A Comparison of 35 gas-liquid point models

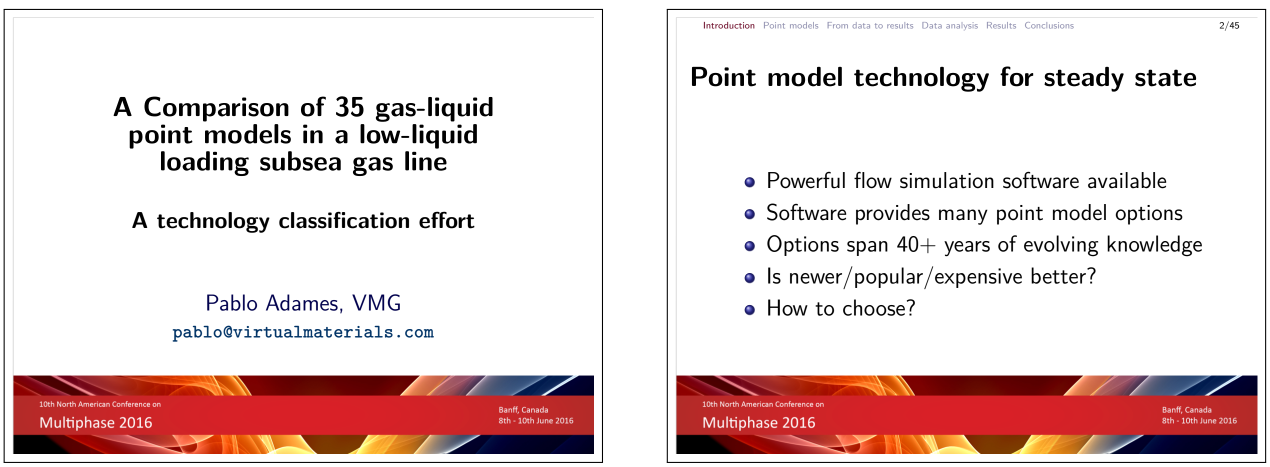
A technology classification effort

Pablo Adames

June 08, 2016

1 Introduction

This document was created to contain both the annotations and the actual slides for the presentation of the paper of the given title at the 10th North American Conference on Multiphase Production in Banff, Canada.



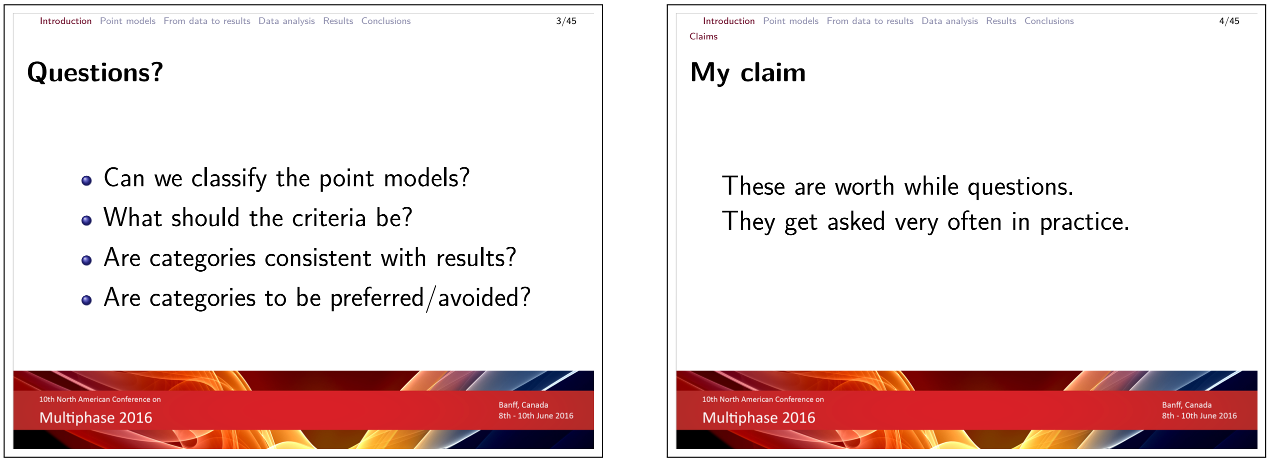
(a) ***Title page*** **(b)** ***Context for this work***

Figure 1: Introductory and context slides

Slide 1b illustrates the context for this work. The main reason was to answer the question of which model to choose when there is a rich list of alternatives using flow simulation software. The many implementations available now range from the very first ones from the 60's to the ones that come from the latest research and software practices.

Slide 2a shows the central question of how to classify the point model options in software that use marching algorithms.

Answering these questions may lead to less uncertainty, lower costs, and safer designs. It may also bring about a deeper understanding of the underlying modelling technology for the software user.

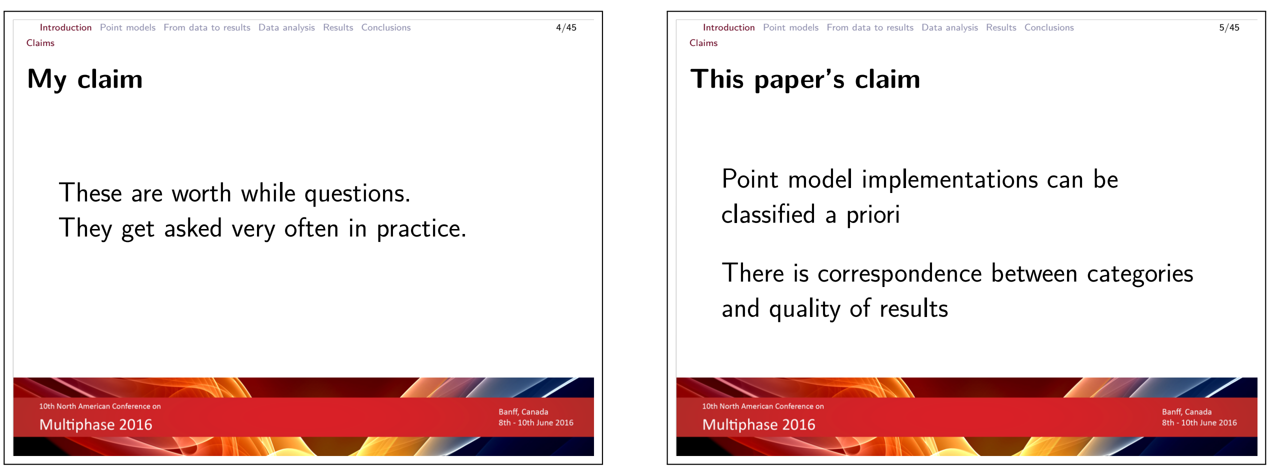


(a) ***Questions*** **(b)** ***Personal claim***

Figure 2: Motivation for this investigation

1.1 Claims

My personal experience is reflected in slide 3a. Any software vendor will get asked this question at least once a month if not once a week. Actually in my experience when users stop asking about these options they have quietly settled for an empirical rule of their own or a single proprietary option. The latter is a very valid choice in daily work but one that limits understanding and thus exposes the user to unheeded use of the technology beyond its limitations and uncertainties. Even using proprietary technology one should strive to understand its driving principles.



(a) ***Personal claim*** **(b)** ***Claim of this paper***

Figure 3: Justification and goal

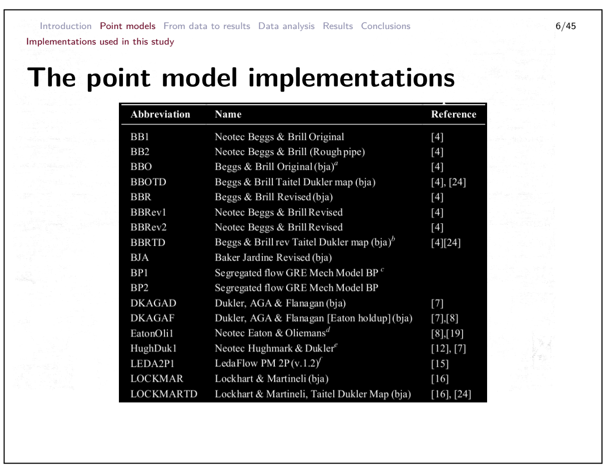
Slide 3b states the achieved goal of this paper. Point models have been classified before by previous authors (see full paper for references). This paper uses a new theoretical classification based on four separate modelling criteria. The pressure drop results were classified with modern data analysis algorithms using an industry standard data processing computer language and the correspondence with the a priori categories was found statistically meaningful.

2 Point models

The context for these point models is the calculation of the total pressure change and liquid accumulation in a pipe line. This is traditionally done by descretizing it into shorter segments. The algorithm for solving all the segments *marches* along solving each one separately. The point model contributes the hydrodynamic gas-liquid volumetric distribution due to slippage and the pressure change in the segment. Other sub-models may deal with heat transfer, phase behaviour, and the transport properties of each phase. Together they make up the pipe simulator technology. We shall be concerned here exclusively with point model technology.

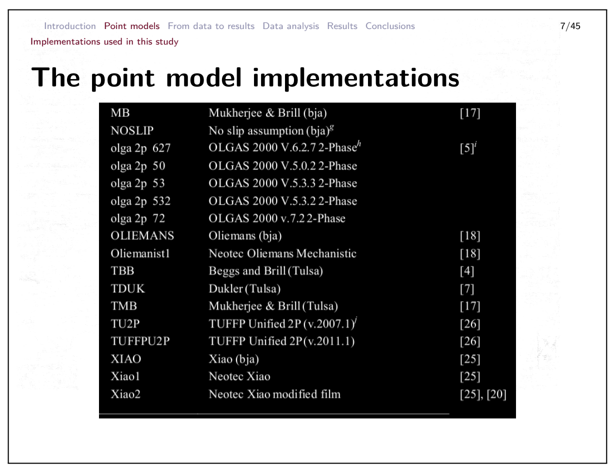
2.1 Implementations used in this study

To classify point model technology one requires to have a fair number of software implementations available. A total of 35 gas-liquid different implementations were available and enumerated in slides 4a and 4b.



Abbreviation	Name	Reference
BB1	Neotec Beggs & Brill Original	[4]
BB2	Neotec Beggs & Brill (Rough pipe)	[4]
BBO	Beggs & Brill Original (bja)[a]	[4]
BBOTD	Beggs & Brill Taitel Dukler map (bja)	[4], [24]
BBR	Beggs & Brill Revised (bja)	[4]
BBRev1	Neotec Beggs & Brill Revised	[4]
BBRev2	Neotec Beggs & Brill Revised	[4]
BBRTD	Beggs & Brill rev Taitel Dukler map (bja)[b]	[4][24]
BJA	Baker Jardine Revised (bja)	
BP1	Segregated flow GRE Mech Model BP [c]	
BP2	Segregated flow GRE Mech Model BP	
DKAGAD	Dukler, AGA & Flanagan (bja)	[7]
DKAGAF	Dukler, AGA & Flanagan [Eaton holdup] (bja)	[7],[8]
EatonOli1	Neotec Eaton & Oliemans[d]	[8],[19]
HughDuk1	Neotec Hughmark & Dukler[e]	[12], [7]
LEDA2P1	LedaFlow PM 2P (v.1.2)[f]	[15]
LOCKMAR	Lockhart & Martineli (bja)	[16]
LOCKMARTD	Lockhart & Martineli, Taitel Dukler Map (bja)	[16], [24]

(a) ***Point model list, part 1***



MB	Mukherjee & Brill (bja)	[17]
NOSLIP	No slip assumption (bja)[g]	
olga 2p 627	OLGAS 2000 V.6.2.7 2-Phase[h]	[5][i]
olga 2p 50	OLGAS 2000 V.5.0.2 2-Phase	
olga 2p 53	OLGAS 2000 V.5.3.3 2-Phase	
olga 2p 532	OLGAS 2000 V.5.3.2 2-Phase	
olga 2p 72	OLGAS 2000 v.7.2 2-Phase	
OLIEMANS	Oliemans (bja)	[18]
Oliemanist1	Neotec Oliemans Mechanistic	[18]
TBB	Beggs and Brill (Tulsa)	[4]
TDUK	Dukler (Tulsa)	[7]
TMB	Mukherjee & Brill (Tulsa)	[17]
TU2P	TUFFP Unified 2P (v.2007.1)[j]	[26]
TUFFPU2P	TUFFP Unified 2P (v.2011.1)	[26]
XIAO	Xiao (bja)	[25]
Xiao1	Neotec Xiao	[25]
Xiao2	Neotec Xiao modified film	[25], [20]

(b) ***List part 2***

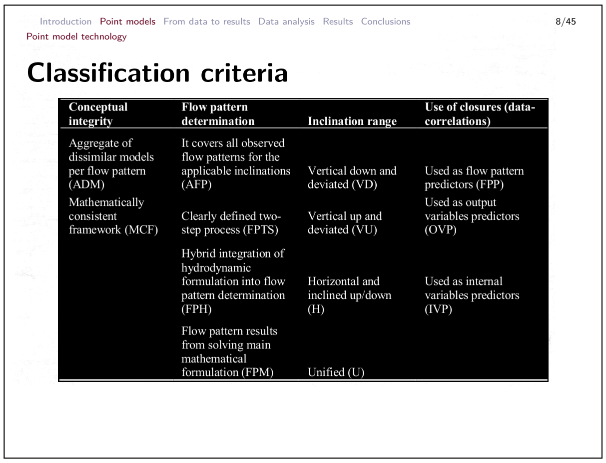
Figure 4: List of 35 point model implementations available for this study. For more detail look at the original paper or the accompanying slides.

If one maintains every other option constant in a flow simulator while varying the point models it is possible to cover the response to a wide rage of technology. The breath of technology depends not only on the number of implementations but also on the mathematical formulations that they represent. Some of them are different interpretations of the same formulation (like B&B) while some others are just evolving versions of basically the same formulation (like OLGAS).

2.2 Point model technology

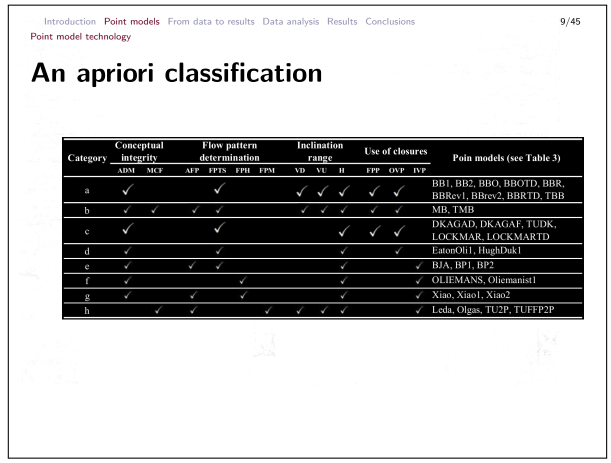
In order to shed some light of the kinds of point models available researchers have traditionally used the term empirical correlation and mechanistic model. A broader criteria appears in slide 5a. The traditional application of empirical correlations comes from their use of pure data correlations as a means of computing output variables like the type of flow pattern, the total pressure gradient, or the liquid holdup.

In this table three other criteria are used and will be mentioned in detail in the following slides.



Conceptual integrity	Flow pattern determination	Inclination range	Use of closures (data-correlations)
Aggregate of dissimilar models per flow pattern (ADM)	It covers all observed flow patterns for the applicable inclinations (AFP)	Vertical down and deviated (VD)	Used as flow pattern predictors (FPP)
Mathematically consistent framework (MCF)	Clearly defined two-step process (FPTS)	Vertical up and deviated (VU)	Used as output variables predictors (OVP)
	Hybrid integration of hydrodynamic formulation into flow pattern determination (FPH)	Horizontal and inclined up/down (H)	Used as internal variables predictors (IVP)
	Flow pattern results from solving main mathematical formulation (FPM)	Unified (U)	

(a) ***Technological aspects considered***



Category	Conceptual integrity		Flow pattern determination				Inclination range			Use of closures			Poin models (see Table 3)
	ADM	MCF	AFP	FPTS	FPH	FPM	VD	VU	H	FPP	OVP	IVP	
a	✓			✓			✓	✓	✓	✓	✓		BB1, BB2, BBO, BBOTD, BBR, BBRev1, BBrev2, BBRTD, TBB
b	✓	✓	✓	✓			✓	✓	✓	✓	✓		MB, TMB
c	✓			✓					✓	✓	✓		DKAGAD, DKAGAF, TUDK, LOCKMAR, LOCKMARTD
d	✓			✓					✓		✓		EatonOli1, HughDuk1
e	✓		✓	✓					✓			✓	BJA, BP1, BP2
f	✓				✓				✓			✓	OLIEMANS, Oliemanist1
g	✓		✓		✓				✓			✓	Xiao, Xiao1, Xiao2
h		✓	✓			✓	✓	✓	✓			✓	Leda, Olgas, TU2P, TUFFP2P

(b) ***Theoretical categories***

Figure 5: An analysis of point model technology

2.3 A priori point model groups

Slide 5b shows an *a priori* classification of the 35 point models used in this study following common characteristics from slide 5a. Later in this presentation it will be shown that the classification algorithms generate groups of very similar composition as the ones shown in this slide, thus stating the claim of slide3b.

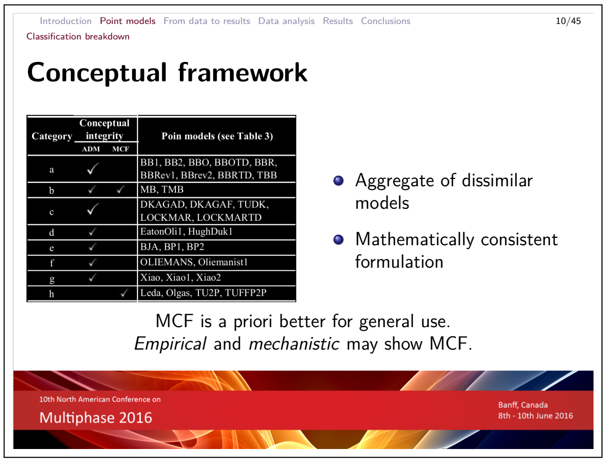
2.4 Classification breakdown

Following is a more detailed presentation of each of the four main columns of the table shown on slide 5a.

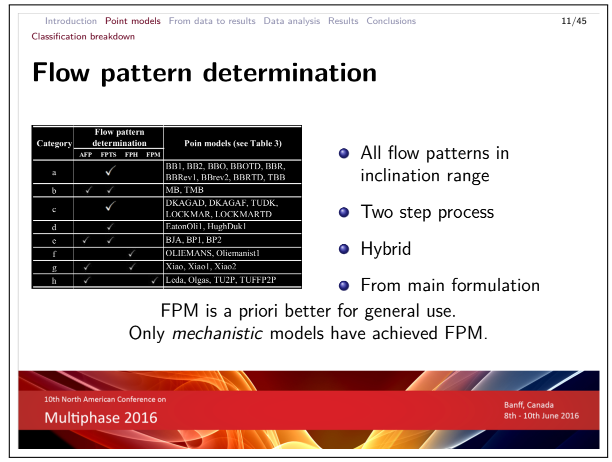
2.4.1 Conceptual integrity

Slide 6a shows two features of mathematical formulations that can de used to discriminate its inherent coherence. When the formulas for pressure drop and liquid holdup applicable to all flow patterns that may be found for a certain system are derived from a general governing idea, they are deemed more coherent that if all are separately derived from different ideas.

Both traditionally considered *empirical*, like group **b**, and *mechanistic* models like those in group **h**, may show structural consistent formulations, see slide 6a.



(a) ***Features of model coherence***



(b) ***Features related to flow pattern effects***

Figure 6: Conceptual integrity and flow pattern aspects of point model technology

2.4.2 Flow pattern dependency

The first column, AFP, in slide 6b refers to a feature of point models that allow them to predict the presence of all experimentally observed flow patterns for some set of input conditions. Point models that can do this are called comprehensive for the range of input specified.

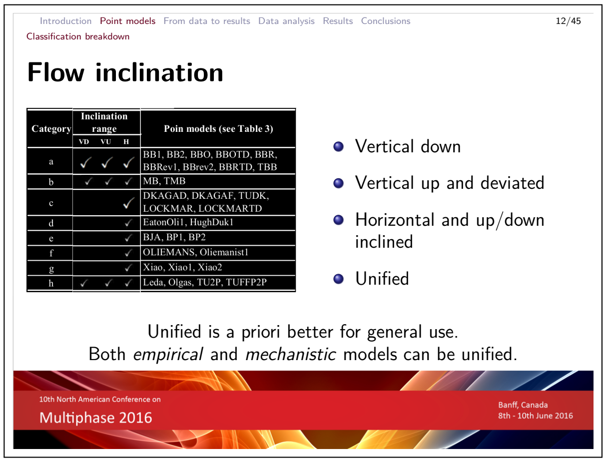
The next three columns refer from left to right to the level of separation between the calculation step where the existence of a flow pattern is made and the step where the solution of the pressure drop and liquid holdup happen for the flow pattern that was predicted.

The first column is for the point models where these are two clearly separate and potentially dissimilar processes. This was a very common practice in the early decades of development of this technology. The last column is for the models where these steps are highly or fully integrated and the one in between is for those in a hybrid state of integration.

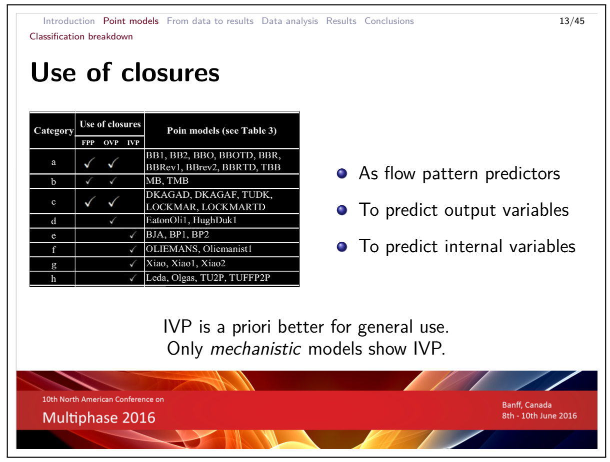
Mechanistic models tend to have hybrid or full integration with the potential of minimizing large discontinuities.

2.4.3 Inclination effects

Slide 7a shows the ranges of inclinations as a feature that serves to categorize point models. However it does not discriminate between empirical and mechanistic models because groups in both ends of that scale display this feature. When a point model covers all inclinations is called unified.



(a) ***The effect of inclination***



(b) ***Closures as sources of empiricism***

Figure 7: Features related to inclination effects and use of closures in point model technology.

2.4.4 Use of closures

Slide 7b shows the strongest feature traditionally used to classify models as empirical or mechanistic. Using pure data correlations for anything but internal model variables puts more weight on the underlying structure of the formulation to define the output.

If that structure is built upon the laws of physics then the model can be more predictive and less reliant on the range of the data used to develop its closures.

3 From data to results

Another requirement to classify technology is to have sufficient representative data to validate the a priori classification. All with the purpose to identify which types of technology perform better.

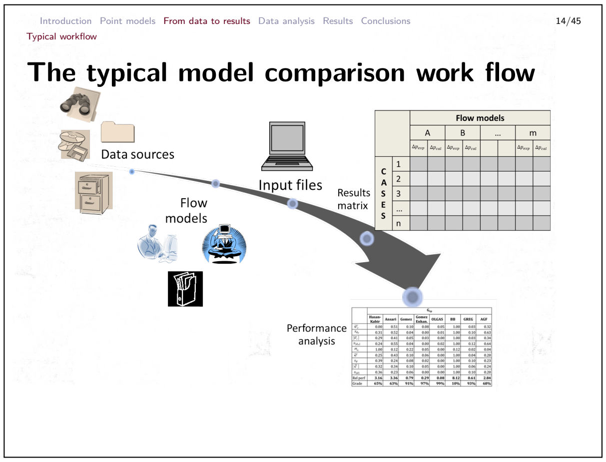
3.1 Typical workflow

Slide 8a shows the sequence of tasks from obtaining, and preparing the data to obtaining and selecting the point models to generating input files or case files to run in software, then generate results matrices and from them performance indices. With the indices one can compute grades and feed them to classification algorithms.

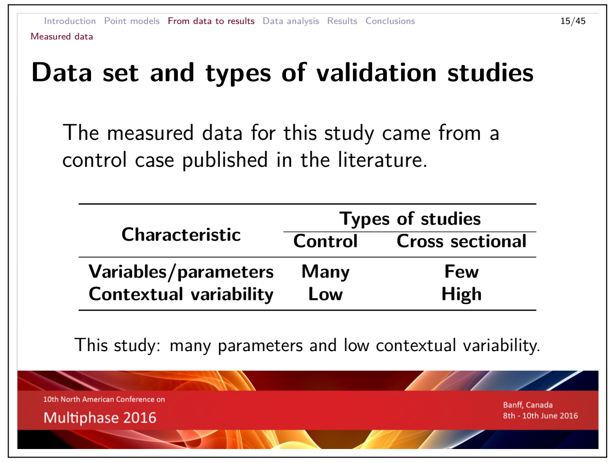
3.2 Measured data

Slide 8b talks of the data used for this study. It came from a control case published in the literature.

A control case restricts the depth of the phenomena studied to the conditions and specifics of the case study. This is in contrast with a cross-sectional study where many different cases represent very



(a) ***From data to performance***



(b) ***A control study***

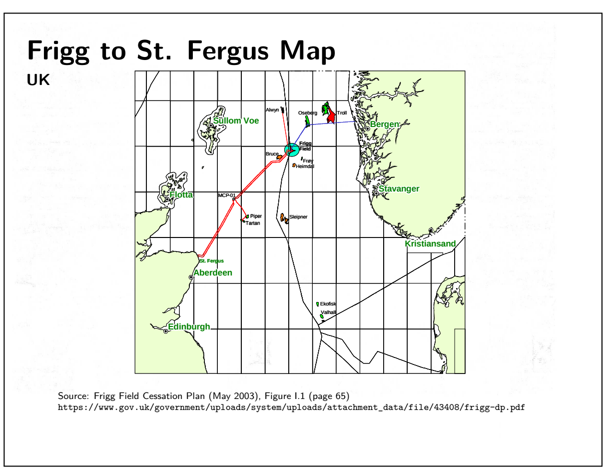
Figure 8: Workflow and control study to generate classification data.

different conditions. Control studies allow to focus on more variables/parameters at a time while cross sectional studies are aimed at fewer variables/parameters but across more contextual scenarios.

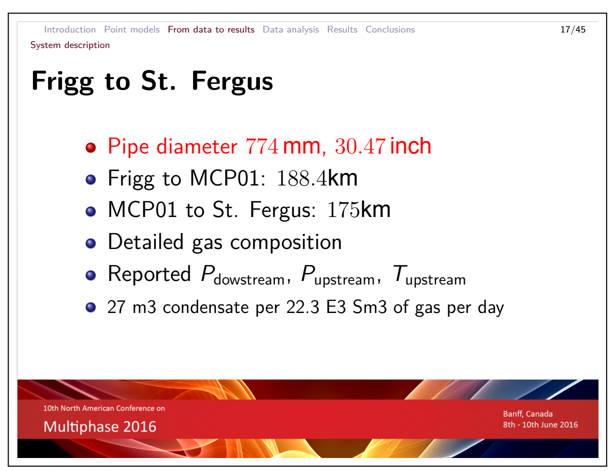
The case had data from around 1982-83 from the Friggs to St. Fergus system in the North sea between the Uk and Norway. The field has being already decommissioned. It is an case of transmission of a produced gas fluid over around 360 km in a cold subsea environment. The cumulative liquid production and the fluid characterization show that it is a very challenging low liquid loading case.

3.3 System description

Slide 9a gives the location of the field and a sense for the distance to the St. Fergus terminal in the UK.



(a) ***Geographical context***



(b) ***Line specifications***

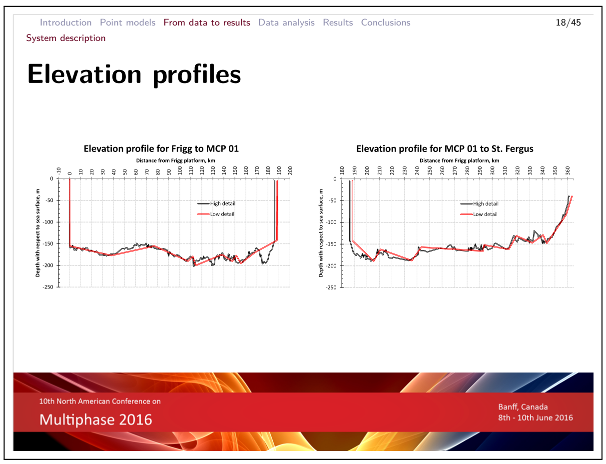
Figure 9: Frigg Field and St. Fergus terminal locations and some system features.

Slide 9b in turn presents the main features of the production tie-back broken down in two sections. The reported pressures at both ends allowed to parameterize on the boundary specification to test sensibility of the point models in a flow simulator context.

Slide 10a shows the kind of elevation profiles available for this study. This allowed to sensitize on this feature. Modern technology has made super detailed bathymetric and actual pipeline profiles readily available so that holdup effects can be simulated with higher fidelity.

3.4 Experimental campaign

Slide 10b brings the three main groups of field data corresponding to different flow rates, inlet temperatures, inlet pressures, pipe lengths, and corresponding elevation profiles. There was no attempt to use sensitivity on any of these groups as they are all similar and represent the then normal operation of the field.



(a) ***Pipe elevation profiles***

Type	File	$Q_{\text{eq-gas}}$ (e3sm^{3}d)	P_{upstream} (bar)	$P_{\text{downstream}}$ (bar)	T_{upstream} ($^\circ$C)
Group 1	FriggStFergus01	31.6	143	50	47
	FriggStFergus02	28.7	132	50	47
	FriggStFergus03	27.3	145	84.5	47
	FriggStFergus04	22.5	131	88	47
	FriggStFergus05	21.2	105	50	47
	FriggStFergus06	19.9	108	66	47
	FriggStFergus07	15.6	114	89	47
	FriggStFergus08	33.5	149	49	28
Group 2	FriggMCP01	33.4	149	109	28
	FriggMCP02	38.9	148	89	29
	FriggMCP03	40.9	148	30	33
	FriggMCP04	43.8	148	17	28
Group 3	MCPStFergus01	32.4	109	51	36
	MCPStFergus02	33.4	109	48.5	23
	MCPStFergus03	36.1	117	47.5	33
	MCPStFergus04	38.4	123	48.5	46
	MCPStFergus05	38.9	125	48	6
	MCPStFergus06	40.9	131	48.5	5
	MCPStFergus07	43.8	140	49	5

Group 1: Frigg to St. Fergus; **Group 2:** Frigg to MCP01; **Group 3:** MCP01 to St. Fergus

(b) ***Experimetal campaign***

Figure 10: Elevation data and experimental campaign.

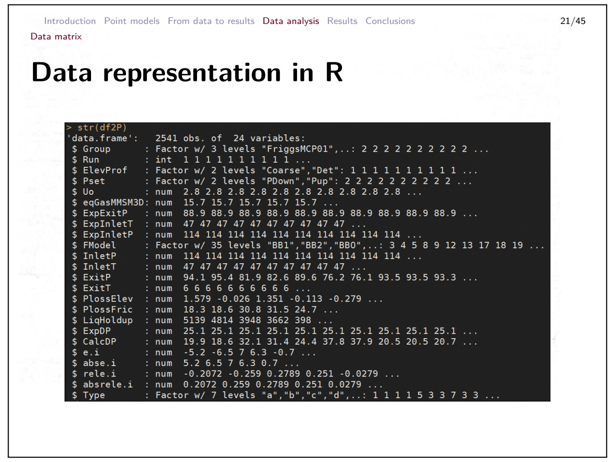
Slide 11a shows the liquid slug catcher at the St. Fergus terminal. Many fields connected to this important gas processing hub in the UK and operations used this huge gas-liquid separator to handle the liquids coming from flow ramp ups, and pigging operations, in the many lines, including the twin Frigg lines, each holding a little over more than 500 tonnes of condensate at any given time.

4 Data analysis

After the data was prepared and the point models identified, two variables were were selected as parameters to test sensitivities in the point model classification and thus reach more robust conclusions:

1. Location of the specified pressure boundary for the simulations: upstream or downstream
2. Type of elevation profile available: coarse or detailed

(a) ***Slug catcher infrastructure at St. Fergus***



(b) ***2541 simulated ΔP and 24 variables***

Figure 11: Liquids handling equipment at St. Fergus and data matrix in R.

4.1 Data matrix

After the different cases were built and ran in the simulator. The total number of combined cases was 19 runs $\times$ 35 models $\times$ 2 pressure specification locations $\times$ 2 profiles $=$ 2660, however due to stalled solutions in some of them, only pressure drop results for 2541 cases were recorded. Slide 11b shows the matrix representation in the statistical language R.

The results matrix had the type of a priori point model as a variable for each combination of run, point model, and elevation profile. This made easier to keep track of the results of the classification algorithms in order to test if they predicted the a priori groups in common clusters.

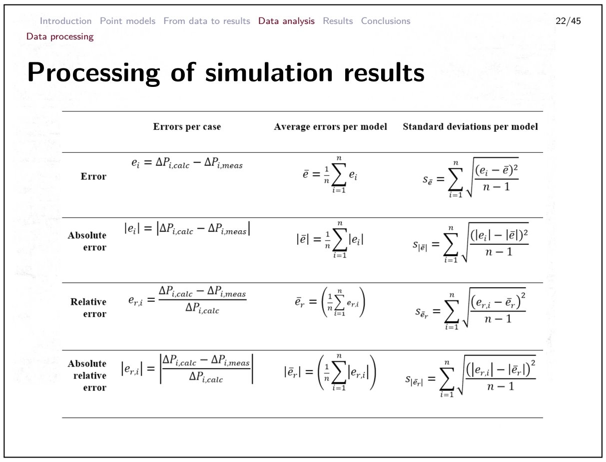
4.2 Data processing

Further processing of the simulated pressure drop results can be achieved using four basic types of errors. The last transformation is to turn each error into statistical variables and then normalize them to be used as indexes for performance comparisons and for classification.

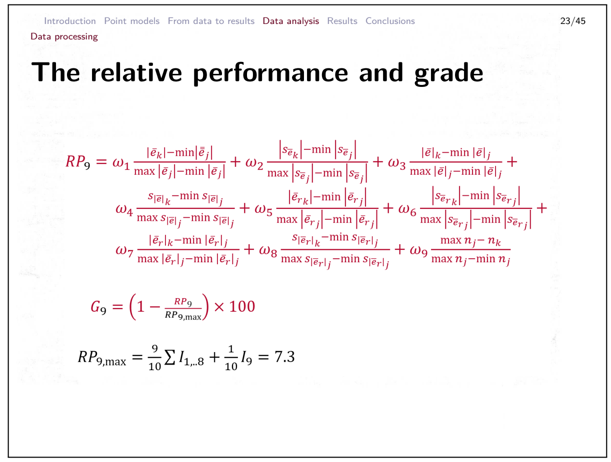
Slides12a and 12b show the formulas used for these transformations. The relative performance formula was adjusted with weighting factors to lower the contribution of the index for successful cases solved because some point models obtained too high a score on this account alone while being very poor predictors otherwise.

5 Results

From this point on the slides will show the bar plots of grade for different combinations of filters applied to the enlarged results matrix with all the derived variables obtained in the previous step. Since some of these filters mask interactions of some of the parameters involved in the filters only comparisons of the more refined filters serve the purpose of identifying the best conditions to improve



(a) ***Derived variables from simulated*** Δp



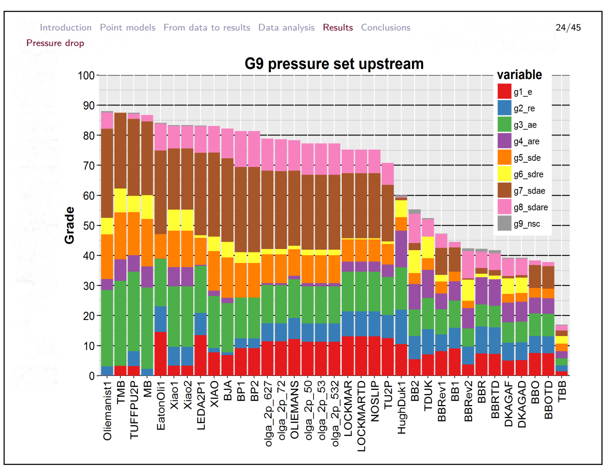
(b) ***Relative performance and grade***

Figure 12: Derived variables from pressure drop and the creation of indexes from them to compute relative performance and grade.

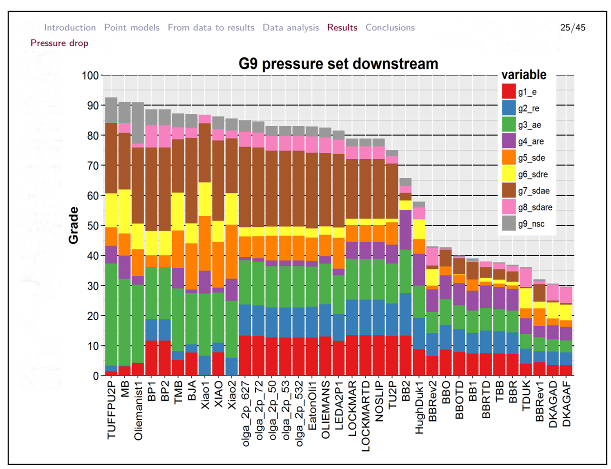
the overall accuracy and improve the robustness of the flow simulator regardless of the point model employed.

This last transformation made the maximum distance possible between point models in terms of relative performance indexes equal to 7.3 instead of the unweighed theoretical maximum of 9.0 for nine indices.

5.1 Pressure drop



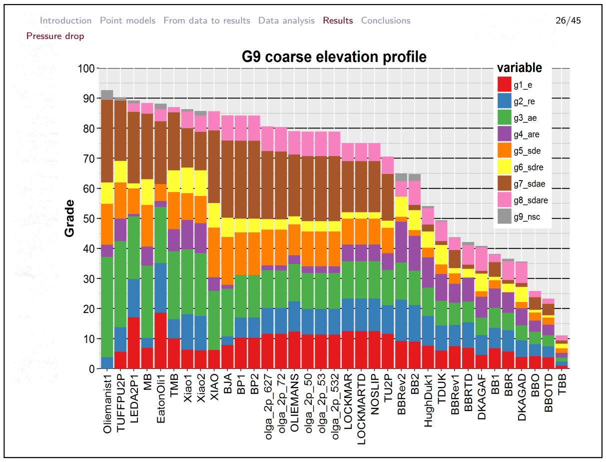
(a) ***Pressure set upstream***



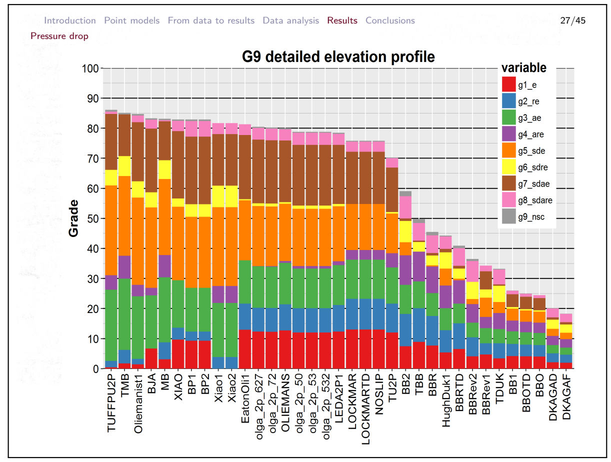
(b) ***Pressure set downstream***

Figure 13: Effect of pressure specification location on Δp.

In order to investigate the effect of the location of the pressure specification, visualizations of the grades broken down by components can be seen side by side in slides 13a and 13b. From these plots one could derive that higher grades can be obtained from running point models with the pressure specification located downstream.



(a) ***Coarse***

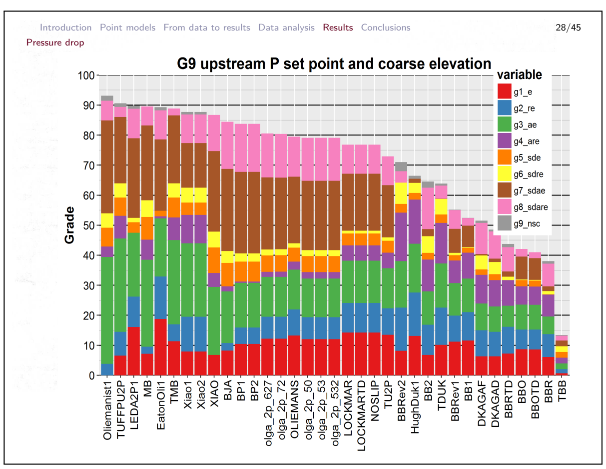


(b) ***Detail***

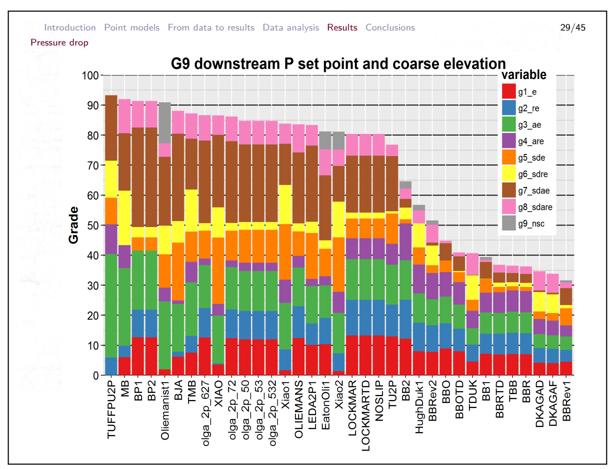
Figure 14: Effect of elevation profile detail on Δp.

The apparent effect of the elevation profile detail can be seen in slides 14a and 14b. From these plots it would seem higher grades can be obtained by the better performers when a coarse profile is used.

However both comparisons have potential mixed effects from interactions between elevation profile and pressure specification location. So it is appropriate to observe the plots with one filter constant while the other is varied.



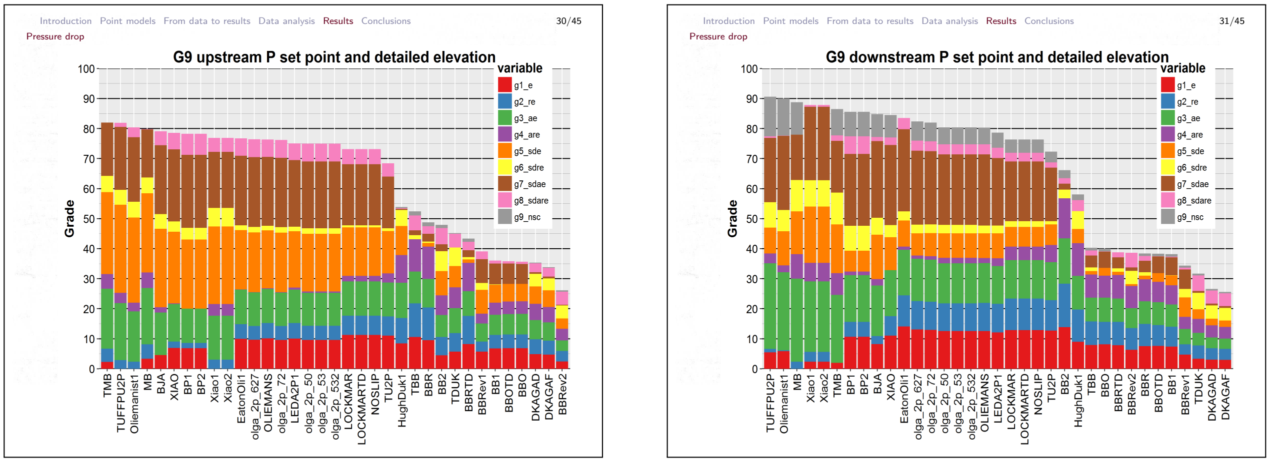
(a) ***Upstream*** p



(b) ***Downstream*** p

Figure 15: Effect on Δp of the pressure specification location maintaining the coarse profile filter.

Slides 15a and 15b maintain the coarse elevation profile constant and one can appreciate that the point models score higher across the top performers when the pressure is set downstream.

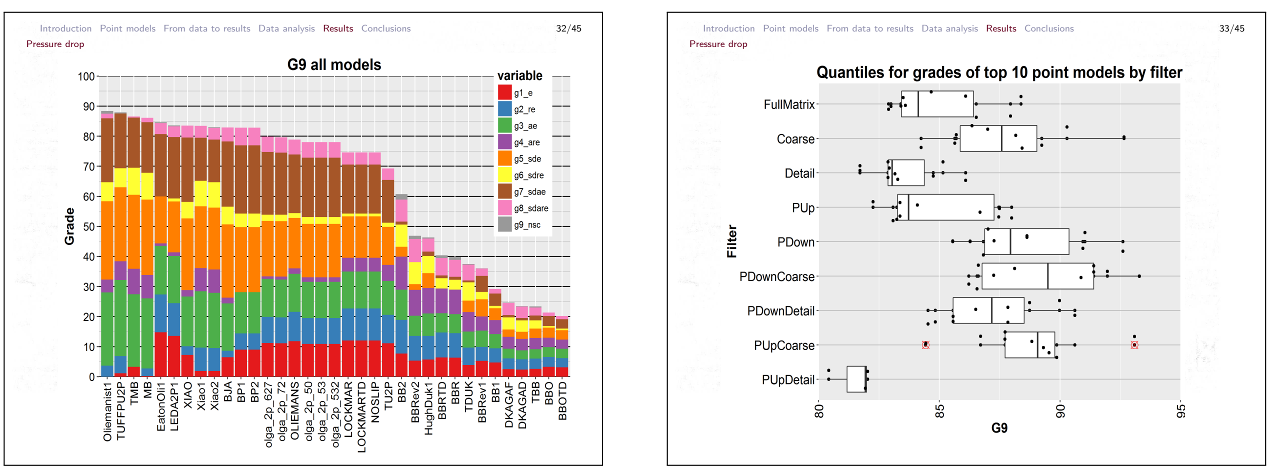


(a) *Upstream* p **(b) *Downstream* p**

Figure 16: Effect on Δp of the pressure specification location maintaining the detail profile filter.

Slides 16a and 16b contrast the effect of the same detail elevation profile filter while varying the pressure specification location. Similarly but now even more notoriously the beneficial effect of setting the pressure specification downstream is manifested. Another observation is that the index for number of cases successfully converged is significant in the top performers, meaning that other models are failing in this respect significantly.

From these observations it can be concluded that for this system the sensitivity to pressure specification location for a given level of detail in the elevation profile is lower when the pressure is specified downstream. Compare the more consistent results of slides 15b and 16b versus those of varying the elevation profile from coarse to detail with the pressure specification upstream of slides 15a and 16a.



(a) *Aggregate* G_9 *bar plot* **(b) *Box plot***

Figure 17: Aggregate G_9 bar plot and box plot for Δp grades by filters.

Another way of analysing this data is looking at he box plot of the grade results. Slide 17b shows that the filter corresponding to the pressure specification downstream and the detail elevation profile has one of lowest variability, measured as the distance between the 25 and 75% quantiles again highlighting the potential for being a very robust setting to make any point model selection from among the top performers in Δp.

On the other side the configuration with the pressure specification upstream and the coarse elevation profile is the only one presenting outliers, with means a user should avoid it is no other information about the point model selection is to be made.

Another way of analysing the performace of the point models from the Δp results is shown in slide 18a. This list of top 10 models overall shows the best performers for pressure drop prediction in a system with very low liquid loading, and very shallow angles: less than 1 percent steeper than 2 degrees from results not shown in this paper. It should not be generalized and most importantly it does not show the whole picture because holdup has not being considered yet.

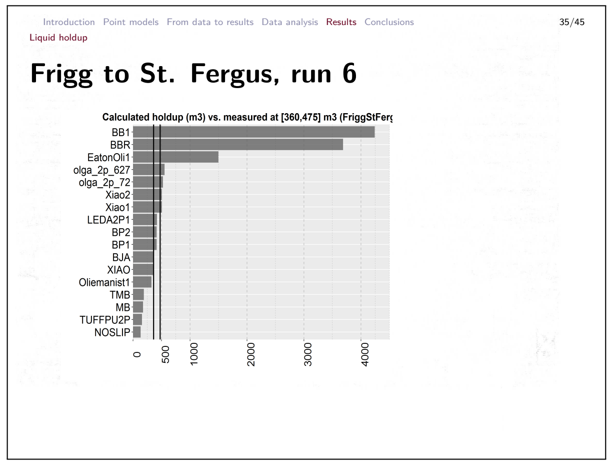
Introduction Point models From data to results Data analysis Results Conclusions 34/45

Pressure drop

Point models for low liquid loading

Position	Point model	Avg G9	Times in top 10
1	TUFFPU2P	91.1	9
2	Oliemanist1	90.9	9
3	MB	89.7	9
4	TMB	89.0	9
5	XIAO	87.2	9
6	BJA	87.0	9
7	Xiao1	87.2	8
8	Xiao2	87.0	8
9	BP1	88.1	5
10	BP2	88.1	5
11	EatonOli1	89.1	4
12	LEDA2P1	89.0	4
13	olga_2p_627	89.1	1
14	olga_2p_72	88.8	1

(a) ***Top 10 models overall***



(b) ***Liquid holdup run 6***

Figure 18: Performance from another angle and holdup data.

5.2 Liquid holdup

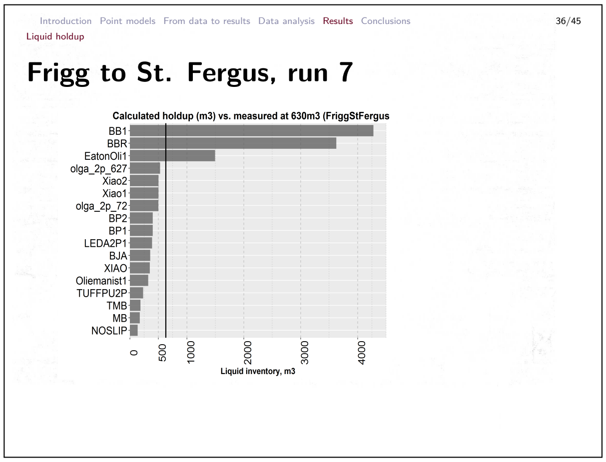
Slides 18b and 19a show that looking at the prediction of liquid retention in the line is paramount to determine the better models or technology to use. However the lack of sufficient reported data for holdup makes estimating G_9 performance impossible as illustrated in slide 19b.

5.3 Insufficient holdup data

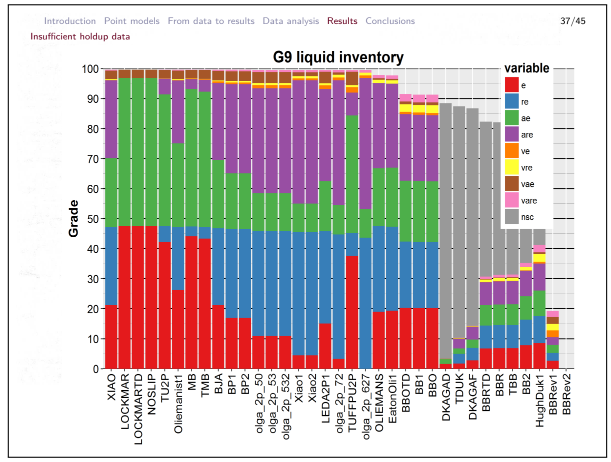
Some take-away points from the comparison of the top 10 performers for Δp and two Beggs and Brill models with the two data points for holdup:

1. 4 of the 14 top performers in Δp made it to within 25% of the two reported holdup values
2. Some empirical models should be completely avoided for holdup predictions
3. Seemingly good pressure predictors can under-perform severely in this low liquid loading case

One can never understate the importance of validation data to establish the accuracy of the estimation of line liquid inventory. This is a key aspect for the safe and effective operation at the plant.



(a) ***Liquid holdup run 7***



(b) ***Failure of using G_9 with only two data points***

Figure 19: Liquid holdup results and lack of resolution in holdup G_9 due to insufficient data points.

5.4 Classification

Classification is a primary task in knowledge creation. Modern data science offers easy to use and powerful techniques to perform this task. It is still an area of development and although there are strong theoretical foundations there are still discussions as to what the most appropriate techniques are for specific types of problems.

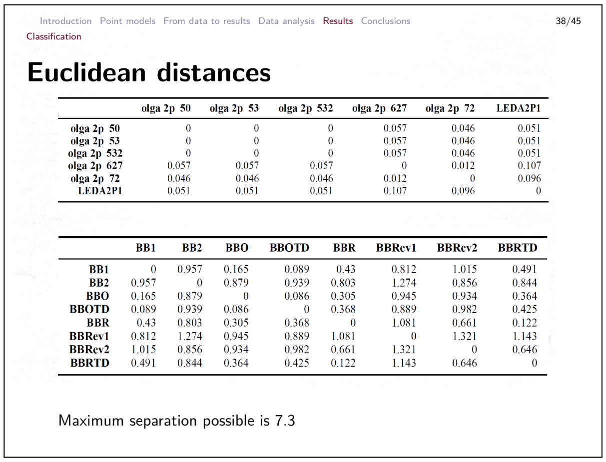
5.5 Hierarchical clustering

This classification technique uses a bottom up approach where each model starts as it own cluster but are joined with the one closest as judged by some metric of distance and then the resulting clusters united with those closest again by the same metric until only one remains.

Common to every classification technique is the measurement of the distance between two points or between two clusters of points. Applying Euclidean distance one can compute the normalized distances from the 9 indexes used for performance measurements.

Slide 20a shows the mutual distances between OLGAS and Leda point models and between the Beggs and Brill implementations. Considering that the maximum separation possible is 7.3 the OLGAS and Leda results look very close together with a maximum separation of 0.1 compared to the 10 times more separated Beggs and Brill point models.

The results of hierarchical clustering are visualized as dendograms, like the one shown in slide 20b where the distance between clusters is the average link, defined as the average distance between each point in one cluster and each point in the other cluster. Several distance metrics were tried in the R code and this one was chosen to give visually clearer separation between clusters.

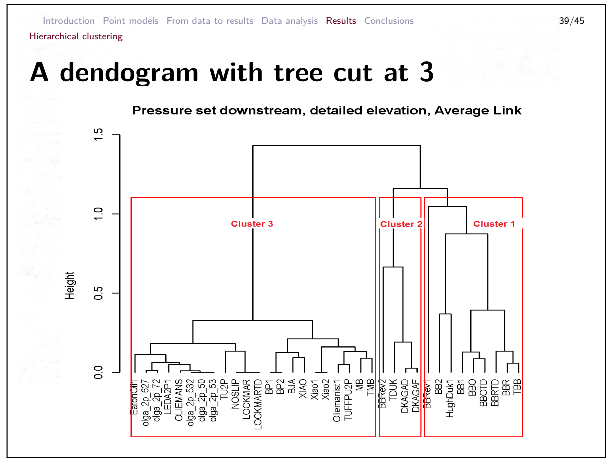


	olga 2p 50	olga 2p 53	olga 2p 532	olga 2p 627	olga 2p 72	LEDA2P1
olga 2p 50	0	0	0	0.057	0.046	0.051
olga 2p 53	0	0	0	0.057	0.046	0.051
olga 2p 532	0	0	0	0.057	0.046	0.051
olga 2p 627	0.057	0.057	0.057	0	0.012	0.107
olga 2p 72	0.046	0.046	0.046	0.012	0	0.096
LEDA2P1	0.051	0.051	0.051	0.107	0.096	0

	BB1	BB2	BBO	BBOTD	BBR	BBRev1	BBRev2	BBRTD
BB1	0	0.957	0.165	0.089	0.43	0.812	1.015	0.491
BB2	0.957	0	0.879	0.939	0.803	1.274	0.856	0.844
BBO	0.165	0.879	0	0.086	0.305	0.945	0.934	0.364
BBOTD	0.089	0.939	0.086	0	0.368	0.889	0.982	0.425
BBR	0.43	0.803	0.305	0.368	0	1.081	0.661	0.122
BBRev1	0.812	1.274	0.945	0.889	1.081	0	1.321	1.143
BBRev2	1.015	0.856	0.934	0.982	0.661	1.321	0	0.646
BBRTD	0.491	0.844	0.364	0.425	0.122	1.143	0.646	0

Maximum separation possible is 7.3

(a) ***Euclidean distances***



(b) ***Dendogram***

Figure 20: Euclidean distances for two groups of point models sharing similar formulations and dendogram with tree cut at three branches.

5.6 Hierarchical clusters

Slide 21a shows the composition of the three clusters obtained form cutting the dendogram from slide 20b at three clusters. 73% of cluster 3, where the top performers for Δp are, belong to categories **e** to **h**, the ones with a lowest degree of empiricism judged by the use of closures and the integration of flow pattern and hydrodynamic calculation steps. In contrast cluster 1 with an average 43.6 G_9 is made up of 89% group **a** point models, while cluster 2 with a representative G_9 of 30.6 is made up of 25% group **a** and the rest group **c**, perhaps the most basic type of models.

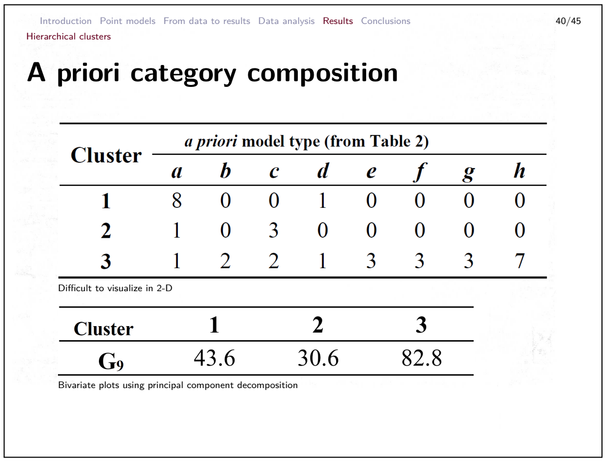
5.7 Partitioning clustering

Partitioning is a top bottom strategy where a fixed number of clusters is assigned arbitrarily to each cluster at the start. Then at every iteration the members are shuffled according to their distance to the center of gravity of each cluster until the differences within iterations converges within a tolerance. There are techniques for estimating the appropriate number of clusters.

5.8 Partitioning clusters

K-means is a representative technique of this classification technique. Slide 21b shows the results for k=3 using the default implementation and options of k-means in R. The second cluster has an even larger mean G_9 than in the hierarchical clustering technique used before. The names of the models shown in the table for clusters using partitioning around medoids are only labels to be able to refer to the cluster.

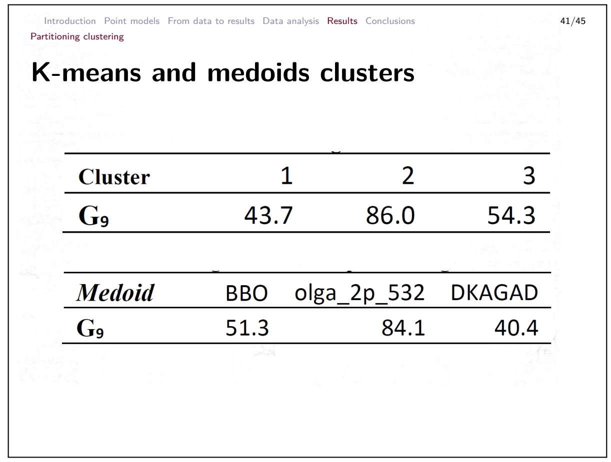
Slide 22a shows a visualization of the clusters obtained with partitioning around medoids (PAM) using a bivariate plot where the axis have the two main principal components that explain 75.5% of all variability. Medoids are representative points for their clusters as opposed to centroids which are



Cluster	*a priori* model type (from Table 2)							
	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*
1	8	0	0	1	0	0	0	0
2	1	0	3	0	0	0	0	0
3	1	2	2	1	3	3	3	7

Cluster	1	2	3
G_9	43.6	30.6	82.8

(a) ***Hierarchical cluster composition***



Cluster	1	2	3
G_9	43.7	86.0	54.3

Medoid	BBO	olga_2p_532	DKAGAD
G_9	51.3	84.1	40.4

(b) ***Partitioning clusters***

Figure 21: A priori category recovery after using hierarchical clustering and K-means results.

a mean value. The sum of all the separations between members of a cluster and their medoid can use any metric and thus it can be applied to any type of variable and it is less susceptible to outliers.

6 Conclusions

In summary one can say confidently that there is no silver bullet. Different point models may be satisfactory for this type of systems for both pressure drop and holdup prediction. The better ones when one considers pressure drop and liquid holdup are: olgas_2p_627, olgas_2p_72, Xiao1, and Xiao2.

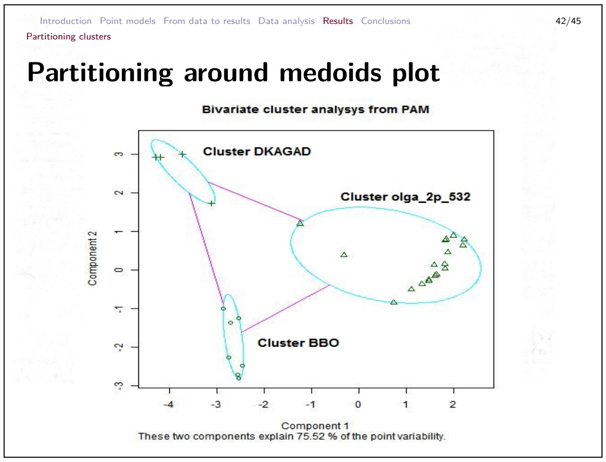
The lack of sufficient holdup data for validation prevents confirming or generalizing these results any more.

6.1 Model performance

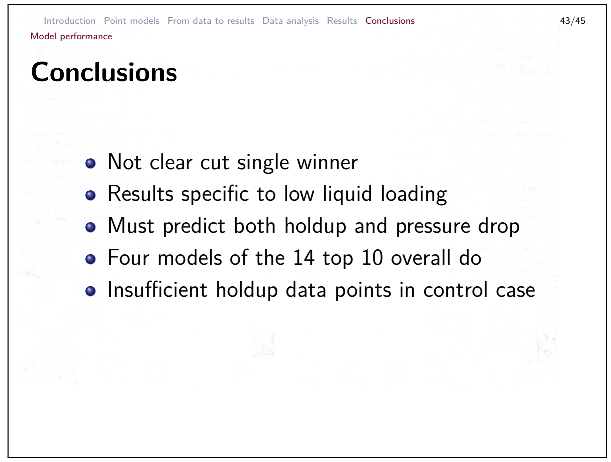
The top performers in pressure drop may not be the best for holdup in a low liquid loading case like this. Always use both Δp and holdup as criteria for performance.

6.2 Simulation strategies

Always prefer to simulate with the pressure specified downstream to minimize sensibility to other uncertainties like elevation profile detail. Try to have as much elevation detail as possible. Always compare more than two point models, specially if no validations are available or if no data are available. Pressure drop alone is not sufficient.



(a) ***Bivariate plot for partitioning around medoids***

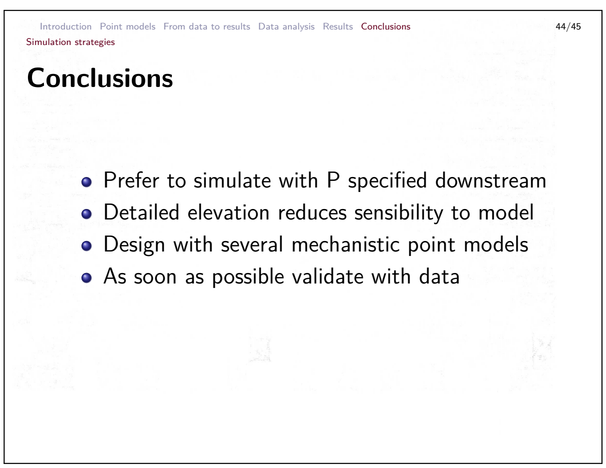


(b) ***Conclusions on model performance***

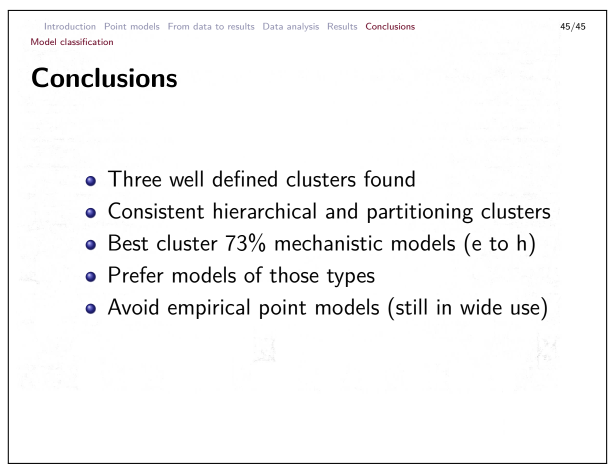
Figure 22: A priori category recovery after using hierarchical clustering and K-means results.

6.3 Model classification

Select models from groups like **e** to **h**. They present the higher levels of flow pattern integration and less empirism. The classification with hierarchical clustering produced clusters with very clear point model association to a priori classifications.



(a) ***Simulation***



(b) ***Classification***

Figure 23: Conclusions on simulation and classification.