A data-driven workflow for predicting horizontal well production using vertical well logs

Jorge Guevara, Matthias Kormaksson, Bianca Zadrozny
IBM Research, Brazil

Ligang Lu, John Tolle, Tyler Croft, Mingqi Wu, Jan Limbeck, Detlef Hohl
Shell Inc

Workshop de eScience   UFABC, 2017

Motivation

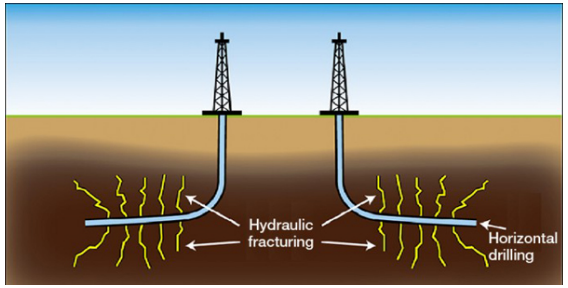
- In recent years, interest in unconventional resource exploration has grown substantially due horizontal drilling and hydraulic fracturing techniques.

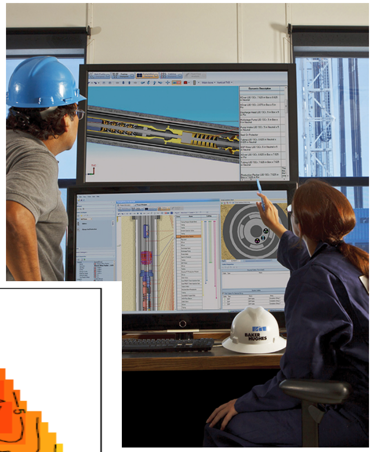
- Currently, the industry is in a state of ”trial-and-error” and does not have a systematic methodology for choosing optimal engineering parameters for individual wells and for unconventional reservoirs as a whole.

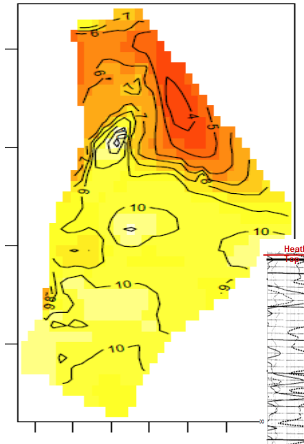
- Only immature research results are available regarding the physical characteristics of “sweet spots”, i.e., locations with high potential for oil and gas.

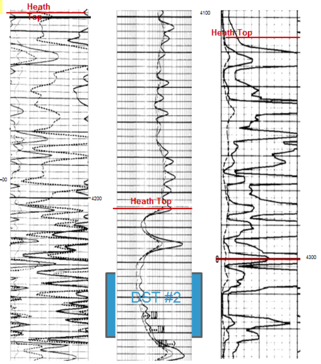
- With drilling costs at an all time high, **choosing the optimal completion parameters and right locations for new wells is a crucial issue**.

- There is a huge amount of available data that the industry has been collecting which opens up a **great opportunity to explore data-driven approaches**.







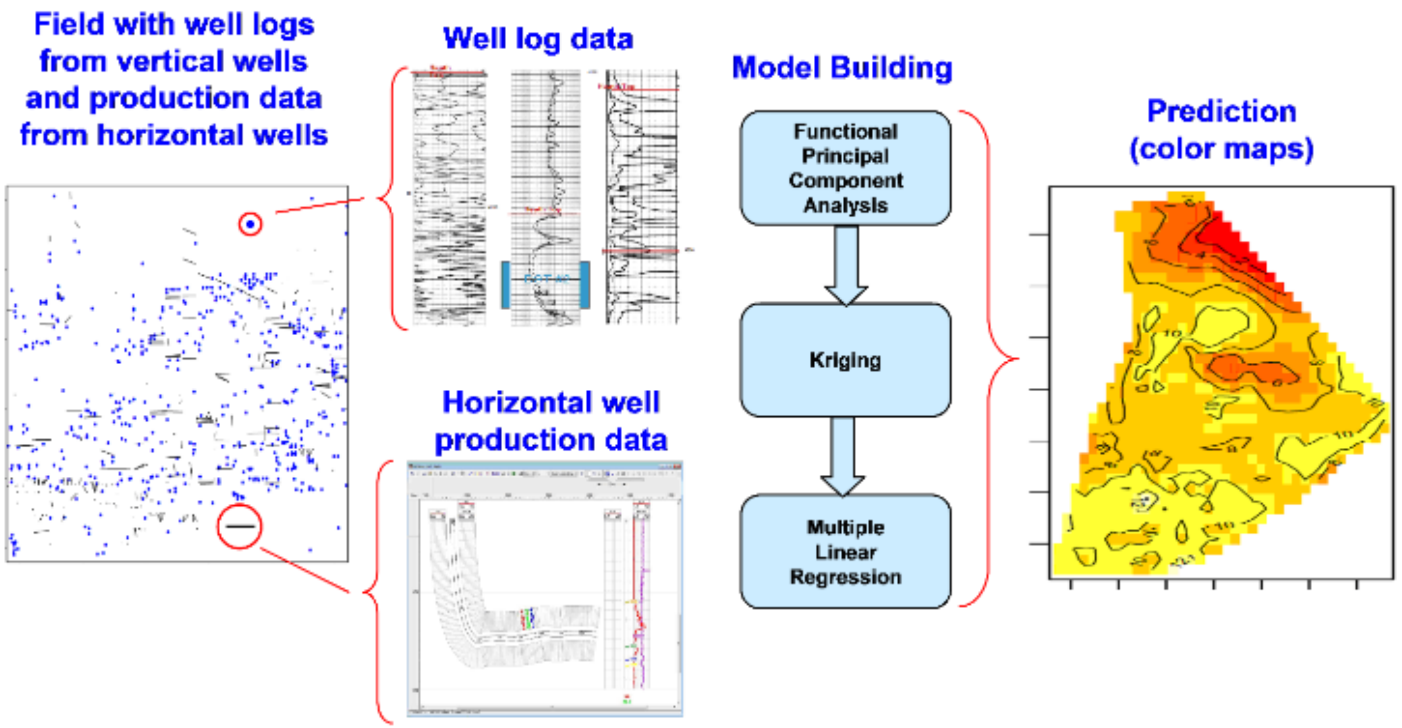


Project Objectives

- The overall objective of the project is to develop a methodology/workflow which may **help guide decisions on where to drill** and **how to complete a well** for unconventional exploration.

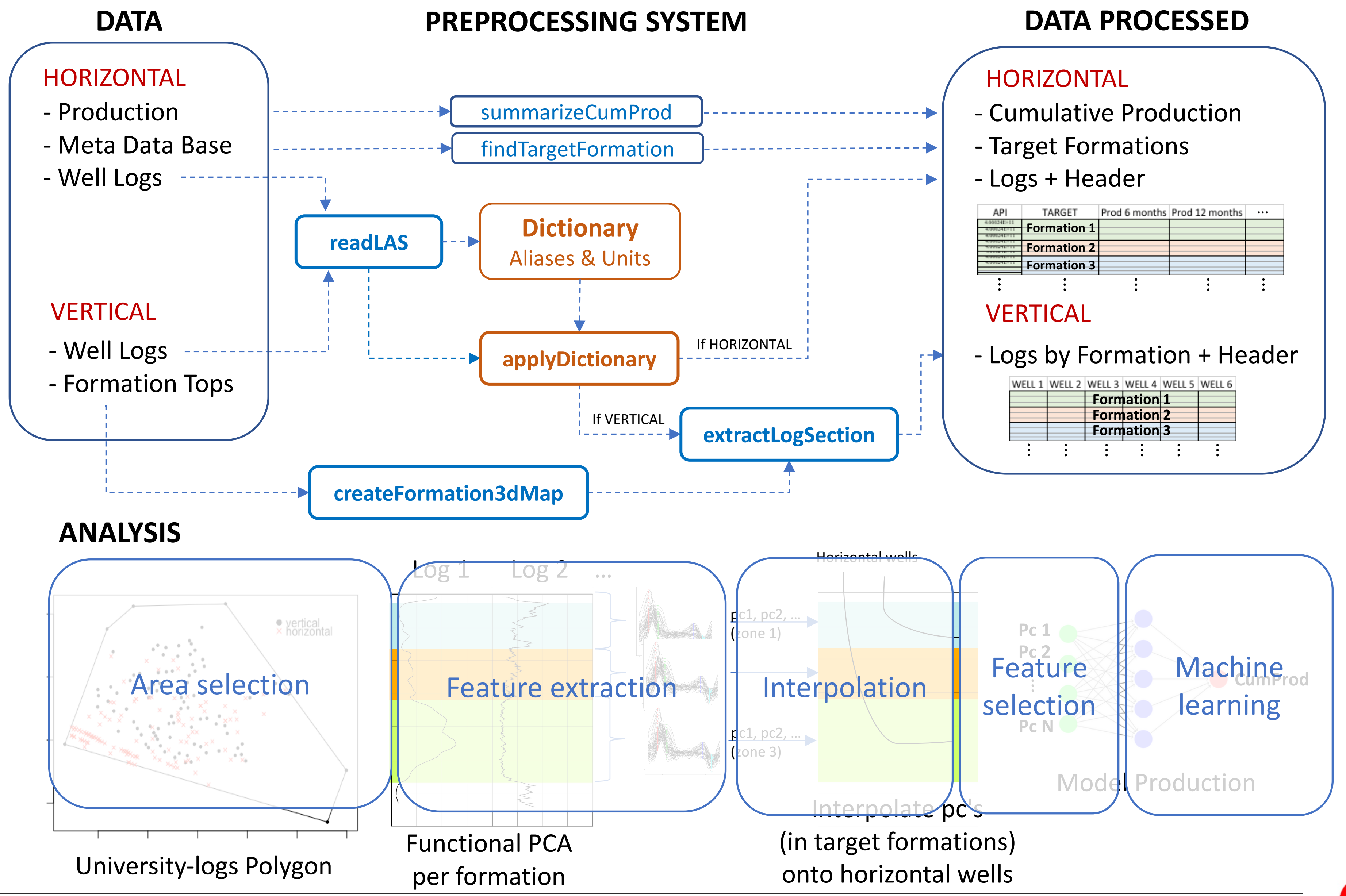
- Specific goals of the project include:
  a. Developing a methodology to build a **predictive model** using available Shell Data to help identify the most probable "sweet–spot" candidates for determining drilling locations and well completion parameters; and
  b. Developing **a software based workflow** of the developed methodology in (a) that includes data integration and interpolation steps; and
  c. Developing an **interactive method** to further evaluate the most probable "sweet-spot" candidates to help guide decisions about the drilling location and the well completion parameters.

Previous work

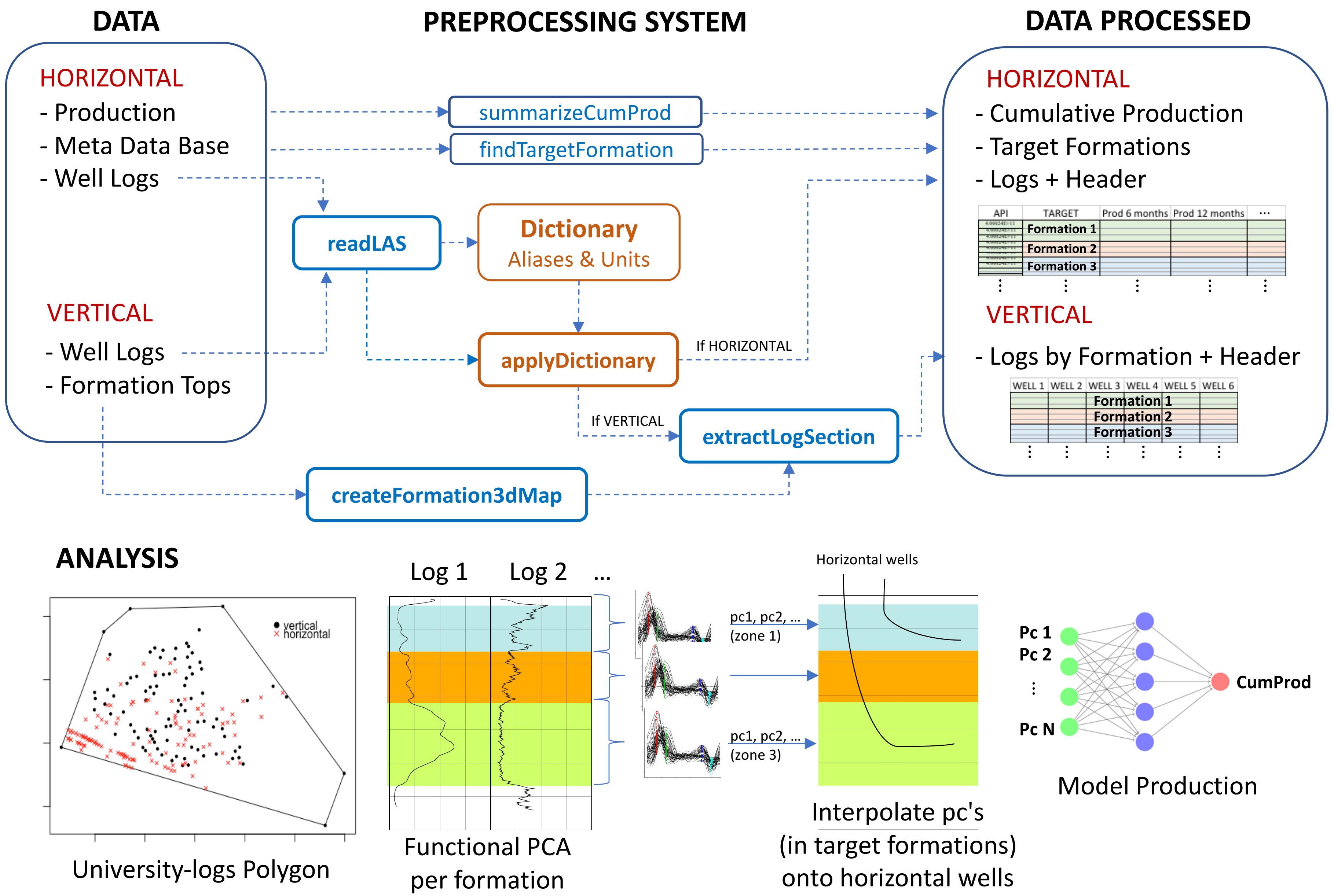


- IBM has developed a data-driven sweet spotting technique for shale plays previously explored with vertical wells.
- The technique involves two main steps:
  1. extracting features from high-dimensional vertical well logs using functional Principal Component Analysis (fPCA)
  2. building models that predict sweet spots in shale plays by correlating extracted features with production data from horizontal wells.
- Method tested with well log data from 2020 vertical wells and production data from 702 horizontal wells in a single field.
- Some important questions were left unanswered in this previous work, which Shell and IBM would like to address jointly in this project: use of features from horizontal well logs, effect of completion parameters.

Methodological Workflow



Methodological Workflow



Dictionary

LAS files

HEADER

```
~Curve Information
#___________________________________________________________________________
#
#LOGNAME            .UNIT     LOG_ID                          : DESCRIPTION
#___________________________________________________________________________
MD                  .ftUS                                     :  {F}
BSPG_MATRIX_POR     .v/v                                      : Description {F}
BVO                 .v/v                                      : Volumetric
Photoelectri factor {F}
BVW                 .v/v                                      : Description {F}
CALI                .in                                       :  {F}
DEN                 .g/cm3                                    :  {F}
DEN_SYN             .g/cm3                                    :  {F}
DENCOR              .g/cm3                                    :  {F}
```

LOG

```
~Ascii
6639        -9999       0.04365472 0.1061043  21.7864    2.4114     2.4114
0.0512      0           -9999      72.1787    -9999      39.6021    58.22269
0.189       -9999       0.149759   2.1292     -9999      1.479581   0
0.7085003   0.005328601 -9999       -9999      -9999      -9999
6639.5      -9999       0.04490386 0.1054577  21.7671    2.4104     2.4104
0.0507      0           -9999      71.6503    -9999      38.1671    55.88174
0.1981      -9999       0.1503615  2.1521     -9999      1.476274   0
0.7013607   0.0053156   -9999      -9999      -9999      -9999
6640        -9999       0.04487788 0.1052427  21.7532    2.4108     2.4108
0.0503      0           -9999      71.0466    -9999      36.3773    52.96199
0.2228      -9999       0.1501205  2.1598     -9999      1.471877   0
0.7010544   0.005302601 -9999       -9999      -9999      -9999
6640.5      -9999       0.0479381  0.1051342  21.7433    2.4059     2.4059
0.0519      0           -9999      70.7577    -9999      35.5752    51.6535
0.259       -9999       0.1530723  2.1637     -9999      1.467901   0
0.6868271   0.00528955  -9999      -9999      -9999      -9999
```

Summary from 700 horizontal LAS files:

log	unit	freq
GR	counts	46
GR	AAPI	108
GR	API	82
GR	GAPI	52
GR	api	18
GR	gAPI	100
GR	aapi	8
DEPTH	FT	302
DEPTH	F	16
DEPTH	ft	27
DEPTH	feet	16
DEPT	FT	74
DEPT	.1IN	1

⋮ ⋮

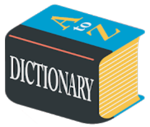
Total=2,041 mnemonics

Total=201 mnemonics

LAS files

Registered properties

New properties



Translate

Crains Petrophysical Dictionary

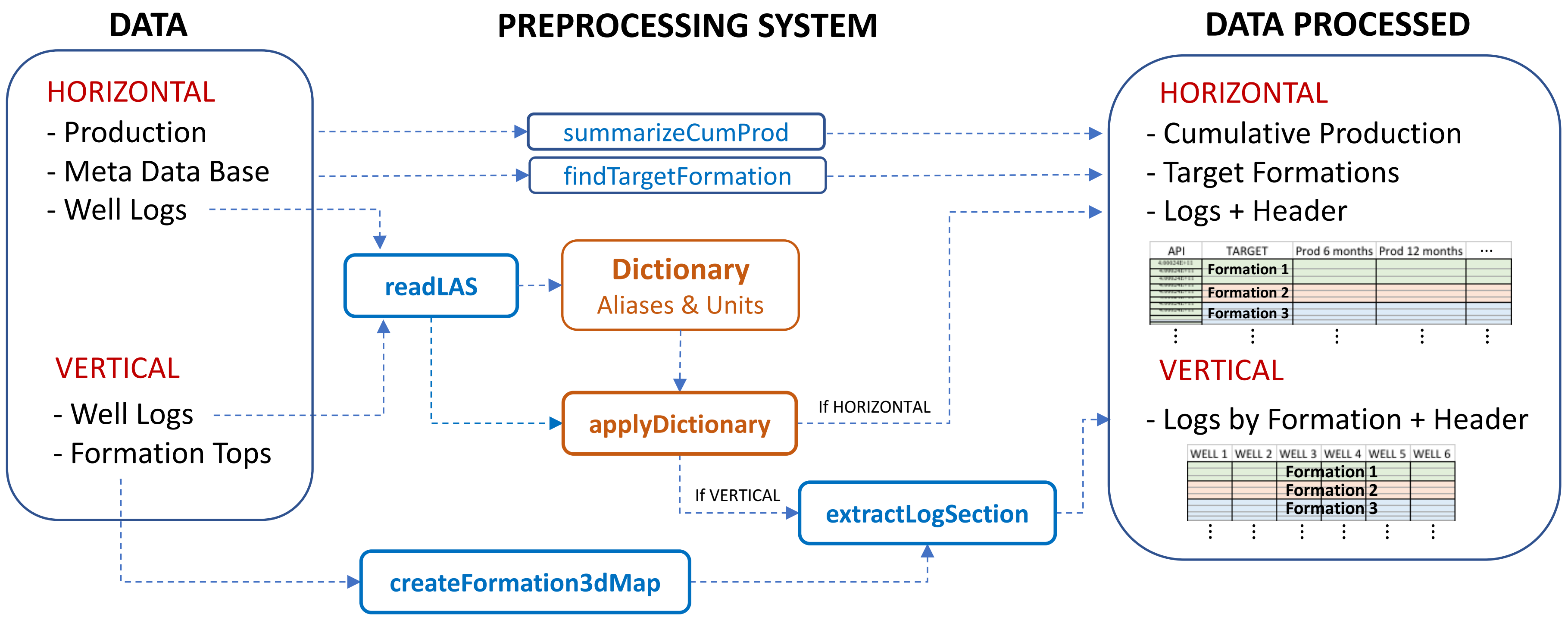
log	alias	source
GR	GR	Crains
DEPTH	DEPTH	Matthias
DEPT	DEPTH	Matthias
TVD	NA	NA
CUST	NA	NA
ROP	NA	NA
TEMP	NA	NA
INCL	NA	NA
SHALE	NA	NA
VS	NA	NA
SILT	NA	NA
C1	CAL	Crains
C2	CAL	Crains
SALT	NA	NA

User Edits?!?

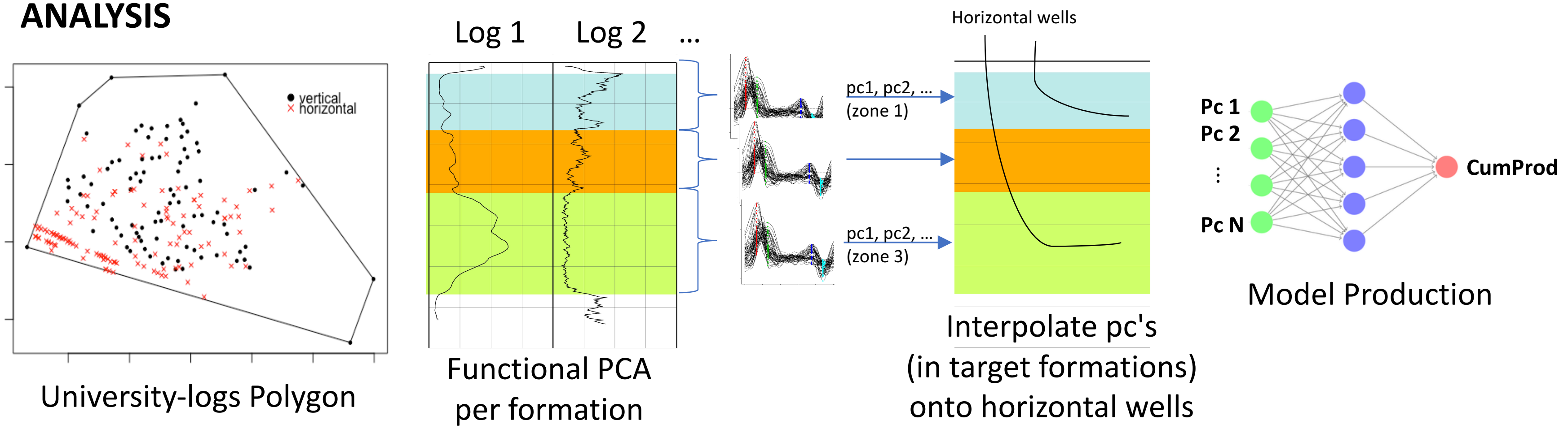
Log/Unit Dictionary with Descriptions

log	unit	freq	Name	Raw/Calculat	Description/Comments
BVO	v/v	90	Bulk Volume Hydrocarbon	Calculated	porosity * (1-water saturation) = BVO.
BVW	v/v	90	Bulk Volume Water	Calculated	porosity * water saturation = BVW.
CALI	in	34	Caliper	RAW	Measures hole size diameter. Shows areas where the borehole is washed out which rapidly degrades the quality of the data. Borehole enlargement can be a result of rock properties, poor mudweight selection, overpressured formation.
CALI	IN	56			
DEN	g/cm3	34	Density	RAW	Bulk density of the formation. The key log for calculating porosity.
DEN	G/CC	2			
DEN	G/C3	54			
DEN_SYN	g/cm3	24	Synthetic Density	Calculated	Will match DEN for >95% of the borehole length. Areas with poor data due to borehole enlargement are clipped out and a relationship is built with other logs to fill in the gaps. We can probably delete this.
DEN_SYN	G/CC	40			
DEN_SYN	G/C3	26			
DENCOR	g/cm3	33	Density correction	RAW	Logging vendor has an algorithm to correct the density log based on borehole enlargement, this is the correction applied. This is a good way to rapidly spot questionable density data.
DENCOR	G/CC	2			
DENCOR	G/C3	55			

Methodological Workflow



ANALYSIS

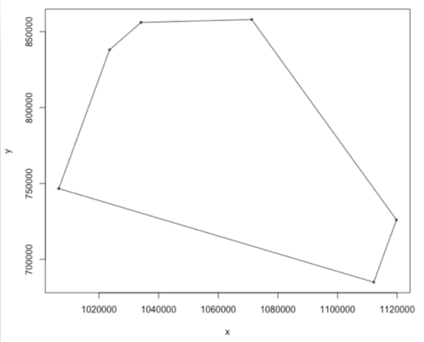


Horizontal Production Data Pre-processing (SummarizeCumProd)

Production data:

API_10	Data.Source	Elapsed.Days	Daily.Oil.Prod.Days	Daily.Gas.Prod.Days	Oil.Cum.Mbbl	Gas.Cum.MMcf
4210932622	IHS_NonDups	790.4	790.4	790.4	79.70600	600.796
4210932622	IHS_NonDups	820.8	820.8	820.8	81.07900	612.395
4210932622	IHS_NonDups	851.2	851.2	851.2	82.34600	624.315
4210932622	IHS_NonDups	881.6	881.6	881.6	83.31200	635.929
4210932624	Shell OFM	-1.0	0.0	0.0	0.00000	0.000
4210932624	Shell OFM	0.0	1.0	0.0	0.04500	0.000
4210932624	Shell OFM	1.0	2.0	0.0	0.18700	0.000
4210932624	Shell OFM	2.0	3.0	0.0	0.36400	0.000

Polygon area:



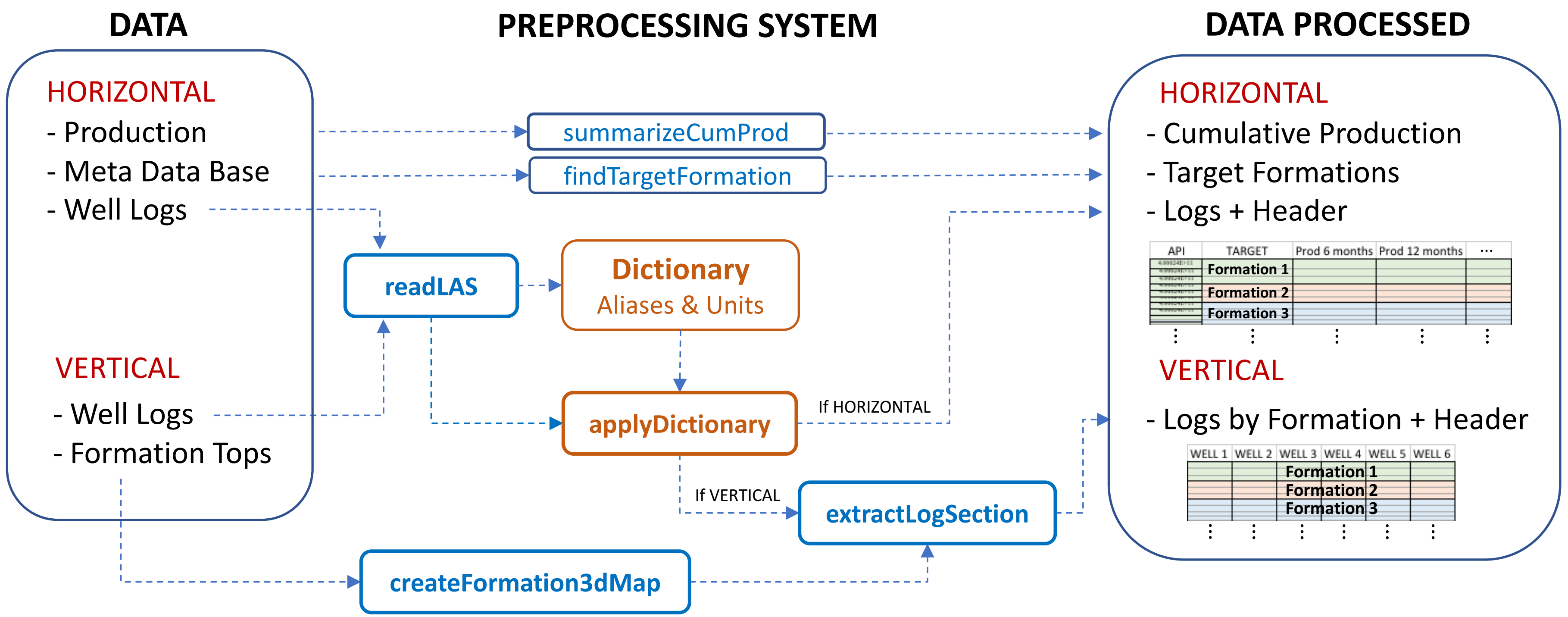
Well log inventory:

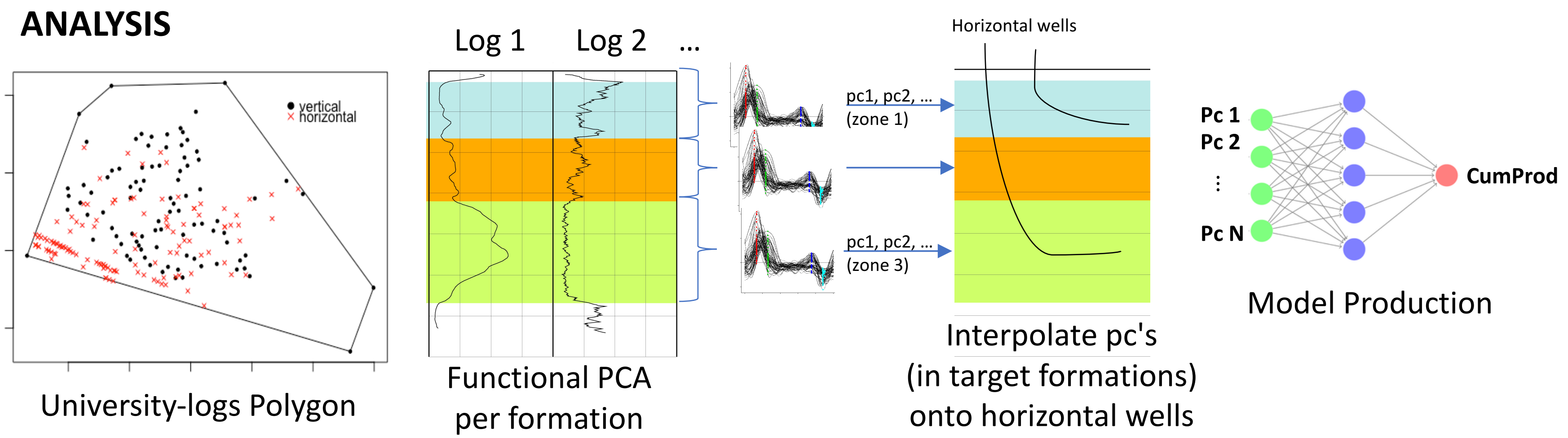
API_10	Surface_X	Surface_Y	Producing_Formation	fileName
4230131393	1013502	761900	Deep Strat	42301313930000_APC_FEE_62-1_LQC_FINAL v2.las
4230131393	1013502	761900	Deep Strat	42301313930000_APC_FEE_62-1_LQC_FINAL v2.las
4230131393	1013502	761900	Deep Strat	42301313930000_APC_FEE_62-1_LQC_FINAL v2.las
4230131393	1013502	761900	Deep Strat	42301313930000_APC_FEE_62-1_LQC_FINAL v2.las
4230131393	1013502	761900	Deep Strat	42301313930000_APC_FEE_62-1_LQC_FINAL v2.las
4230131393	1013502	761900	Deep Strat	42301313930000_APC_FEE_62-1_LQC_FINAL v2.las
4230131393	1013502	761900	Deep Strat	42301313930000_APC_FEE_62-1_LQC_FINAL v2.las

Cumulative production data for oil & gas:

API_10	API_14	Surface_X	Surface_Y	Data.Source	Cum_6_months_	Cum_12_mo	Cum_18_mo	Cum_24_mo	Cum_30_mo	Cum_6_mon	Cum_12_mo	Cum_18_mo	Cum_24_mo	Cum_30_mo
4230130437	4.23013E+13	1044732.656	773734.1143	Shell OFM	0.66543	1.20895	1.74895	2.36995	2.93895	NA	NA	NA	NA	NA
4230130585	4.23013E+13	1035214.076	782513.9318	Shell OFM	16.34063	NA	NA	NA	NA	30.7373019	NA	NA	NA	NA
4230130585	4.23013E+13	1035214.076	782513.9318	Shell OFM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
4230131083	4.23013E+13	1052016.69	769603.4465	Shell OFM	1.90263	3.20529	3.38729	3.57829	6.53529	5.40692023	NA	NA	NA	NA
4230131110	4.23013E+13	1052126.039	771776.567	Shell OFM	1.79085	3.90452	4.49352	5.41152	7.84852	5.53849	6.0538284	NA	NA	NA
4230131122	4.23013E+13	1034507.041	780685.0231	Shell OFM	2.17799	4.28907	NA	NA	NA	NA	NA	NA	NA	NA
4230131148	4.23013E+13	1061127.824	777142.2015	Shell OFM	6.48212	8.83125	12.7488	15.35615	17.53648	16.33762	21.27861	23.99656	27.44268	30.52844
4230131158	4.23013E+13	1039479.017	812676.398	Shell OFM	NA	NA	NA	NA	NA	1379.279	NA	NA	NA	NA
4230131161	4.23013E+13	1050668.217	783115.7514	Shell OFM	0.75788	1.11335	1.38203	1.74524	2.05978	614.103	944.43905	1176.67277	1363.35963	1541.6772
4230131167	4.23013E+13	1044228.083	805721.1455	Shell OFM	NA	NA	NA	NA	NA	394.331	495.62573	523.81055	542.73995	NA
4230131192	4.23013E+13	1061155.982	777148.7146	Shell OFM	1.02499	NA	NA	NA	NA	441.27291	507.35401	594.41018	NA	NA
4230131207	4.23013E+13	1055959.572	784502.8174	Shell OFM	4.55615	7.18582	10.10597	14.11541	17.22011	470.69121	654.12669	793.08407	850.46882	961.70588
4230131210	4.23013E+13	1028529.332	765441.0882	Shell OFM	0.74331	NA	NA	NA	NA	2162.8634	2373.03466	2490.38006	2796.01723	3013.78373
4230131217	4.23013E+13	1062761.074	772822.0658	Shell OFM	1.65933	2.49298	4.06812	6.25848	9.64527	122.00323	156.58284	186.80369	214.06038	218.60659
4230131226	4.23013E+13	1055486.574	748398.9698	Shell OFM	NA	NA	NA	NA	NA	549.26824	891.33208	1082.89903	1227.20493	1366.28338
4230131241	4.23013E+13	1054274.608	761994.3754	Shell OFM	NA	NA	NA	NA	NA	125.57218	216.19874	278.12227	332.60451	374.74833
4230131257	4.23013E+13	1056254.67	752828.3862	Shell OFM	NA	NA	NA	NA	NA	193.33675	273.58681	333.66509	370.02211	422.98195
4230131280	4.23013E+13	1016283.796	768600.3716	Shell OFM	NA	NA	NA	NA	NA	545.4319	785.7633	1006.10099	1179.3161	1320.36526
4230131282	4.23013E+13	1021116.203	771654.6673	Shell OFM	NA	NA	NA	NA	NA	767.7283	1004.88229	1109.41825	1196.08679	1271.65001

Methodological Workflow

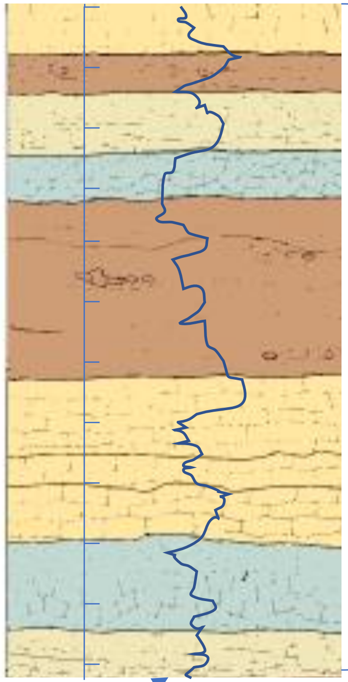


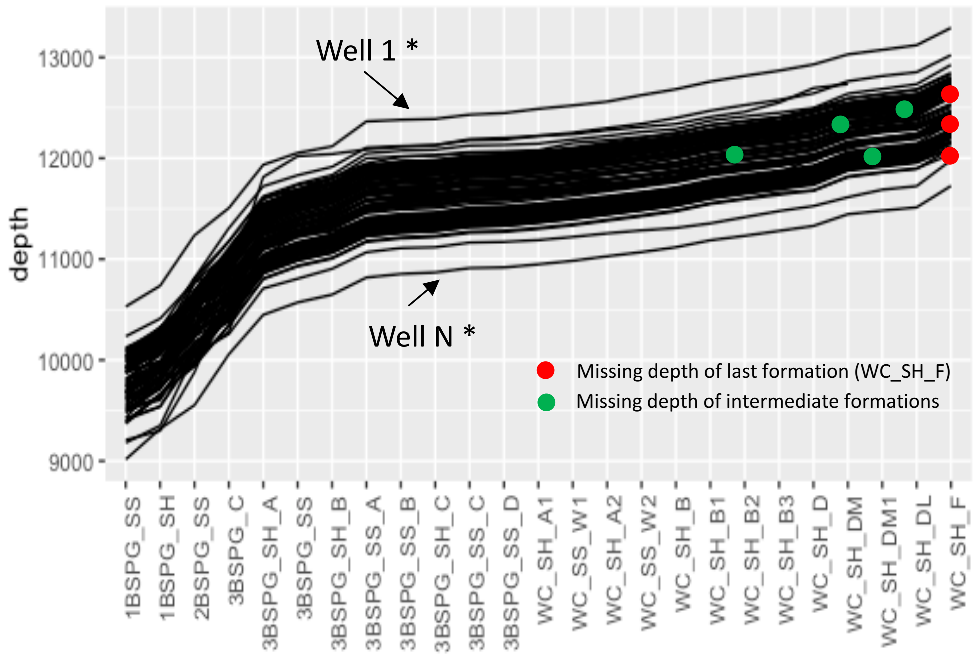


Create (discrete) Formation 3D Map (at vertical wells)

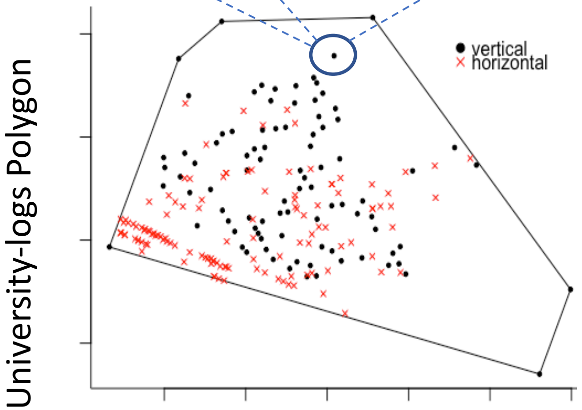
Depth	Formation
10020.83	1BSPG_SS
10294.21	1BSPG_SH
10650.1	2BSPG_SS
11048.6	3BSPG_C
11342.38	3BSPG_SH_A
⋮	⋮
12450	WC_SH_DL
12605	WC_SH_F

Well Log

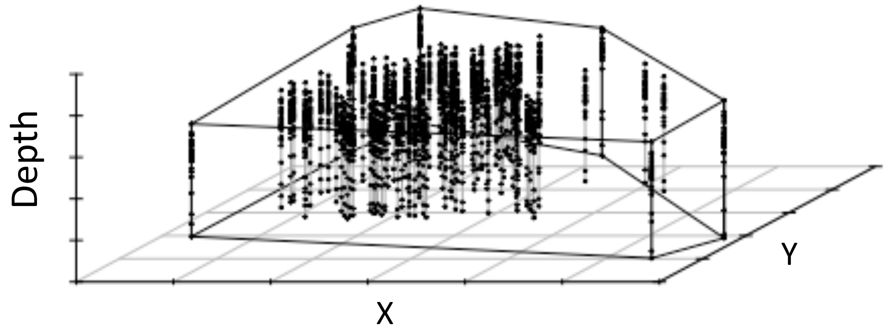




- Missing depth of last formation obtained by Kriging (spatial interpolation)
- Missing depth of intermediate formations obtained by interpolating between formation tops (i.e. interpolate profiles * on above graph)

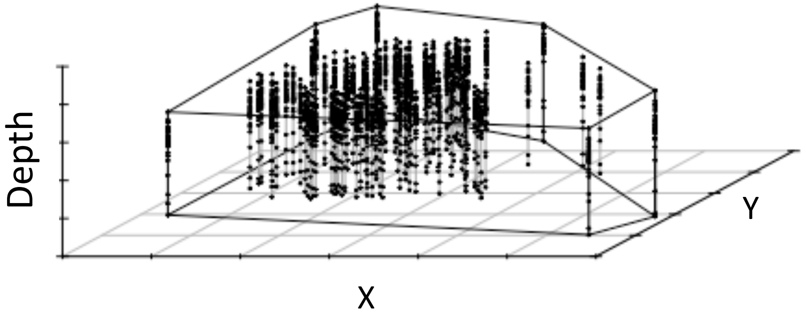


Create 3D map

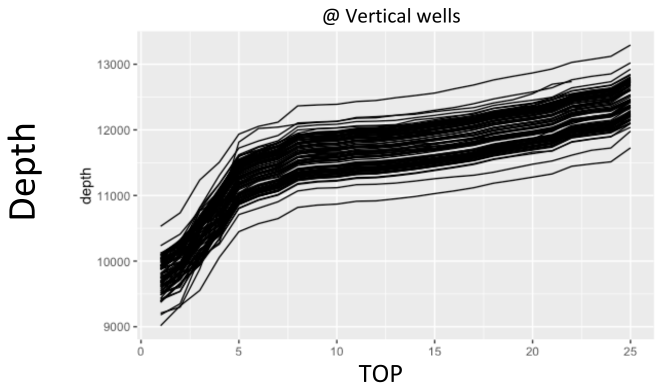


Model Formation 3D Map and interpolate to horizontal wells

1. From previous slide we created a discrete 3D formation map at vertical well locations:



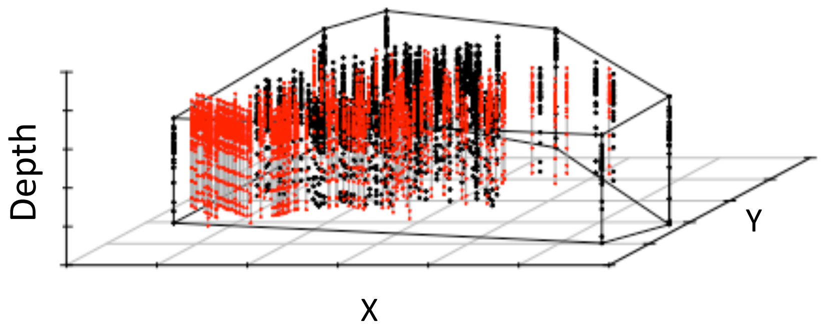
2. We noted that Depth has a clear relationship with Formation Tops at vertical wells:



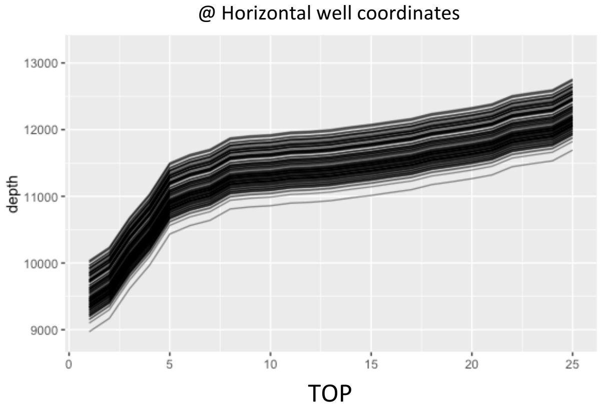
3. Model: $Depth = f_1(TOP) + f_2(X,Y) + Error$

spline

Tensor product spline



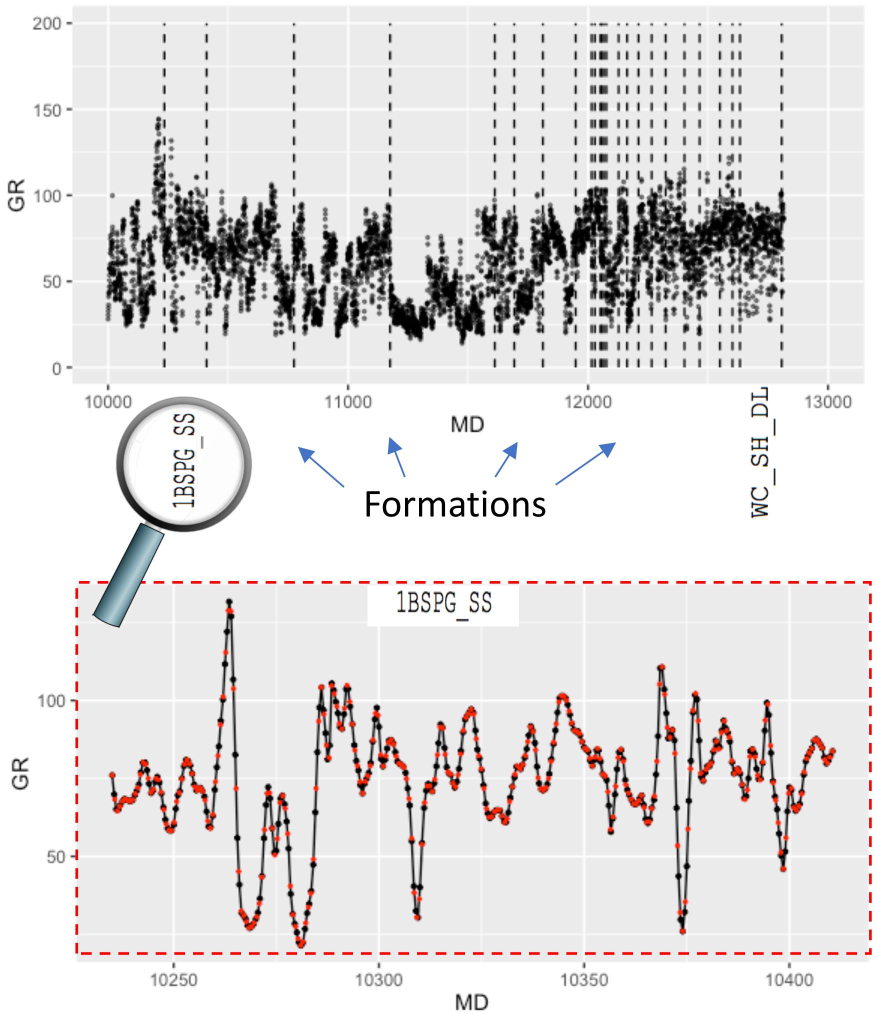
4. 3D formation depths at horizontal well coordinates (red)



5. Depth vs Top at horizontal well coordinates

Extracting Log Sections and normalizing logs by formations

Well 1



Black points: observed well log. Red points: interpolated values at a grid of fixed size n_1.

Normalized well logs (.csv) stratified by formation

Formation	Well 1	Well 2	Well 3	Well 4
1BSPG_SS	76.0215	79.6097	60.0631	85.1999
1BSPG_SS	70.6581053	73.0192982	88.282855	83.8853426
1BSPG_SS	65.8212246	73.1185281	98.9006854	84.2000641
1BSPG_SS	64.6458407	74.9088989	101.772826	82.734702
1BSPG_SS	65.7097931	72.1596746	108.995189	85.2858978
⋮	⋮	⋮	⋮	⋮
1BSPG_SS	81.0441487	79.6692627	82.2807164	79.5946295
1BSPG_SS	82.4105625	72.6555482	99.5709538	72.0131071
1BSPG_SS	83.7583	64.1643	106.0739	59.3989
1BSPG_SH	83.3482	58.7384	103.8597	45.4106
1BSPG_SH	82.6091759	61.7558461	98.5841626	34.4191519

n_1

⋮ ⋮ ⋮ ⋮ ⋮

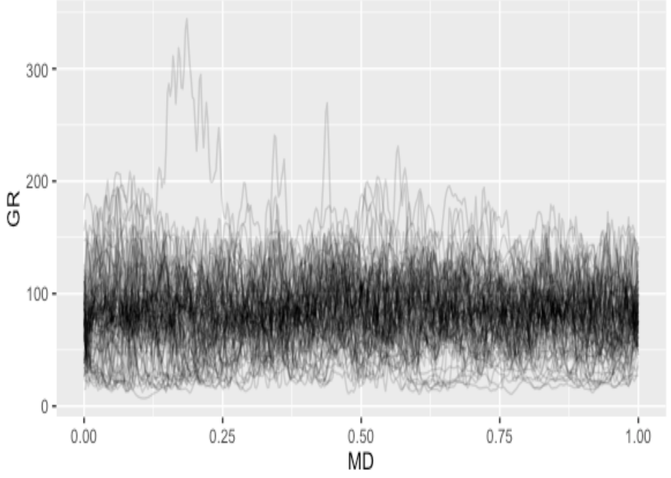
WC_SH_DL	99.9735297	73.6045913	40.7521327	33.0807743
WC_SH_DL	101.065521	73.2267929	54.0674707	31.419708
WC_SH_DL	100.684738	74.1123007	76.5628082	32.9551831
WC_SH_DL	95.1405	68.7398	98.6819	51.9163

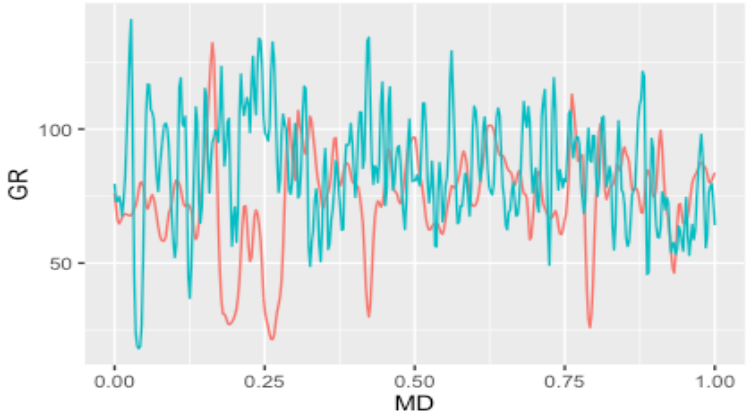
Extracting Features from well logs in each formation (fPCA)

Normalized well logs stratified by formation

Formation	Well 1	Well 2	Well 3	Well 4
1BSPG_SS	76.0215	79.6097	60.0631	85.1999
1BSPG_SS	70.6581053	73.0192982	88.282855	83.8853426
1BSPG_SS	65.8212246	73.1185281	98.9006854	84.2000641
1BSPG_SS	64.6458407	74.9088989	101.772826	82.734702
1BSPG_SS	65.7097931	72.1596746	108.995189	85.2858978
⋮	⋮	⋮	⋮	⋮
1BSPG_SS	81.0441487	79.6692627	82.2807164	79.5946295
1BSPG_SS	82.4105625	72.6555482	99.5709538	72.0131071
1BSPG_SS	83.7583	64.1643	106.0739	59.3989
1BSPG_SH	83.3482	58.7384	103.8597	45.4106
1BSPG_SH	82.6091759	61.7558461	98.5841626	34.4191519

90 University-logs



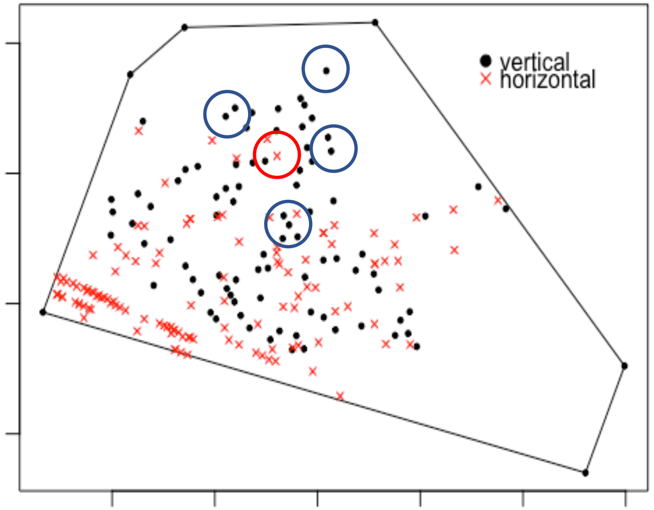


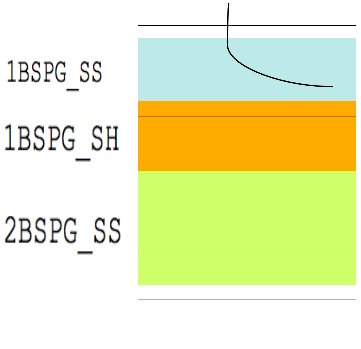
Example of two well logs

Principal component scores in each formation

Formation	Principal Component	Well 1	Well 2	Well 3	Well 4
1BSPG_SS	PC1 (GR)	1.242	0.924	0.744	1.3
1BSPG_SS	PC2 (GR)	1.143	1.203	1.184	1.402
1BSPG_SS	PC3 (GR)	1.302	1.1	0.789	1.203
1BSPG_SS	PC1 (RESDEP)	1.11	1.234	0.0384	0.967
1BSPG_SS	PC2 (RESDEP)	0.998	1.034	0.733	0.893
⋮	⋮	⋮	⋮	⋮	⋮

Interpolate Features (PCs) onto horizontal well coordinates





Interpolate principal components (from formation) at vertical wells (blue circles) onto horizontal wells (red circle) that land in target formation.

Principal component scores in each formation

Formation	Principal Component	Well 1	Well 2	Well 3	Well 4
1BSPG_SS	PC1 (GR)	1.242	0.924	0.744	1.3
1BSPG_SS	PC2 (GR)	1.143	1.203	1.184	1.402
1BSPG_SS	PC3 (GR)	1.302	1.1	0.789	1.203
1BSPG_SS	PC1 (RESDEP)	1.11	1.234	0.0384	0.967
1BSPG_SS	PC2 (RESDEP)	0.998	1.034	0.733	0.893

Interpolated principal components merged with Production Data

API	TARGET	Prod 6 months	Prod 12 months	PC1	PC2	PC3
		Formation 1				
		Formation 2				
		Formation 3				

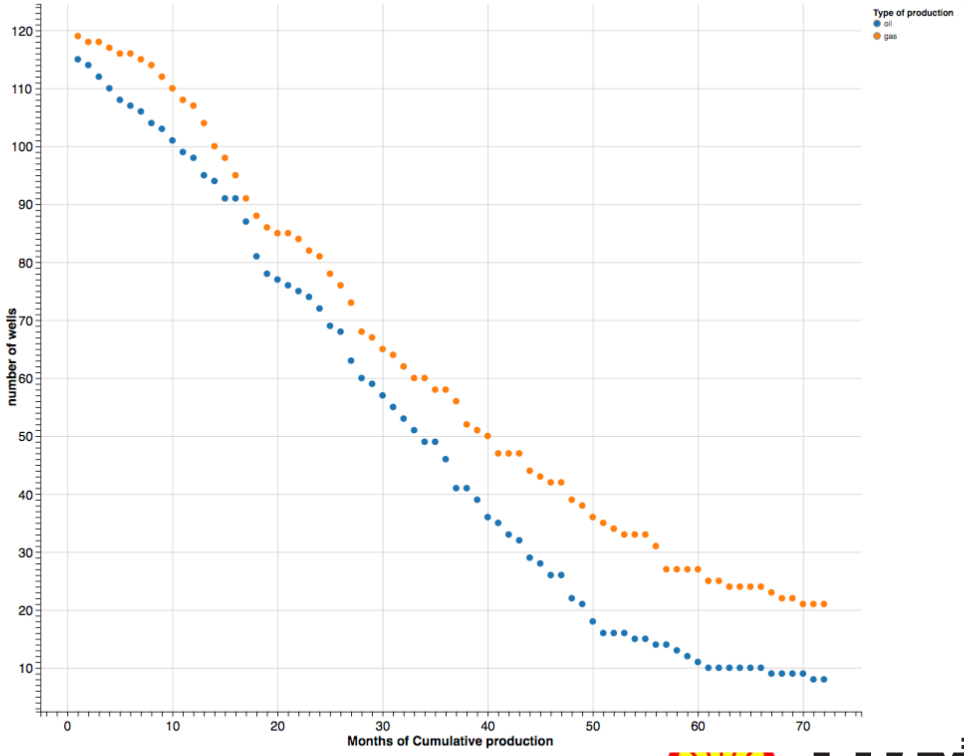
Modeling Production using Principal Component Features

- We observed 98 producing horizontal wells in the University-logs polygon.

- Calculated 10 principal components for vertical wells in each formation for:
  DEN, DENCOR, GR, NEU_LIM, RESDEP, RESSLW, PEF, RESMED, DTC, and DTS

- Interpolated principal components from appropriate target formations onto the corresponding horizontal well coordinates.

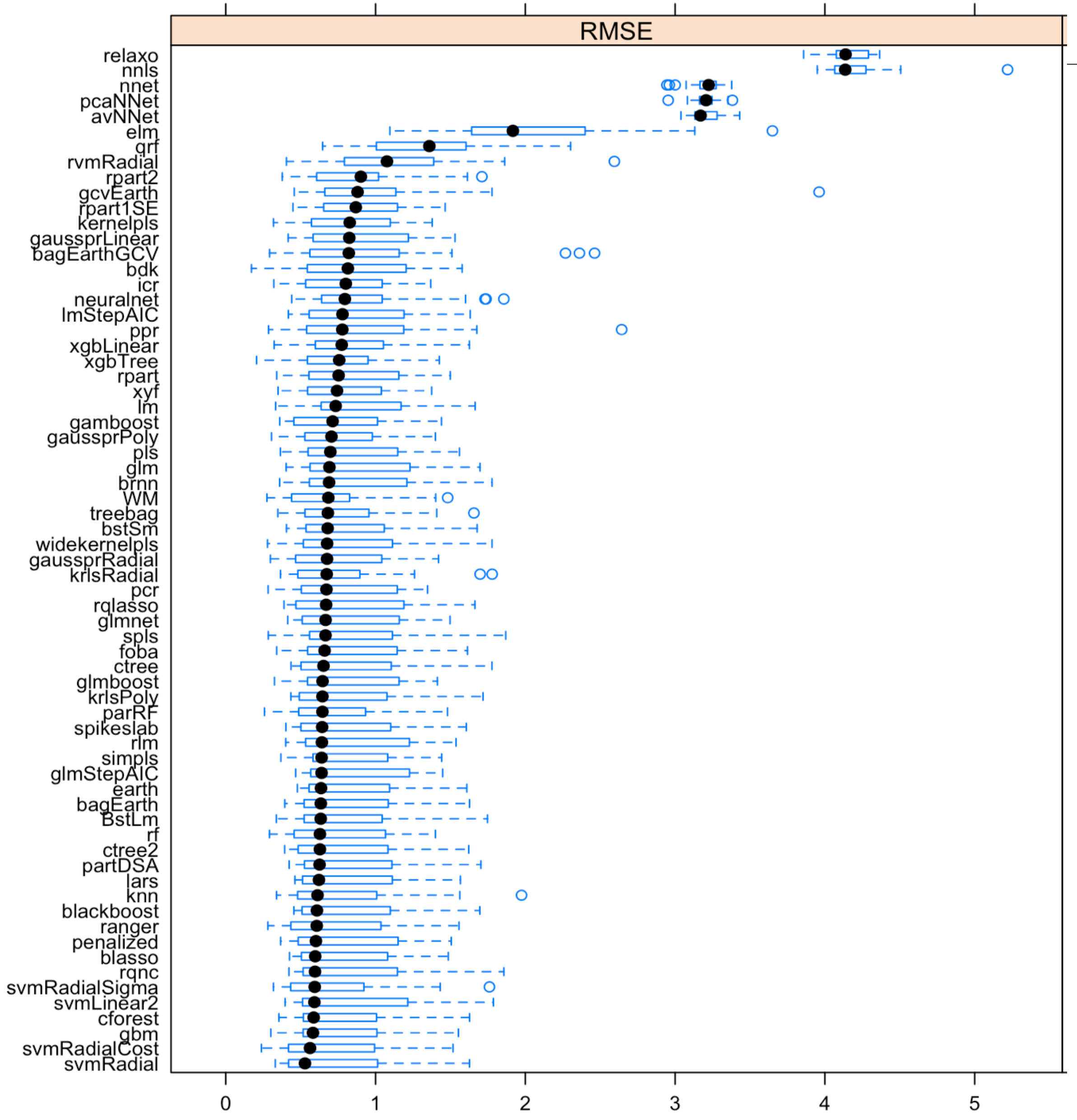
- Cumulative Production (12 months).



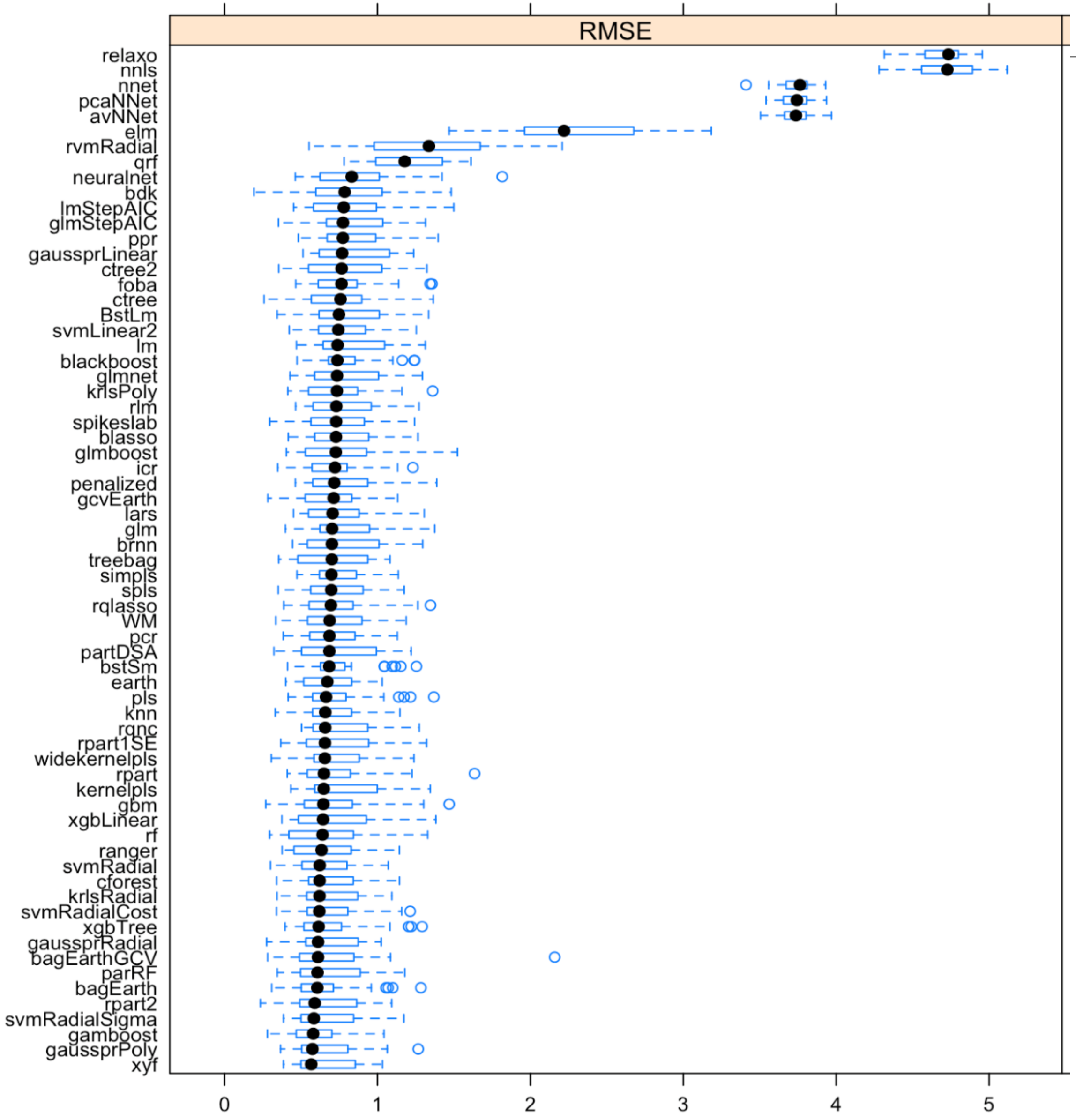
Model selection: repeated 10cv Oil

Experimental setting:

- features = {fPCA's}
- Predicted Variable
  - 12 months of Cumulative production (LOG)
- Model Selection
  - 10 k-fold crossvalidation with 3 repeats
- Preprocessing
  - center and scale
- Metric
  - RMSE
- Comparison of ML models based on Hothorn et al. (2005) and Eugster et al. (2008).



Model selection: repeated 10cv Gas



Experimental setting:

- features = {fPCA’s}
- Predicted Variable
  - 12 months of Cumulative production (LOG)
- Model Selection
  - 10 k-fold crossvalidation with 3 repeats
- Preprocessing
  - center and scale
- Metric
  - RMSE
- Comparison of ML models based on Hothorn et al. (2005) and Eugster et al. (2008).

List of ML models

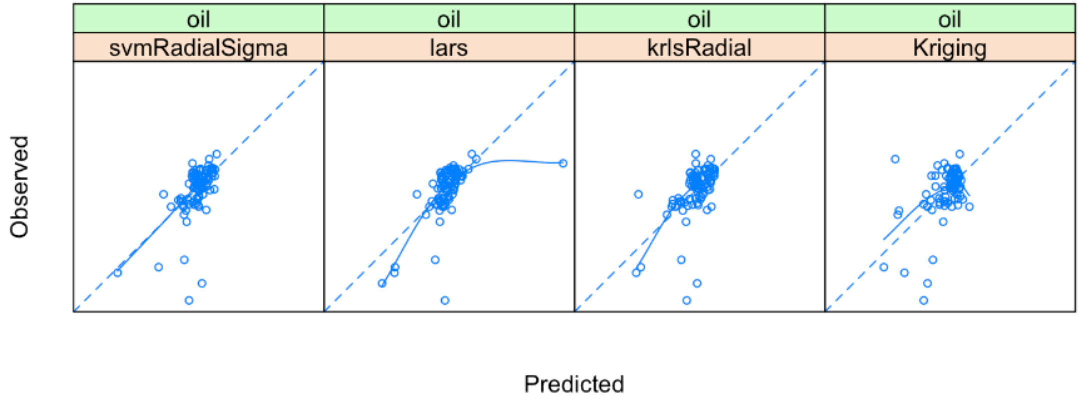
```
"gbm",            #Stochastic Gradient Boosting
"brnn",           #Bayesian Regularized Neural Networks
"icr",            #Independent Component Regression
"lars",           #Least Angle Regression
"lm",             #Linear Regression
"lmStepAIC",      #Linear Regression with Stepwise Selection
"neuralnet",      #Neural Network
"rqnc",           #Non-Convex Penalized Quantile Regression
"nnls",           #Non-Negative Least Squares
"penalized",      #Penalized Linear Regression
"krlsPoly",       #Polynomial Kernel Regularized Least Squares
"pcr",            #Principal Component Analysis
"ppr",            #Projection Pursuit Regression
"qrf",            #Quantile Random Forest
"rqlasso",        #Quantile Regression with LASSO penalty
"krlsRadial",     #Radial Basis Function Kernel Regularized Least Squares
"relaxo",          #Relaxed Lasso
"rvmRadial",      #Relevance Vector Machines with Radial Basis Function Kernel
"foba",           #Ridge Regression with Variable Selection
"rlm",            #Robust Linear Model
"spikeslab",      #Spike and Slab Regression
"blasso",         #The Bayesian lasso
"WM",             #Wang and Mendel Fuzzy Rules
"treebag",        #Bagged CART
"bagEarth",       #Bagged MARS
"bagEarthGCV",    #Bagged MARS using gCV Pruning
"gamboost",       #Boosted Generalized Additive Model
"glmboost",       #Boosted Generalized Linear Model
"BstLm",          #Boosted Linear Model
"bstSm",          #Boosted Smoothing Spline
"blackboost",     #Boosted Tree
"rpart",          #CART
"rpart1SE",       #CART
"rpart2",         #CART
"cforest",        #Conditional Inference Random Forest
"ctree",          #Conditional Inference Tree
"ctree2",         #Conditional Inference Tree
"elm",            #Extreme Learning Machine
"gaussprLinear",  #Gaussian Process
"gaussprPoly",    #Gaussian Process with Polynomial Kernel
"gaussprRadial",  #Gaussian Process with Radial Basis Function Kernel
```

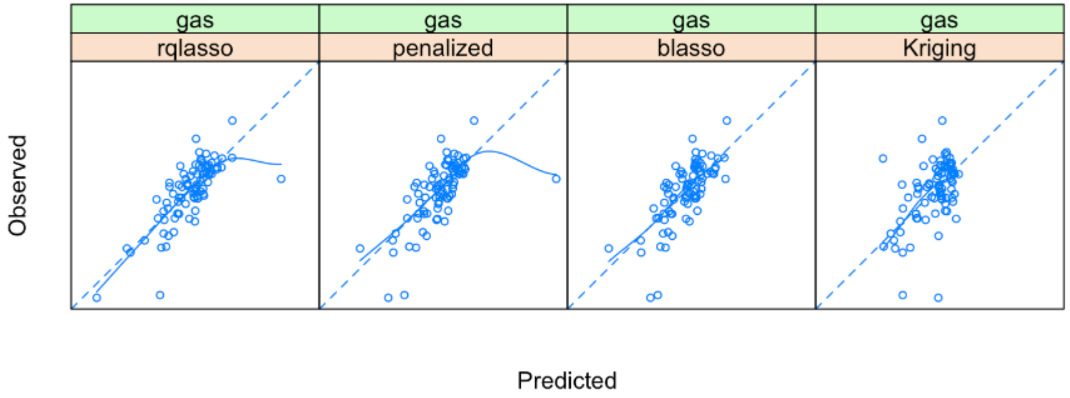
```
"glm",                #Generalized Linear Model
"glmStepAIC",         #Generalized Linear Model with Stepwise Feature Selection
"avNNet",             #Model Averaged Neural Network
# "mlp",                #Multi-Layer Perceptron  (slow)
#"mlpWeightDecay",    #Multi-Layer Perceptron  (slow)
#"mlpWeightDecayML",   #Multi-Layer Perceptron, multiple layers  (slow)
#"mlpML",               #Multi-Layer Perceptron, with multiple layers  (slow)
"earth",                #Multivariate Adaptive Regression Spline
"gcvEarth",            #Multivariate Adaptive Regression Splines
"nnet",                #Neural Network
"pcaNNet",             #Neural Networks with Feature Extraction
"parRF",               #Parallel Random Forest
"kernelpls",           #Partial Least Squares
"pls",                 #Partial Least Squares
"simpls",              #Partial Least Squares
"widekernelpls",       #Partial Least Squares
#"rbfDDA",              #Radial Basis Function Network (slow)
"ranger",              #Random Forest
"rf",                  #Random Forest
"bdk",                 #Self-Organizing Map
"xyf",                 #Self-Organizing Maps
"spls",               #Sparse Partial Least Squares
"svmLinear2",         #Support Vector Machines with Linear Kernel
#"svmLinear",          #Support Vector Machines with Linear Kernel
#"svmPoly",            #Support Vector Machines with Polynomial Kernel
"svmRadial",          #SVM with RBF Kernel
"svmRadialCost",      #Support Vector Machines with Radial Basis Function Kernel
"svmRadialSigma",     #Support Vector Machines with Radial Basis Function Kernel
"xgbLinear",           #eXtreme Gradient Boosting
"xgbTree",             #eXtreme Gradient Boosting
"glmnet",              #glmnet
"knn",                #k-Nearest Neighbors
"partDSA"             #partDSA
```

Nested Leave one out validation

Oil			
Method	Feature Selection	RMSE	Pearson Correlation
svmRadialSigma	Elastic Net	0.723	0.599
lars	built-in	0.724	0.607
krlsRadial	Elastic Net	0.737	0.573
kriging	horizontal production	0.765	0.530

Gas			
Method	Feature Selection	RMSE	Pearson Correlation
rqlasso	built-in	0.522	0.767
blasso	built-in	0.546	0.742
penalized	built-in	0.605	0.679
kriging	horizontal production	0.704	0.507





Summary of Results

- We have developed a workflow in R which currently includes
  - Data pre-processing of vertical well log files (.las) and production data
  - Polygon-based area selection
  - Feature extraction using FPCA
  - Spatial interpolation taking into consideration multiple sub-layers
  - Feature selection and machine learning techniques
- Using this workflow we were able to reproduce the results in IBM's previous work using Shell's data from the University area.

Future Work

- Extracting features from the horizontal well logs
- Consider completion parameters
- Cognitive tools for data integration
  - Cleaning
  - Dictionary
  - (very time consuming process executed by Shell experts)

- Validating reason for missing data in some formations.

- Look for more interpretable machine learning models

- Look for more interpretable geological features

- Experimenting more systematically with different combinations of feature extraction, feature selection and machine learning algorithms.

Paper

Cornell University Library

arXiv.org > cs > arXiv:1705.06556

Search or Article ID inside arXiv | All papers | Broaden

(Help | Advanced search)

Computer Science > Other Computer Science

A data-driven workflow for predicting horizontal well production using vertical well logs

Jorge Guevara, Matthias Kormaksson, Bianca Zadrozny, Ligang Lu, John Tolle, Tyler Croft, Mingqi Wu, Jan Limbeck, Detlef Hohl

(Submitted on 15 May 2017)

In recent work, data-driven sweet spotting technique for shale plays previously explored with vertical wells has been proposed. Here, we extend this technique to multiple formations and formalize a general data-driven workflow to facilitate feature extraction from vertical well logs and predictive modeling of horizontal well production. We also develop an experimental framework that facilitates model selection and validation in a realistic drilling scenario. We present some experimental results using this methodology in a field with 90 vertical wells and 98 horizontal wells, showing that it can achieve better results in terms of predictive ability than kriging of known production values.

Comments:	Part of DM4OG 2017 proceedings (arXiv:1705.03451)
Subjects:	**Other Computer Science (cs.OH)**
Cite as:	**arXiv:1705.06556 [cs.OH]**
	(or **arXiv:1705.06556v1 [cs.OH]** for this version)

Thank You