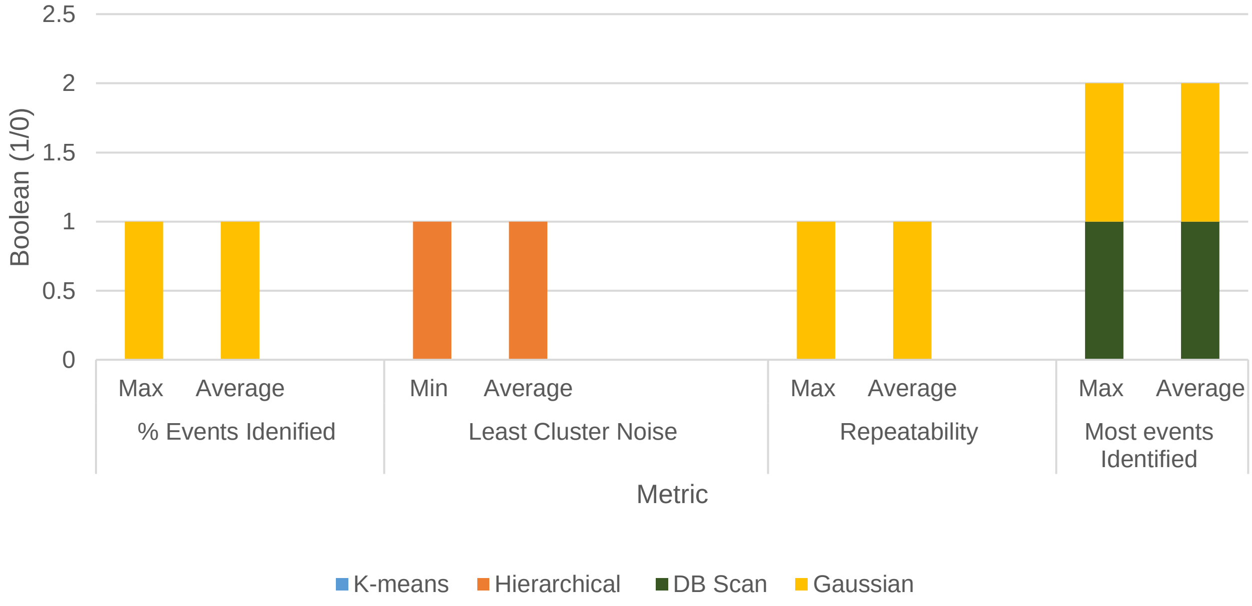
Field Test: Gaussian Method Outperforms DB-scan

Field Test: Gaussian Method



Additional boundaries?

Best Performing Clustering Method For Each Metric



Switch in optimal clustering method

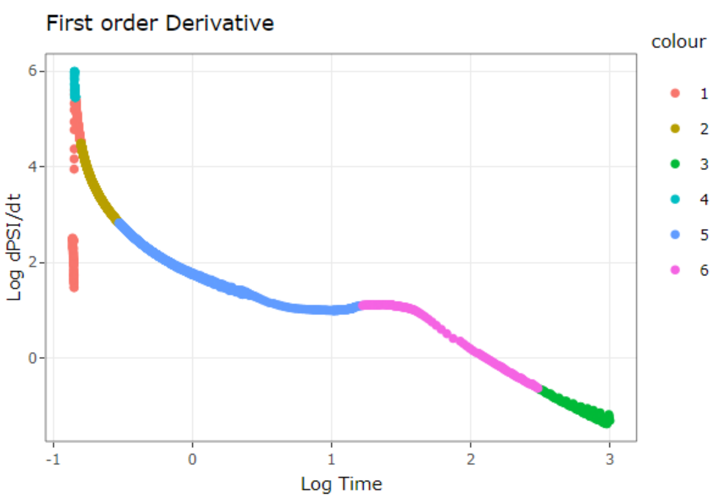
☆ **Gaussian (VEE) -** Ellipsoidal, equal shape and orientation appears to outperform DB scan in this scenario

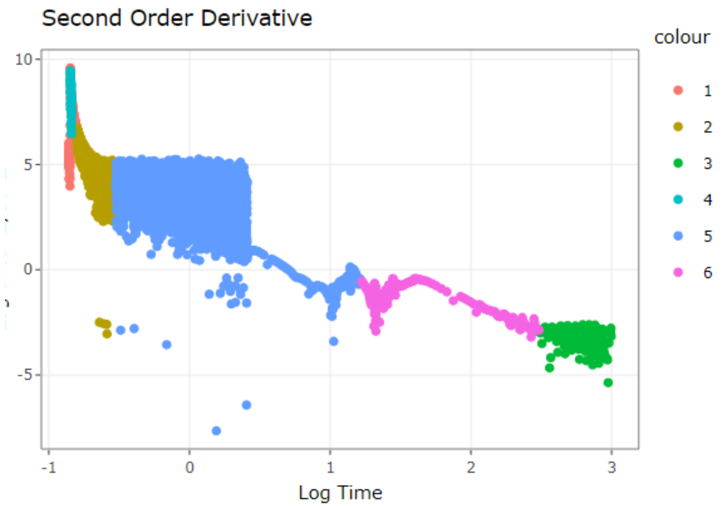
- **Probabilistic nature of Gaussian methods can handle noise?**

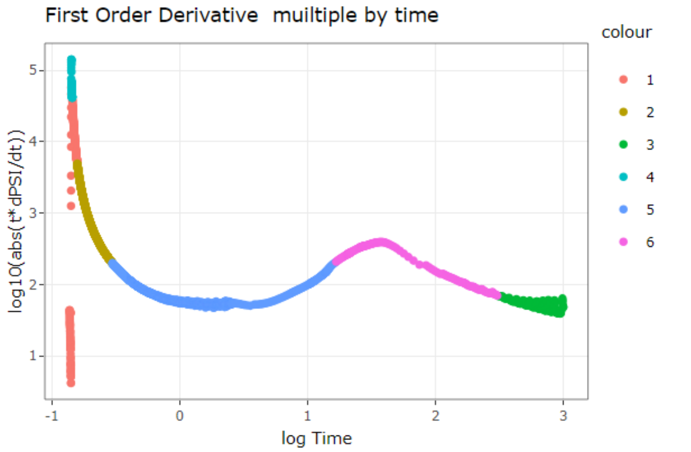
Outline

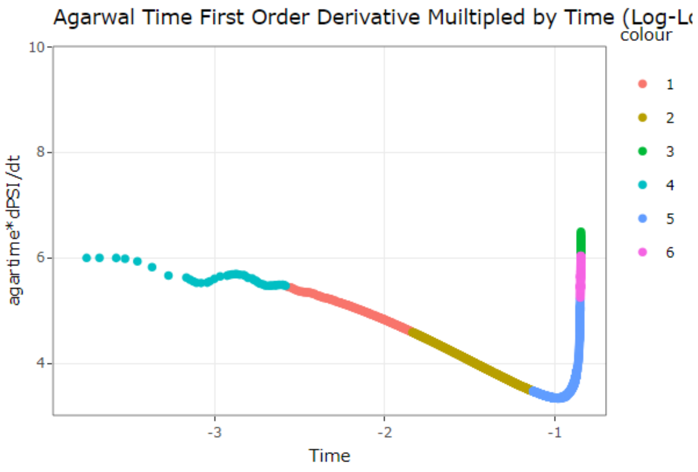
- ❑Introduction: Why DFIT?
- ❑Current techniques and proposed techniques of interpretation
- ❑CREWES DFIT interpretation app
- ❑App calibration
- ❑Synthetic testing + field testing
- ❑Mathematic inferences
- ❑Conclusions

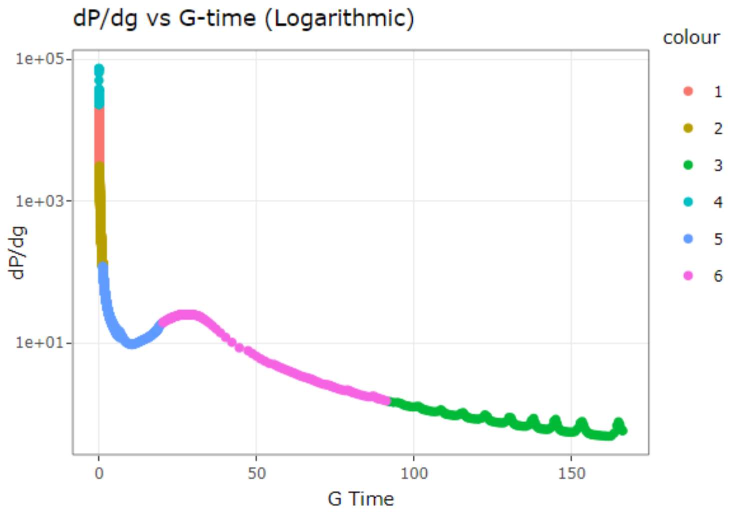
Mathematical Relation - All Parameter Clustering -6clusters K-means

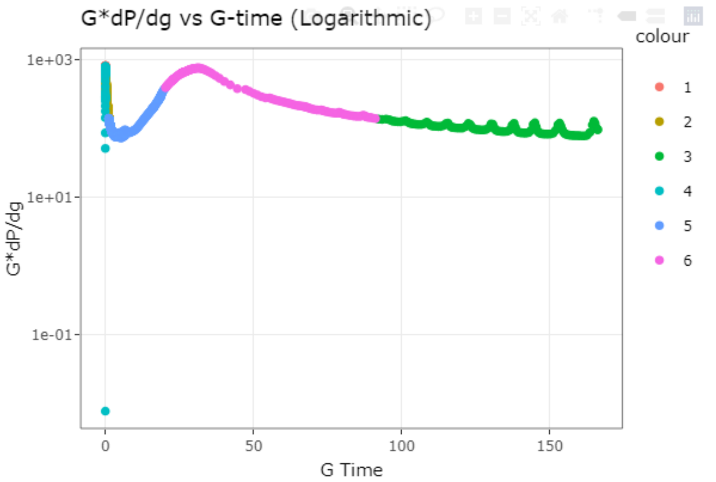






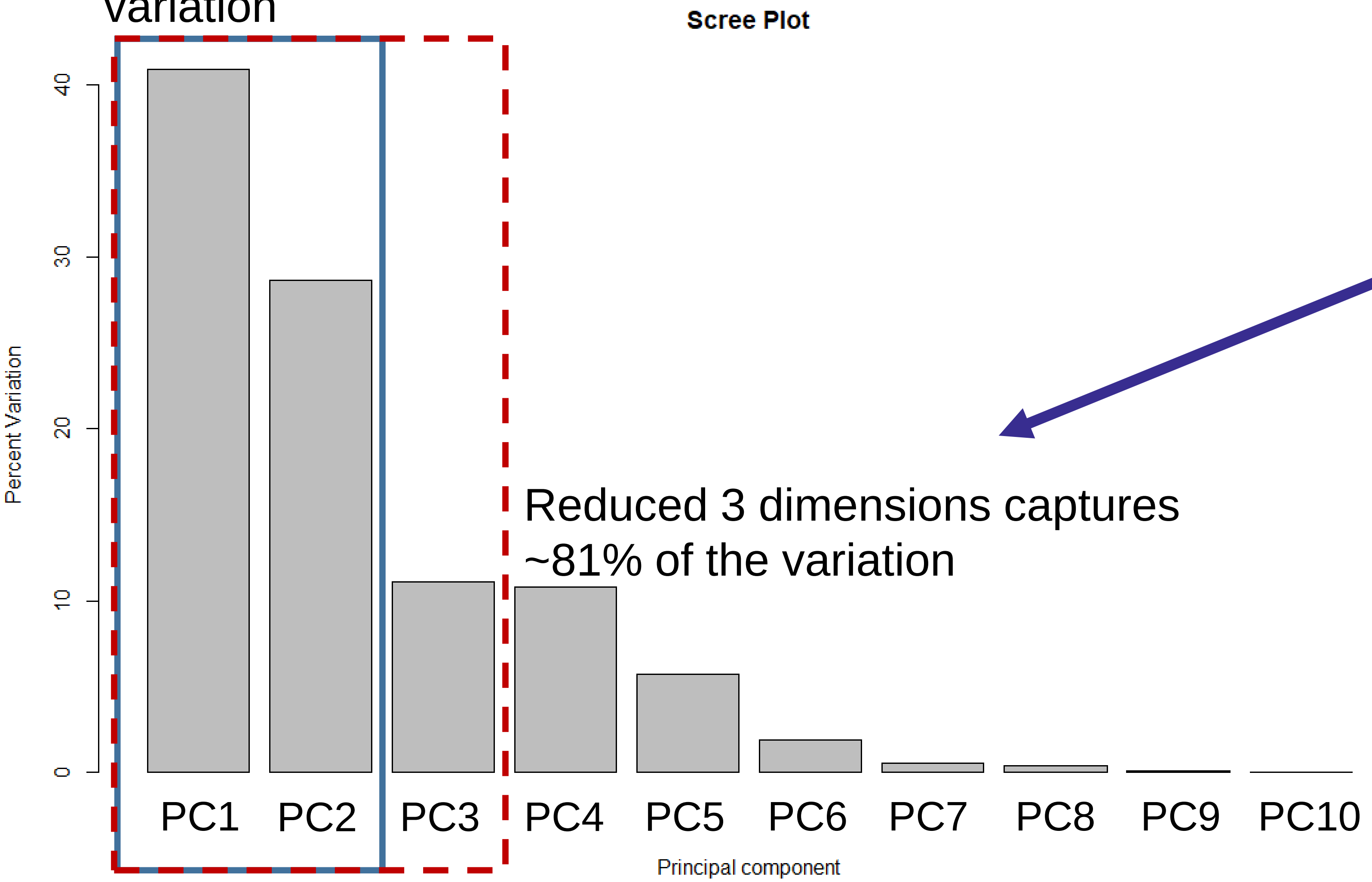






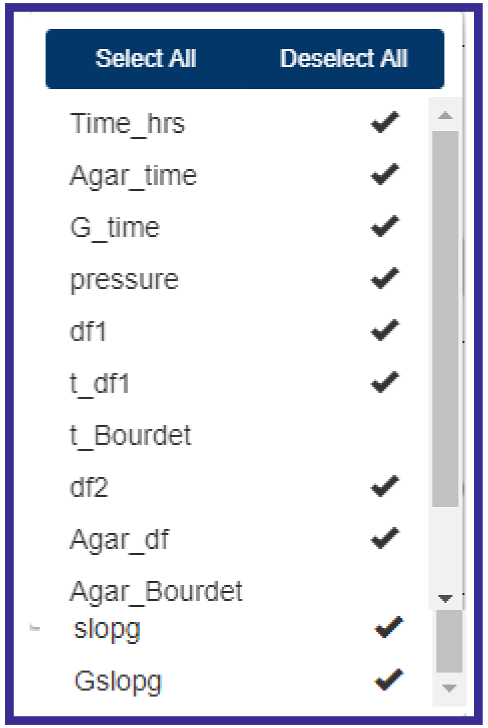
Principle Component Scree Plot – Reducing Dimensions

Reduced 2 dimensions captures ~70% of the variation

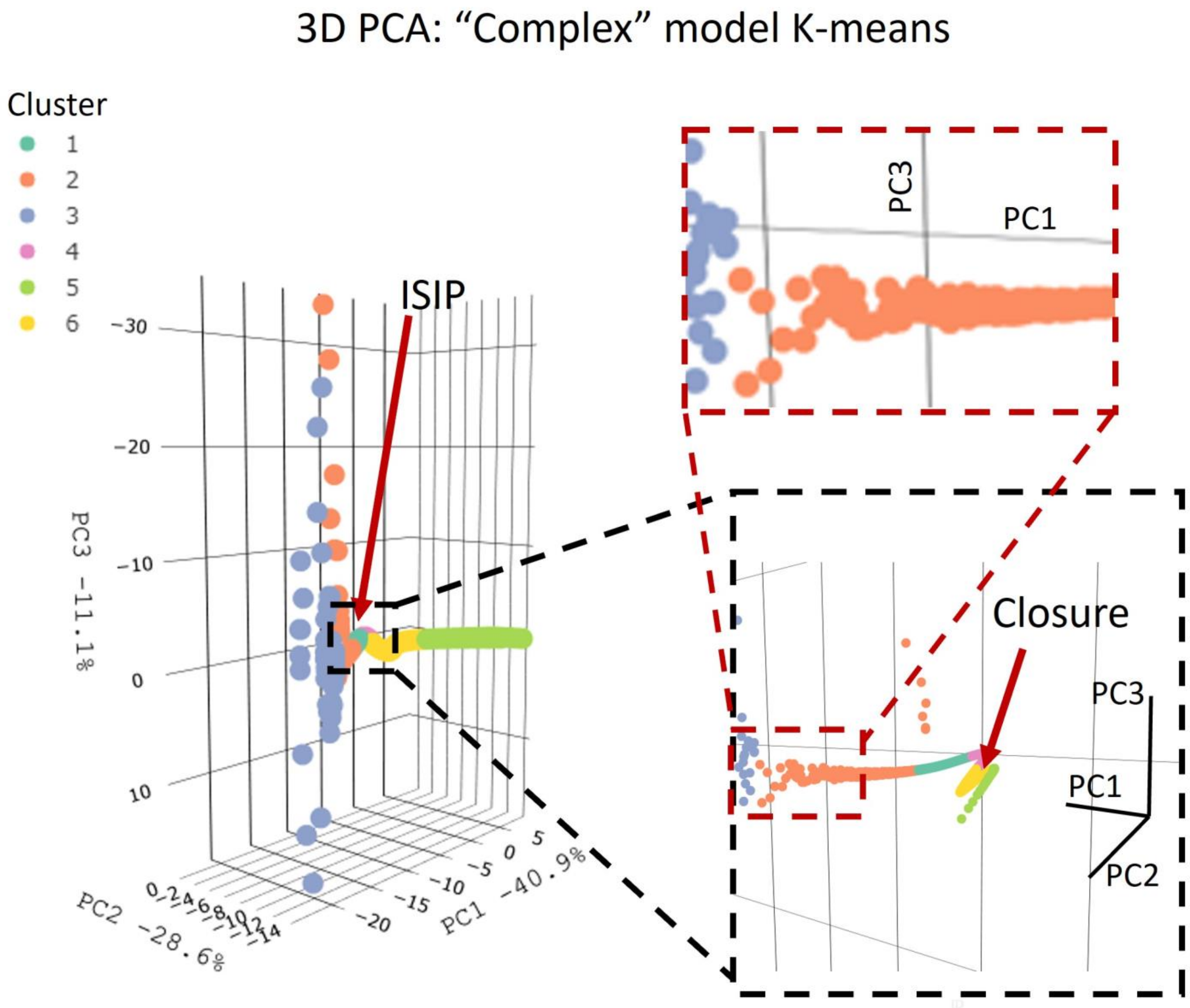


Reduced 3 dimensions captures ~81% of the variation

Input parameters



Understanding the Mathematics:



1) **2D Inflection/change** in slope creating a variation in point density (new cluster)

2) **3D Inflection/change** in slope creating a variation in point density (new cluster)

3) **Change in natural oscillation** of data -> variation in point density

Outline

- ❑Introduction: Why DFIT?
- ❑Current techniques and proposed techniques of interpretation
- ❑CREWES DFIT interpretation app
- ❑App calibration
- ❑Synthetic testing + field testing
- ❑Mathematic inferences
- ❑Conclusions

Conclusions and Ideas:

- Shiny apps provide **a quick method** of optimizing and **visualizing machine** learning applications

- ~**6 clusters appear to capture data variation** for K-means, DB-scan and Hierarchical methods

- **Incorporation of derivative curves** and scaling appears to help algorithms pick more events (although more accurate localized results maybe achieved by using fewer curves).

- **DB scan appears to have a best performance on modeled data**, however, when introduced to **field data, the Gaussian mixture model** appears to outperform

  - Apply a Fast Fourier Transform to data to filter it?

- **Cluster boundary** related to a combination of **natural oscillation, inflection, maxima/ minima** creating variation in point density and therefore cluster position.

Thank You

- We thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19.

- NSERC Alexander Graham Bell Canada Graduate Scholarship (CGS-M)

- Canada First Research Excellence Fund (CFREF) Theme 2 Global Research Initiative in Sustainable Low Carbon Unconventional Resources (GRI)

- NSERC Alliance grant

- Mark McClure (president of ResFrac) for providing academic licenses of ResFrac for use in this study

- Colleagues from CREWES and TOC groups

References

Sullivan, M., B. Zanganeh, A. Springer, and C. Clarkson, 2019, Post-Fracture Pressure Decay: A Novel (and Free) Stage-Level Assessment Method: SPE/AAPG/SEG Unconventional Resources Technology Conference, Denver, Colorado, USA, July 2019, DOI:10.15530/urtec-2019-970.

Cramer, D.D., and D.H., Nguyen, 2013, Diagnostic Fracture Injection Testing Tactics in Unconventional Reservoirs: SPE Hydraulic Fracturing Technology Conference, DOI:10.2118/163863-MS.