MSc Petroleum Engineering
Project presentation

Artem Potlog

Depositional facies prediction using machine learning

Heriot-Watt University

School of Energy Geoscience Infrastructure and Society
Institute of GeoEnergy

Prof. Vasily Demyanov
Farah Rabie

OUTLINE

PROBLEM STATEMENT

Manual picks from wireline logs
Different cut-offs & mental rules
Heterogeneous log suites/quality
Hard labels hide uncertainty

Inconsistent labels lead to biased volumes/forecasts

Slow, hard to QC

Problem: Develop and evaluate ML workflow that predicts depositional facies from logs.

Approach: Unsupervised clustering to find the number of facies; supervised classification predict depo. facies

PROBLEM STATEMENT

Manual picks from wireline logs

- Different cut-offs & mental rules
- Heterogeneous log suites/quality
- Hard labels hide uncertainty

→

- Inconsistent labels lead to biased volumes/forecasts
- Slow, hard to QC

ML workflow

- Unsupervised clustering: estimate # of groups; spot overlaps/outliers
- Supervised classification: per-class probabilities

→

- Consistent, probabilistic facies picks
- Explicit uncertainty for QC and re-interpretation

Problem: Develop and evaluate ML workflow that predicts depositional facies from logs.

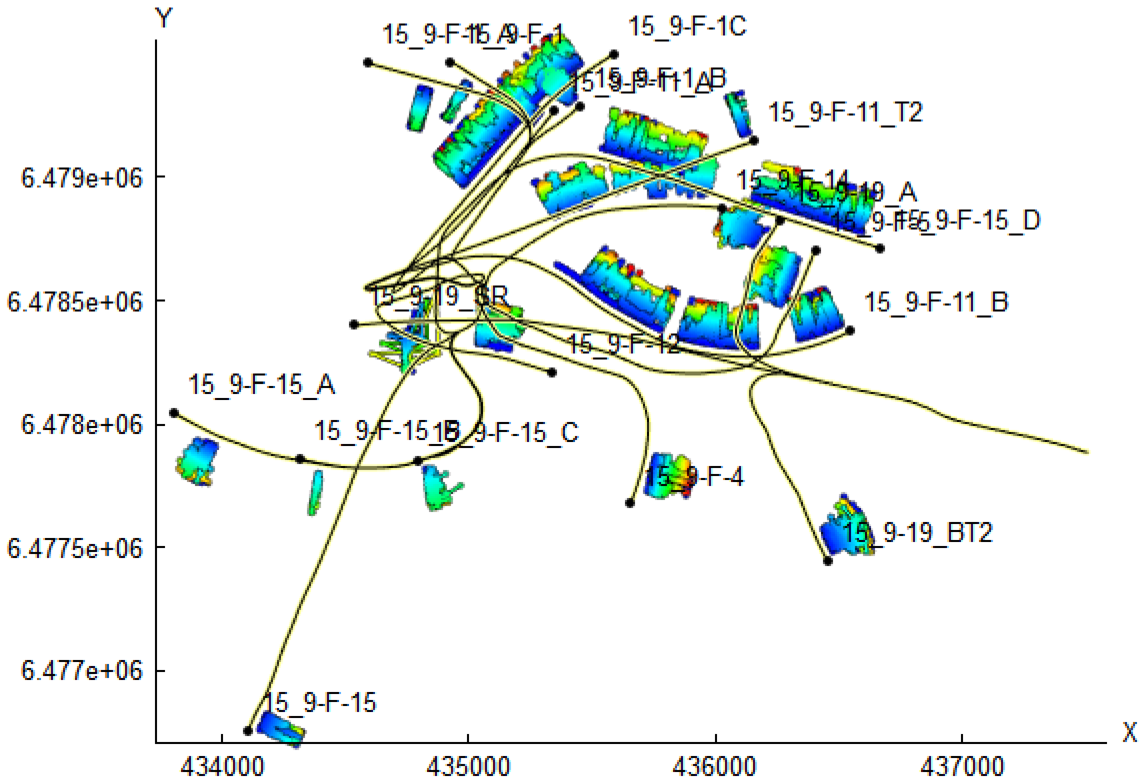
Approach: Unsupervised clustering to find the number of facies; supervised classification predict depo. facies

DATA

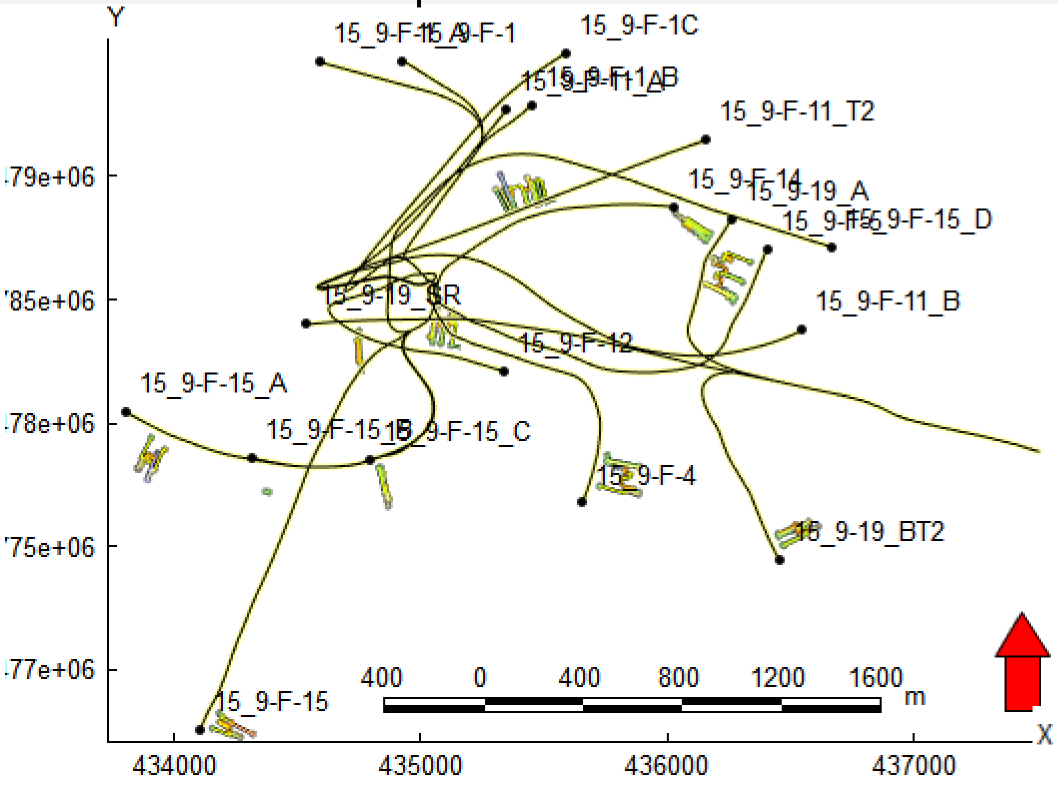
DATA IS SPARSE

Across Volve field wells (Equinor public data set) log coverage is heterogeneous, some curves are missing or partial, core control is scarce.
This uneven coverage produces ambiguous labels and **uncertainty that originates in the data**

Logs

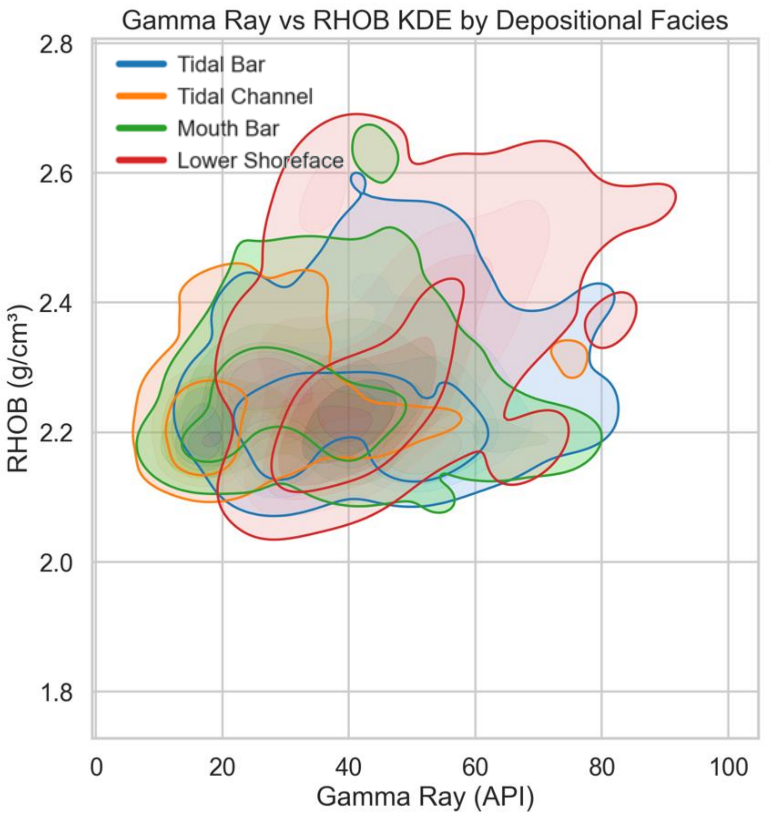


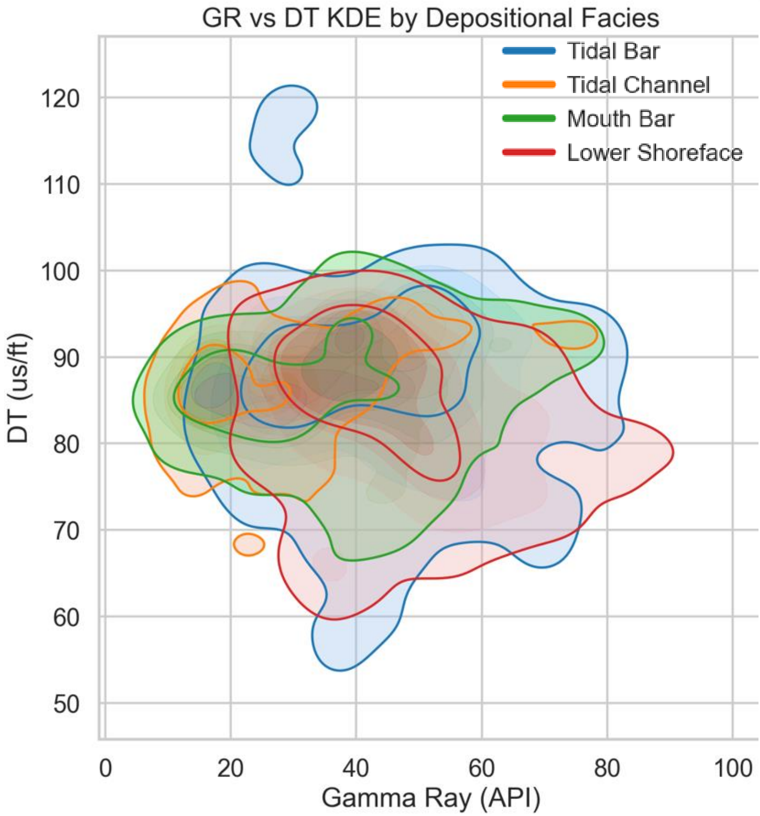
Depositional facies

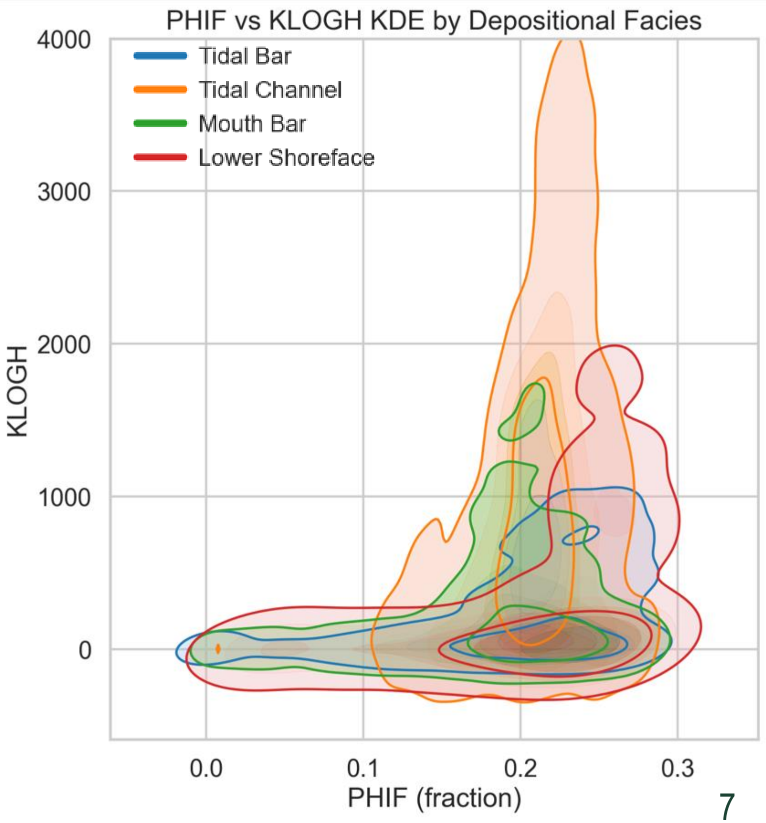


LOG CROSS-PLOTS STRONGLY OVERLAP

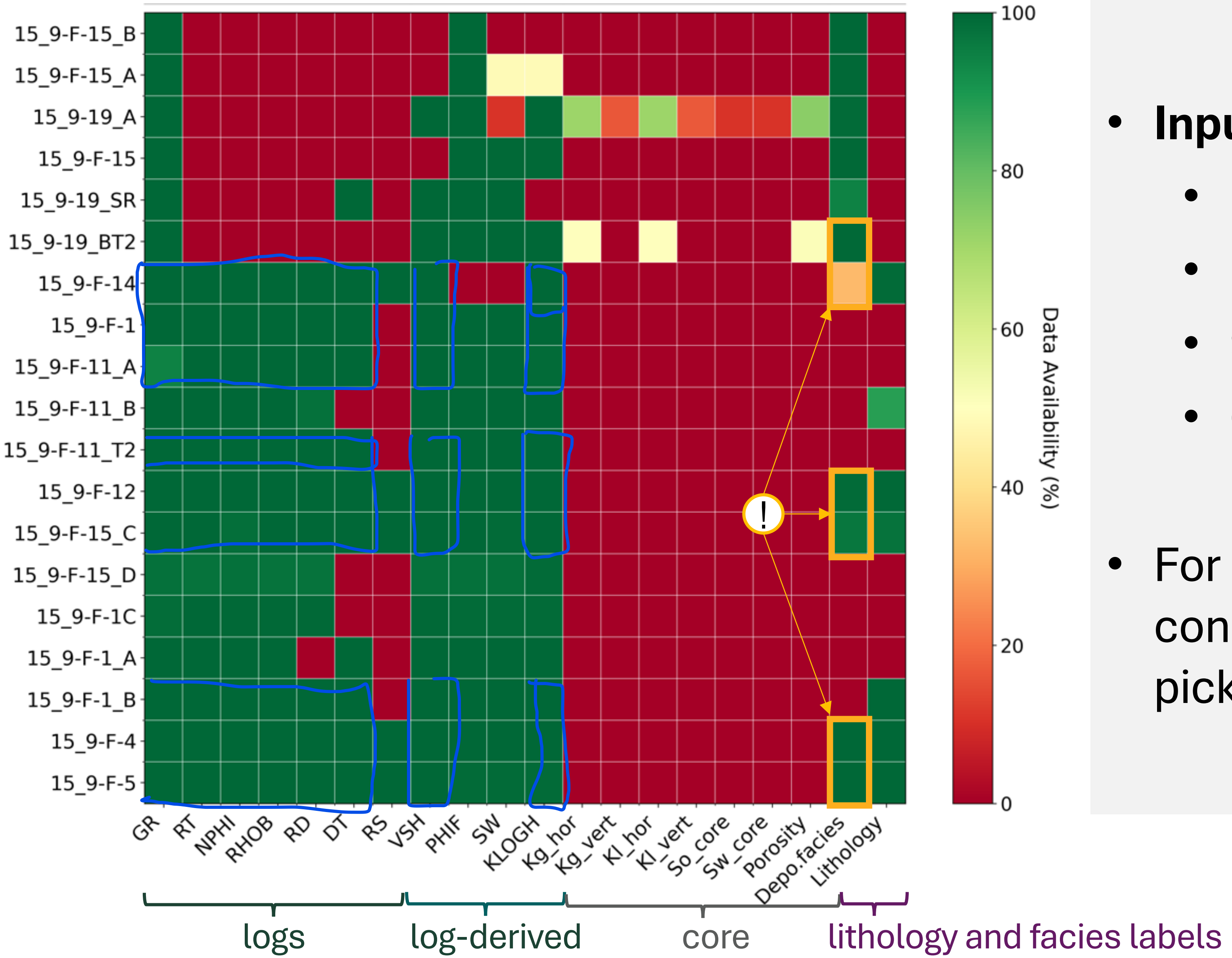
We plotted different logs on cross-plots based on depo. facies and see a strong overlap in all pairs. Same combination can appear facie or another –> part of the uncertainty comes from the logs. For ML this means limited cluster separability







DATA SCARCITY



- **Inputs:**
  - logs;
  - log-derived curves;
  - 9 depo. facies;
  - 6 lithology labels.

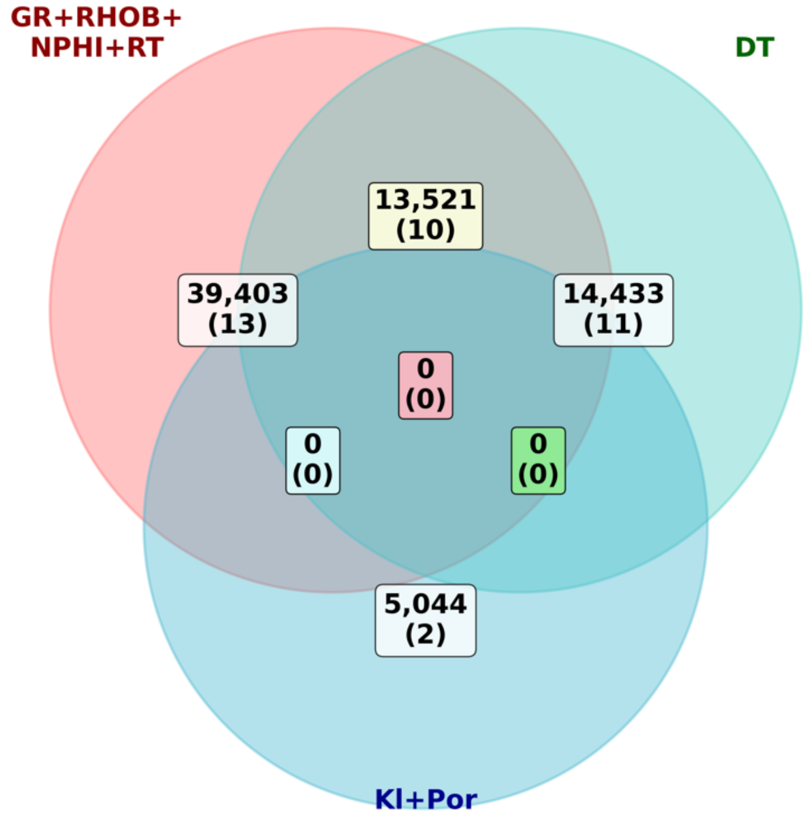
- For modelling we use wells with consistent depo. facies vs logs picks

LITHO AND DEPO FACIES LABELS ARE LIMITED

5k samples in 5 wells (out of 19) are available for unsupervised validation and supervised learning.

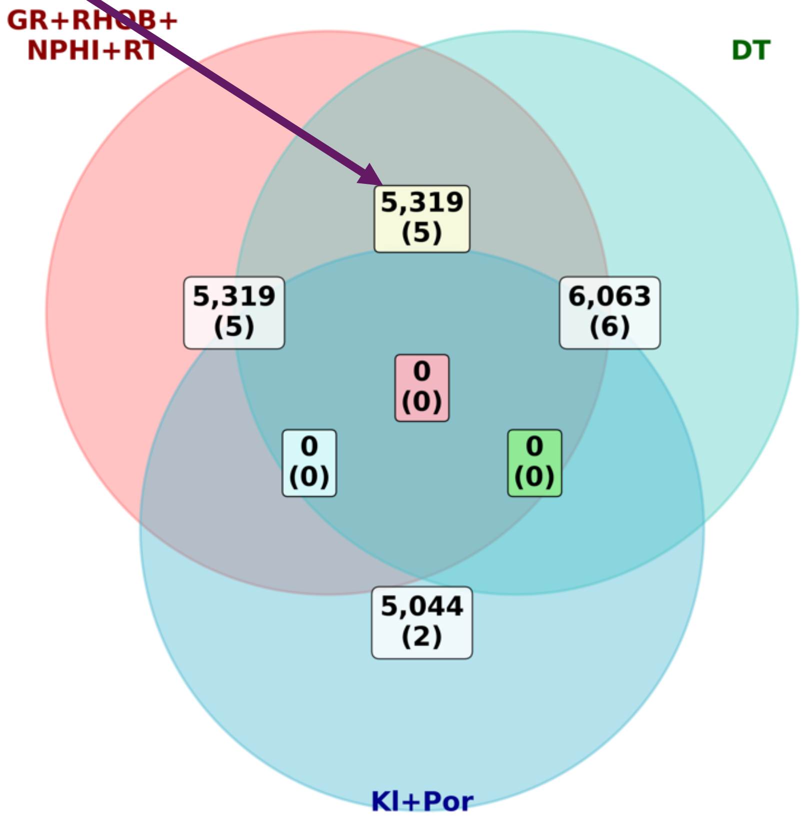
In the project we also test different log combinations sacrificing informativeness (less logs) but increasing available samples.

Without depo. facies and lithology labels



Numbers show rows (wells) for each condition

With depo. facies and lithology labels



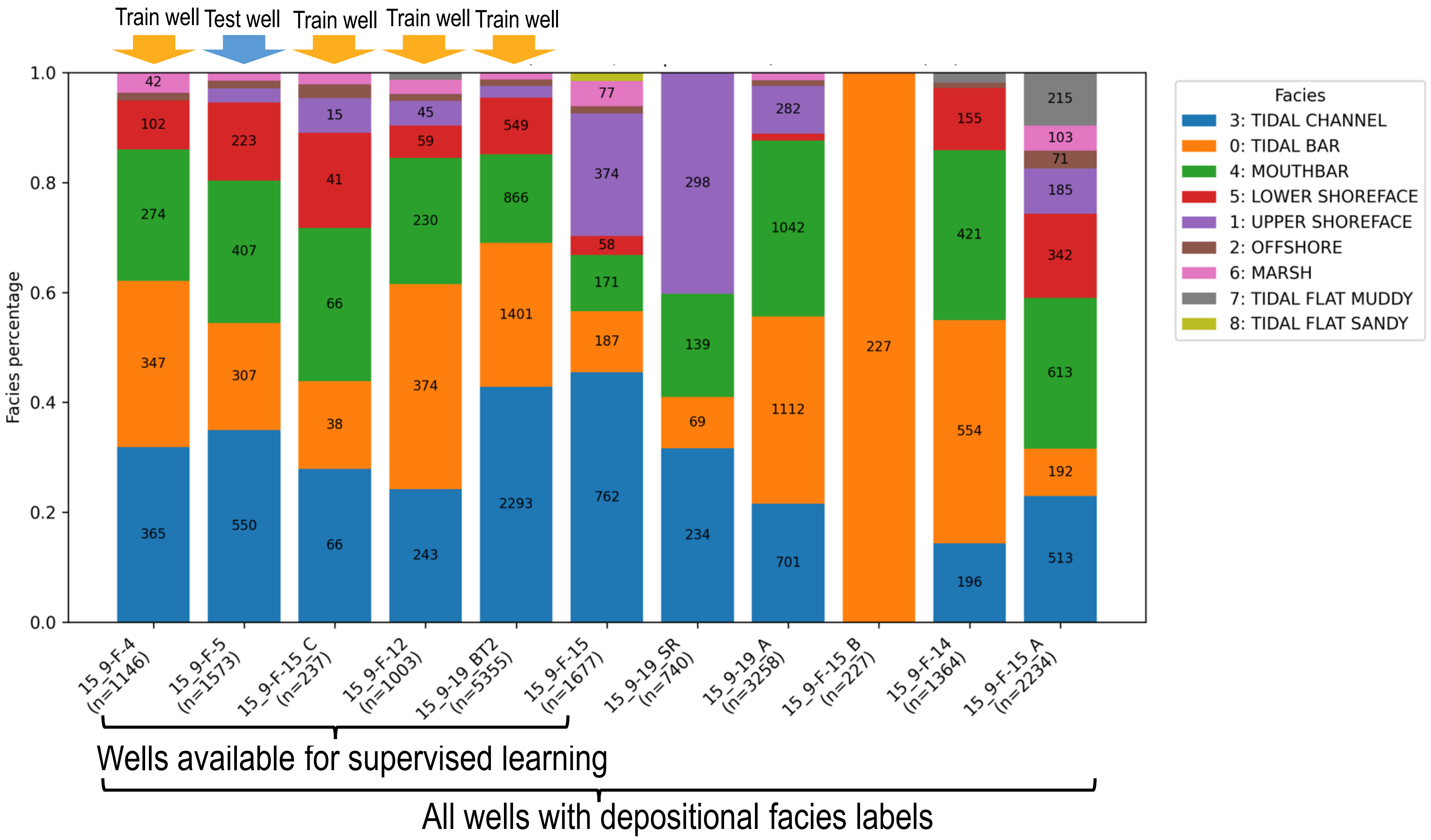
Numbers show rows (wells) for each condition

WELLS PICK FOR TRAINING/TESTING DICTATED BY FACIES REPRESENTATION

Not all wells available for supervised learning have a good representation of depo. facies, which dictates wells pick for training/testing purpose

Well F-5 (or 15_C) is used for testing purpose

Wells F-4, 15_C, F-12, BT2 – training wells



WORKFLOW AND METHODS

WORKFLOW AND METHODS

1) **Clustering** to separate the logs into groups and study if these groups are meaningful;
2) **Classification** to predict depositional facies facies

Clustering

Algorithm: K-means.

Cluster Selection (k): Silhouette score and Elbow chart.

Interpretation: Principal Component

Validation: Against known depositional facies

Classification

Algorithm: Random Forest

Protocol: Training on wells with strong facies representation, testing on held-out wells.

Metrics: Accuracy, Precision, Recall, f1

WORKFLOW AND METHODS

1) **Clustering** to separate the logs into groups and study if these groups are meaningful;
2) **Classification** to predict depositional facies facies

Data preprocessing

Input Features: logs, log-derived values, lithology

Engineered Features: GR Slope, GR Serration, relative position.

Transformations and scaling

KNN imputation for missing values

Categorical lithology one-hot encoded.

Clustering

Algorithm: K-means.

Cluster Selection (k): Silhouette score and Elbow chart.

Interpretation: Principal Component

Validation: Against known depositional facies

Classification

Algorithm: Random Forest

Protocol: Training on wells with strong facies representation, testing on held-out wells.

Metrics: Accuracy, Precision, Recall, f1

CLUSTERING

CLUSTERING OBJECTIVE

Explore structure with clustering. Use unsupervised methods to estimate the number of groups and compare clusters with known facies to spot overlaps and label noise.

QUESTION 1

	Why
1. How clustering separate groups? (Do classes form natural groupings?)	Reveal natural splits and set target granularity

QUESTION 2

	Why
2. How similar or different are these groups?	What drives separability

QUESTION 3

	Why
3. How interpretable are the groups with respect to depositional facies?	Ensure clusters map to geology, not just numerical groupings

CLUSTERING EXPERIMENTS

Questions	Why	Experiments
1. How clustering separate groups? (Do classes form natural groupings?)	Reveal natural splits and set target granularity	Run in different feature sets to study log informativeness, k results, Principal components, confusion matrices
2. How similar or different are these groups?	What drives separability	
3. How interpretable are the groups with respect to depositional facies?	Ensure clusters map to geology, not just numerical groupings	

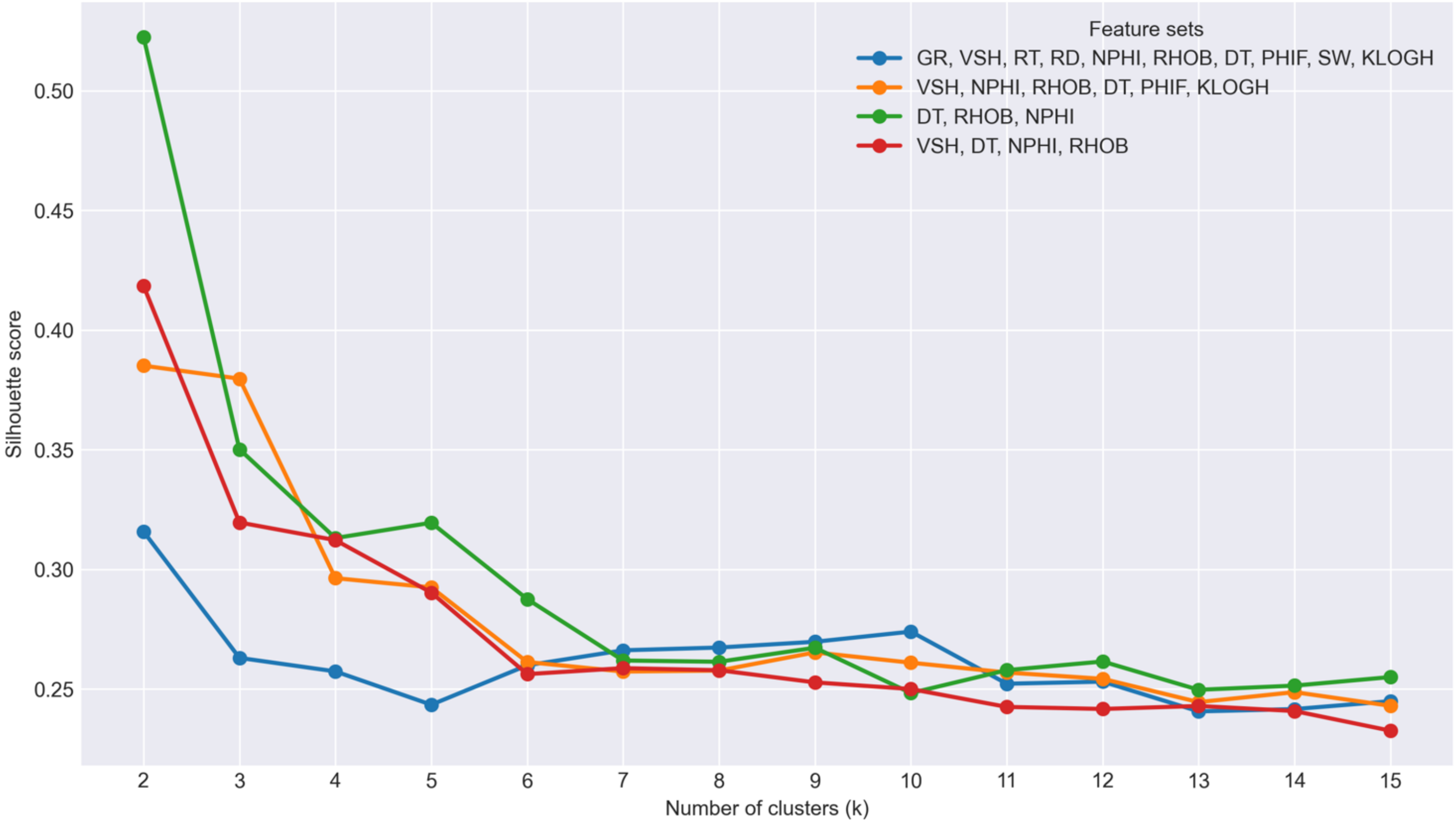
FOUR CLUSTERING EXPERIMENTS OUTLINE

Feature set is: 1) all logs; 2) all logs without fluids; 3) porosity/sorting logs; 4) cleanliness/porosity logs

Experiments	Why
1. All logs: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH	Broadest view of variance
2. All without fluids: GR (and/or VSH), NPHI, RHOB, DT, PHIF, KLOGH	To emphasize rock property space (cleanliness, porosity/sorting, mobility proxy) and de-emphasize fluids
3. Porosity/sorting only: DT, RHOB, NPHI	Check whether groups emerge purely from sorting/porosity differences
4. Cleanliness and porosity: VSH, DT, PHIF (or NPHI and RHOB)	Designed to highlight clean, well-sorted sands versus heterolithic sands

HOW CLUSTERING SEPARATE GROUPS, HOW SIMILAR THEY ARE?

Silhouette score shows that logs form two natural clusters driven by rock properties: clean/porous vs shaly/compact. Beyond two clusters depositional facies do not form distinct natural groupings



1. All logs:
fluid variability cuts across depositional patterns and weakens unsupervised separation.

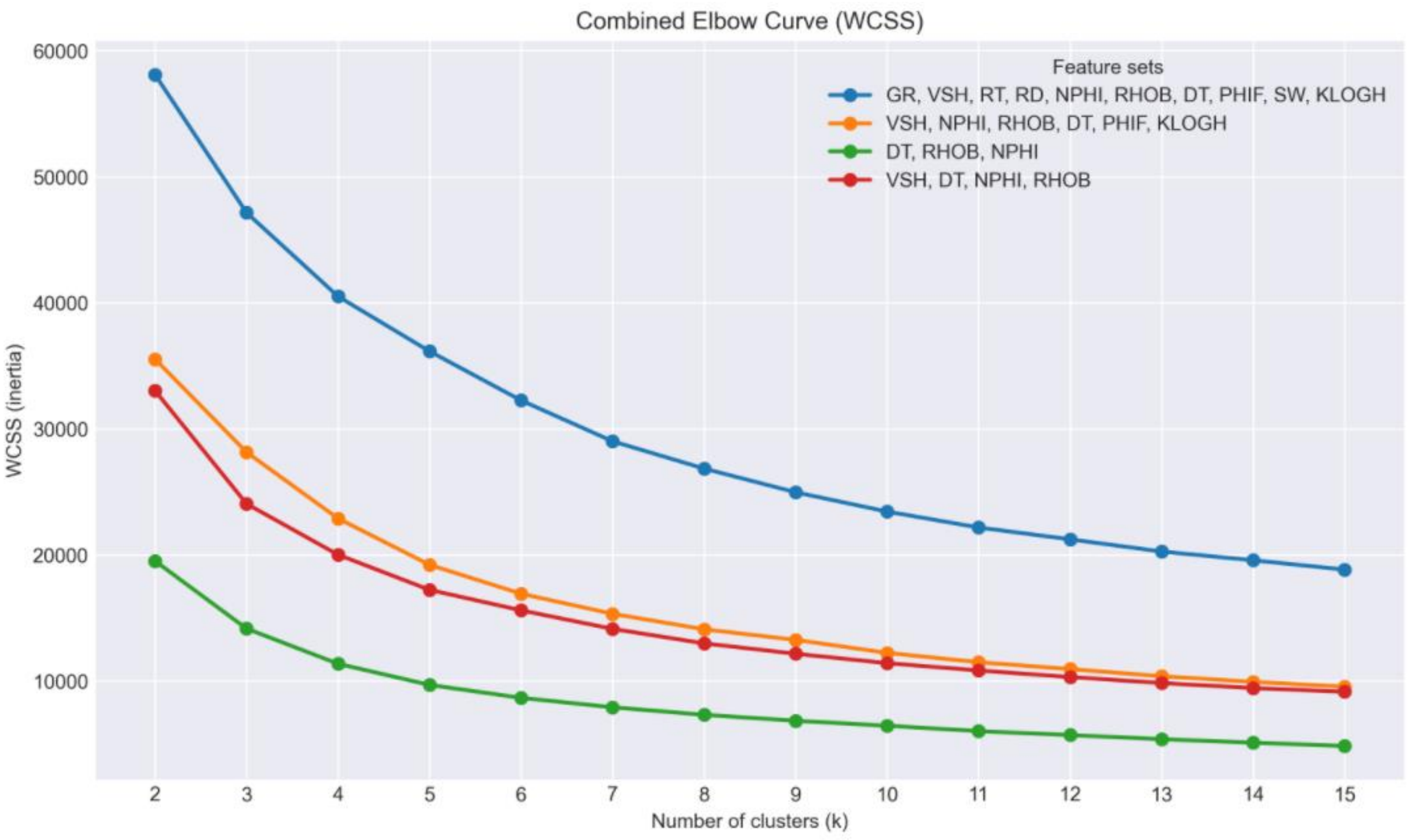
2. VSH, NPHI, RHOB, DT, PHIF, KLOGH:
third, moderate flow-quality mode appears when permeability is added. But sand vs sand distinction remains blended

3. DT, RHOB, NPHI: very clear binary clean-shaly split

4. VSH, DT, NPHI, RHOB: in spaces dominated by porosity/compaction vs cleanliness, the data separate into clean sand vs shaly/tight

HOW CLUSTERING SEPARATE GROUPS, HOW SIMILAR THEY ARE?

From the elbow curve we see that logs without fluids show 2-3 clusters at best



1. All logs

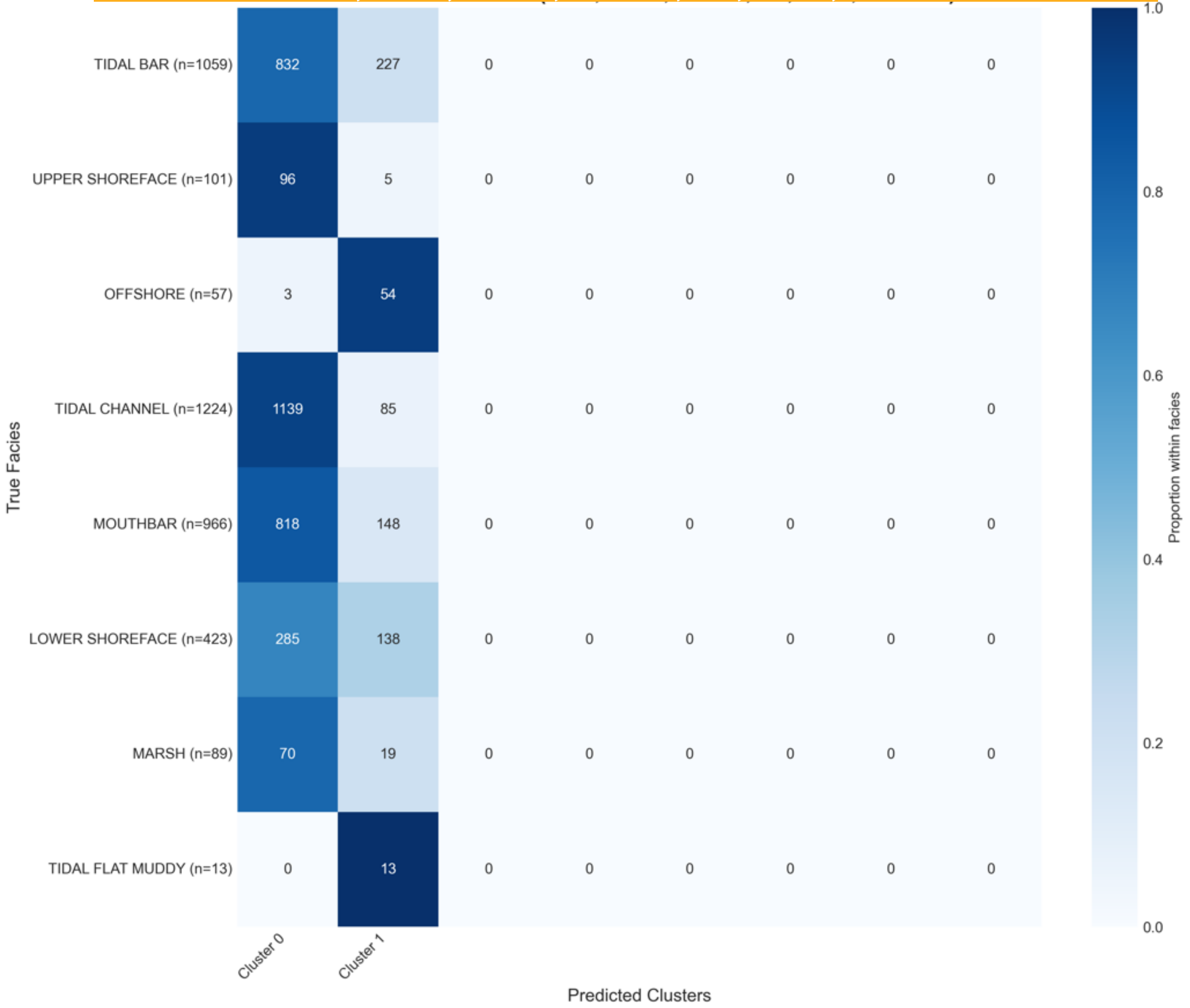
2. VSH, NPHI, RHOB, DT, PHIF, KLOGH

3. DT, RHOB, NPHI

4. VSH, DT, NPHI, RHOB

HOW INTERPRETABLE ARE THE GROUPS?

k = 2, VSH, NPHI, RHOB, DT, PHIF, KLOGH

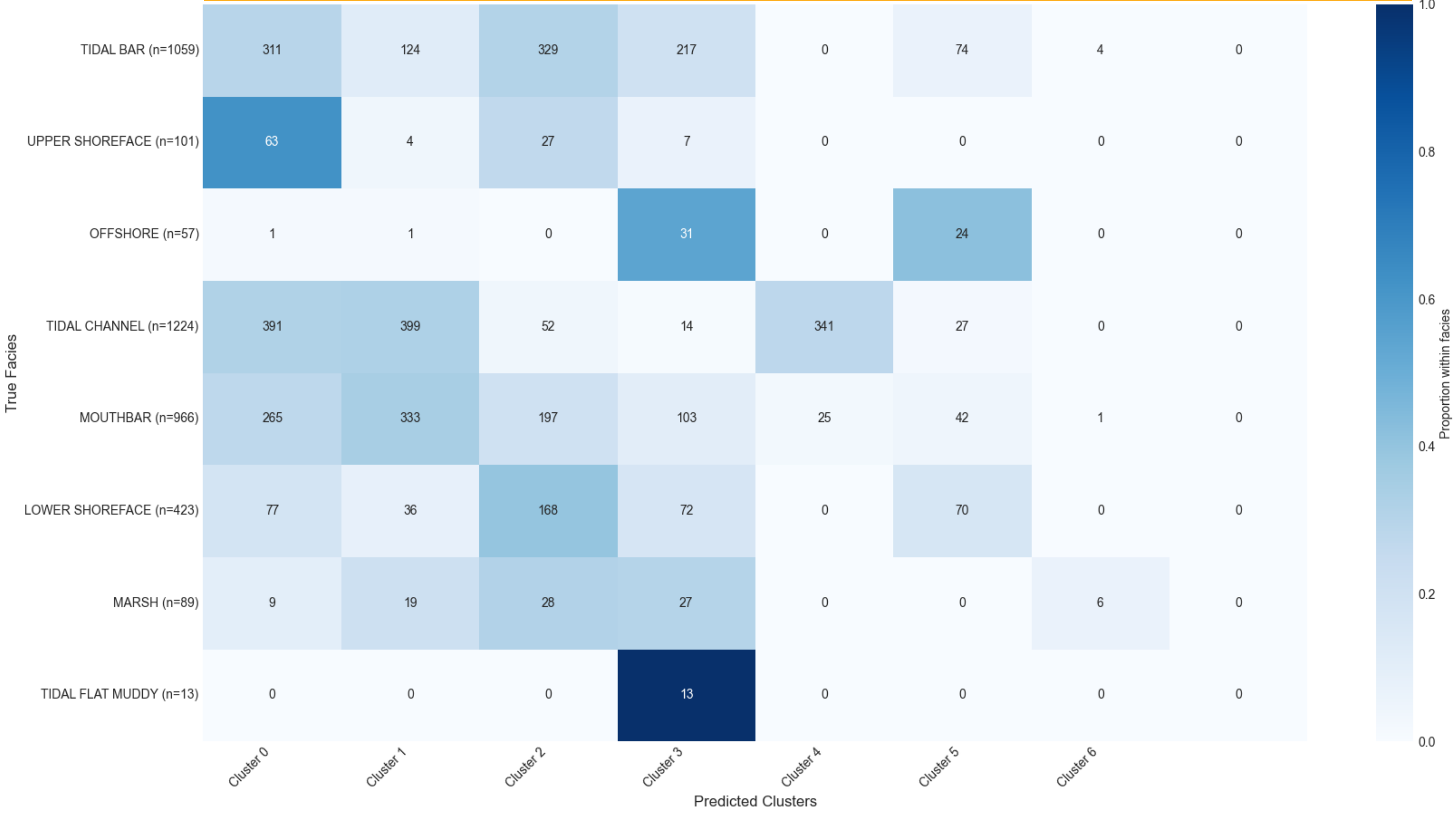


Clusters are geologically interpretable at a coarse level: clean sand vs shaly separation.

On the confusion matrices for k = 2 one cluster concentrates most sand depo. facies (Tidal Bar, Tidal Channel, Mouth Bar, Upper Shoreface), while other captures shaly depo. facies (Offshore, Tidal Flat Muddy).

HOW INTERPRETABLE ARE THE GROUPS?

k = 7, VSH, NPHI, RHOB, DT, PHIF, KLOGH



Increasing number of clusters leads to a poor separation: facies spread across several clusters

CLASSIFICATION

CLASSIFICATION OBJECTIVE

Train supervised models to predict depositional facies; use per-well blind tests to check generalization across wells; optimize accuracy

QUESTION 1

	Why
1. Do different feature sets change performance?	Find which log combinations carry separable signal; quantify trade-offs; test if fluids help beyond porosity/cleanliness.

QUESTION 2

	Why
2. Does adding a lithology feature help?	Lithology narrows classes and should reduce sand vs sand confusion.

QUESTION 3

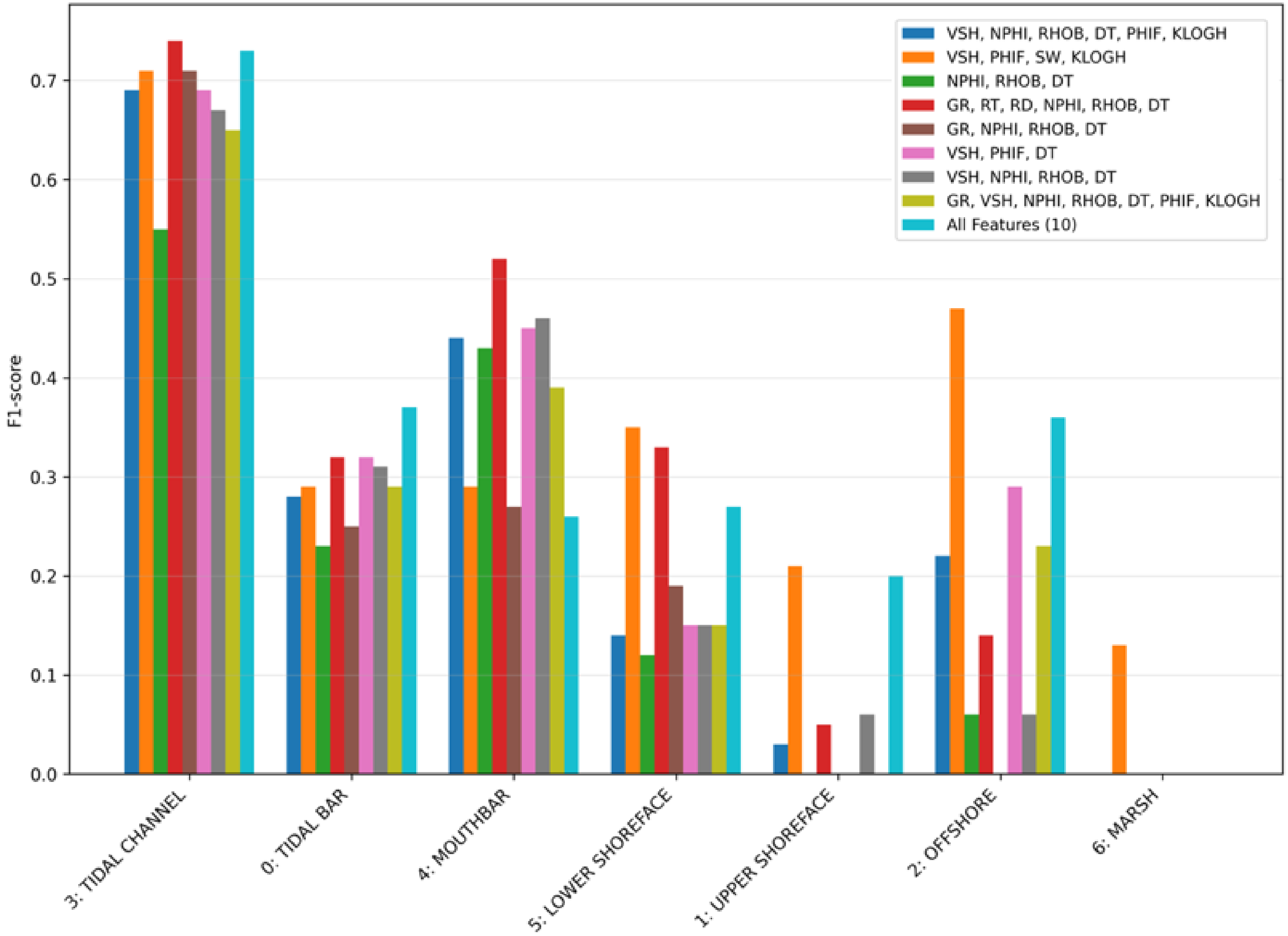
	Why
3. Do engineered features help?	Geology-aware features can surface motifs important in Hugin’s tide-influenced sands

CLASSIFICATION EXPERIMENTS

Questions	Why	Experiments
1. Do different feature sets change performance?	Find which log combinations carry separable signal; quantify trade-offs; test if fluids help beyond porosity/cleanliness.	Run in different feature sets
2. Does adding a lithology feature help?	Lithology narrows classes and should reduce sand vs sand confusion.	For best performing add Lithology
3. Do engineered features help?	Geology-aware features can surface motifs important in Hugin's tide-influenced sands	For best performing w/ Lithology add engineered features

DO DIFFERENT FEATURE SETS CHANGE PERFORMANCE?

Presented are results f1-score per-facie for 9 different features combinations we ran.

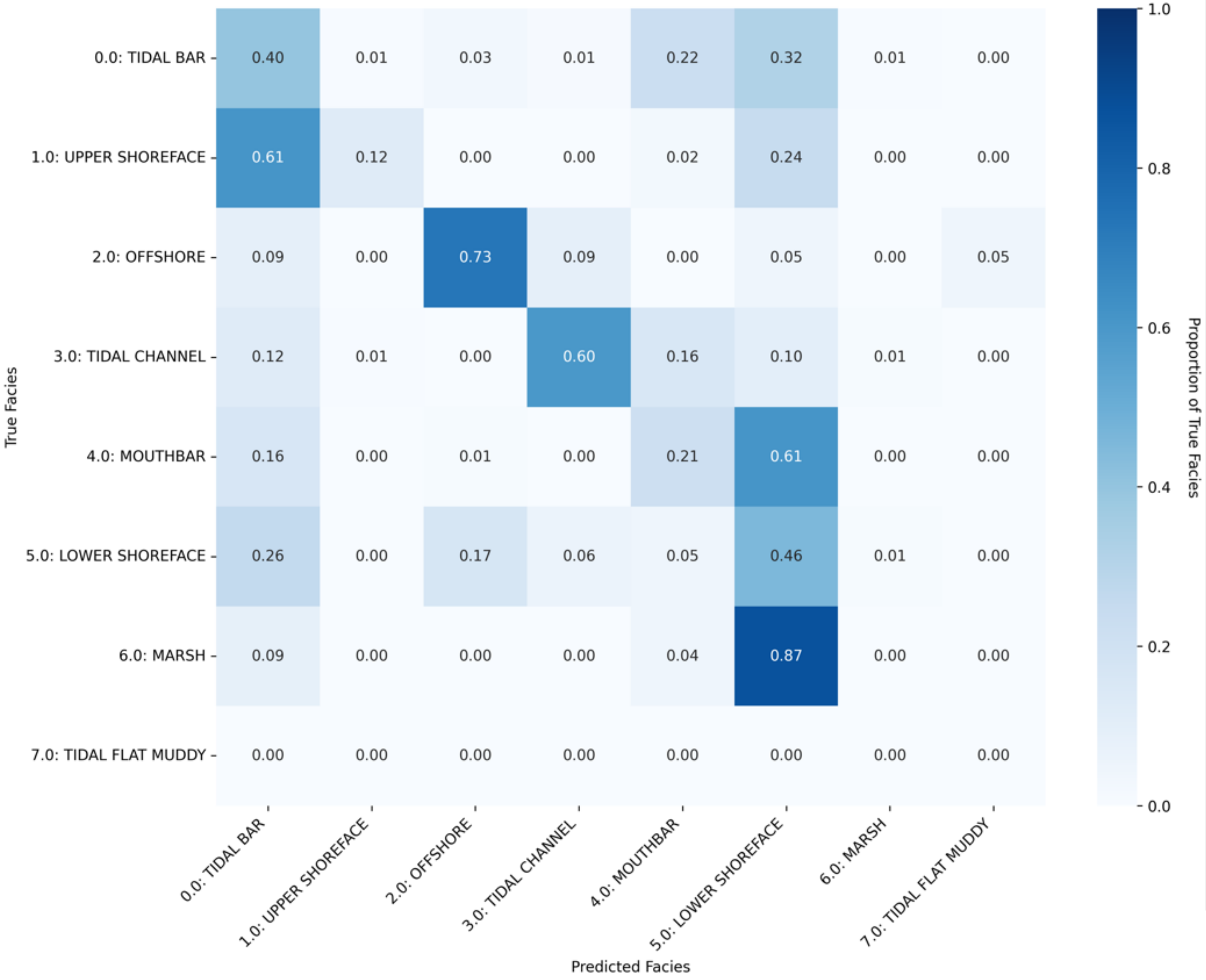


Results: Feature set choice materially changes performance; gains come from combining porosity + cleanliness + fluids

Addition of KLOGH and PHIF introduces better performance in minority classes (higher macro-f1)

MOST CONFUSED CLASSES

All logs set Confusion matrix



In the best performing case most confused pairs of sand vs sand facies:

Tidal Bar vs Mouth Bar

Tidal Bar vs Lower Shoreface

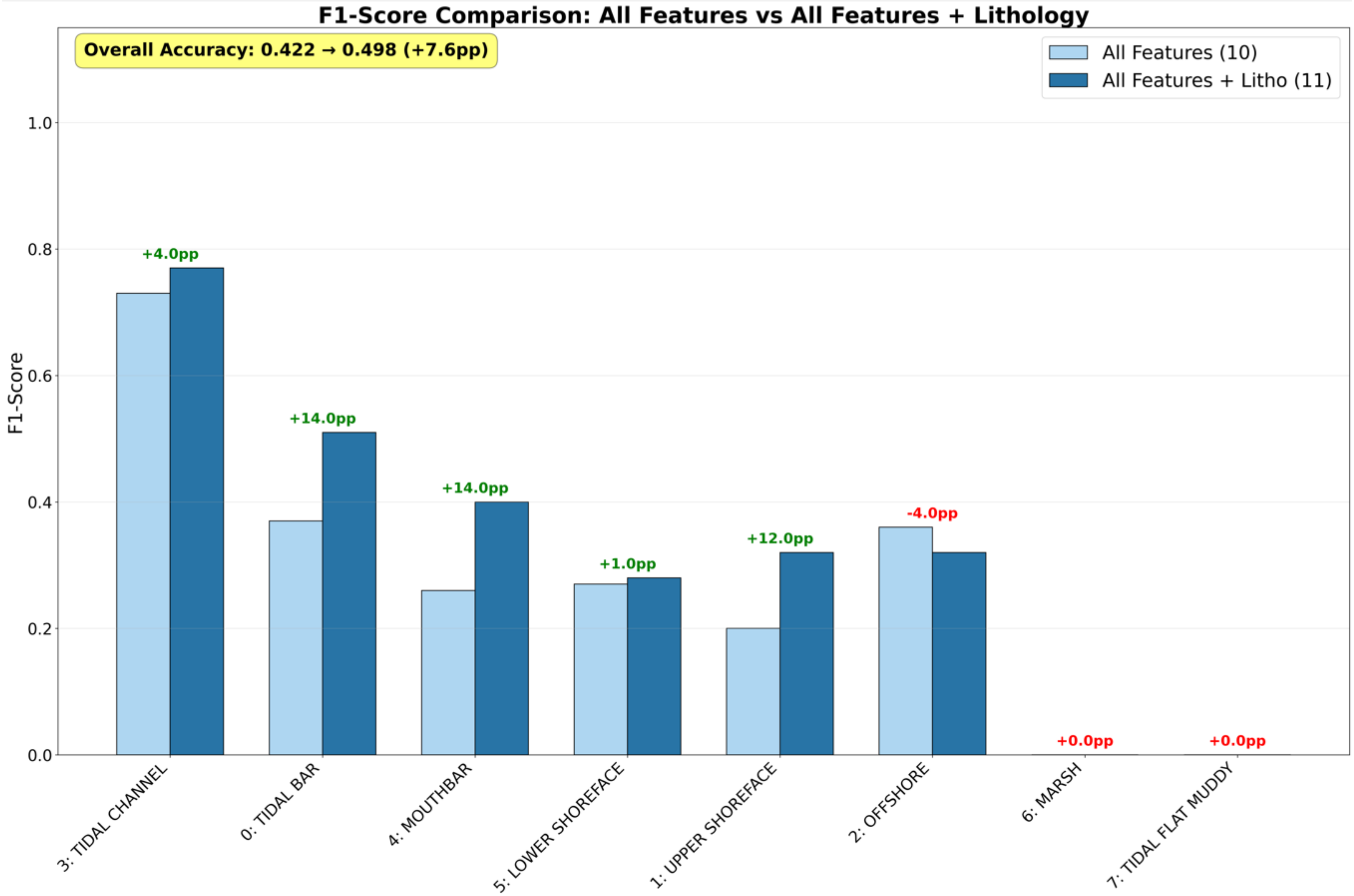
Mouth Bar vs Lower Shoreface

Tidal Channel vs Mouth Bar

Upper shoreface vs Tidal Bar

These results **motivate adding lithology** and engineered context features to tackle sand vs sand ambiguities and confusion

DOES ADDING A LITHOLOGY FEATURE HELP?



Addition of lithology significantly improved overall and per-facie prediction accuracy

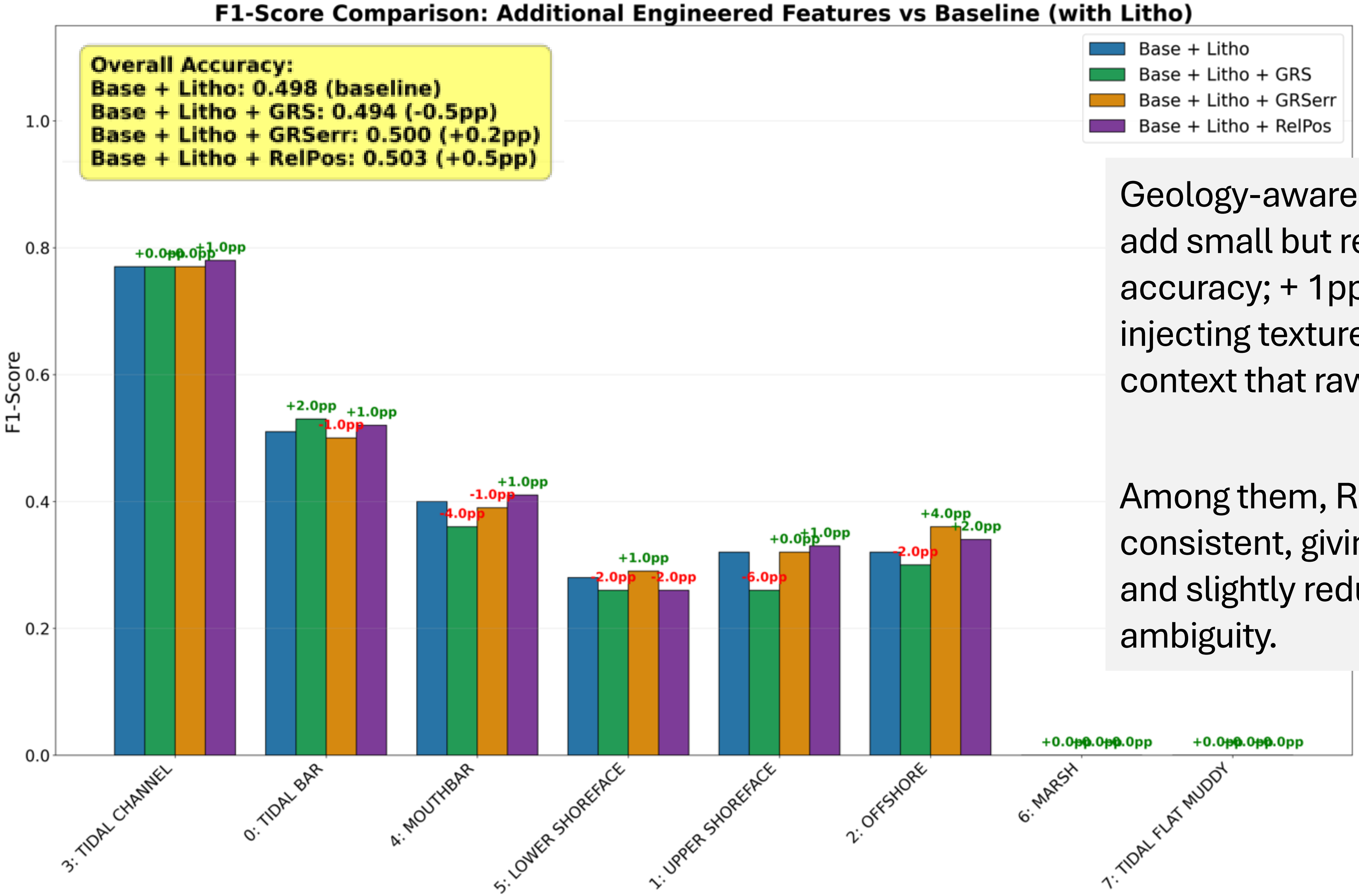
Sand vs sand confusion across problematic pairs also significantly improved

ENGINEERED FEATURES EXPERIMENTS OUTLINE

Engineered features aimed to add geological context which may improve model performance by decreasing confusion among sand vs sand depofacies

Experiments	Why
GR slope: d(GR)/dz in 3 m window	To capture coarsening up (Mouth Bar) vs fining up (Tidal Bar, Tidal Channel)
GR serration: rolling σ(GR) in 3 m window	GR Serration is higher in heteroliths/mud-draped tidal bars
RelPos (relative position in GR cycle)	Help distinguish sands in regressive-transgressive cycles

DO ENGINEERED FEATURES HELP?



Geology-aware engineered features add small but real gains (+small accuracy; + 1pp macro f1) by injecting texture and stacking context that raw logs miss.

Among them, RelPos is the most consistent, giving the best overall lift and slightly reducing sand vs sand ambiguity.

CONCLUSION

CONCLUSION

What worked

- ✓ Clustering separates clean sand vs shaly/tight: good for scoping/QC.
- ✓ Classifier better predicts depo. facies with all features including fluids.
- ✓ Lithology boosted accuracy and decreased sand vs sand confusion.
- ✓ Engineered features mimic geological features and slightly increase accuracy, can address class imbalance.

Limitations and challenges

- Persistent Tidal Bar vs Mouth Bar vs Lower Shoreface confusion.
- Minority facies under-learned.
- Engineered features impact is modest.

Value

- Faster, more consistent depo. facies picks.
- Lightweight workflow with clear diagnostics.

FUTURE WORK

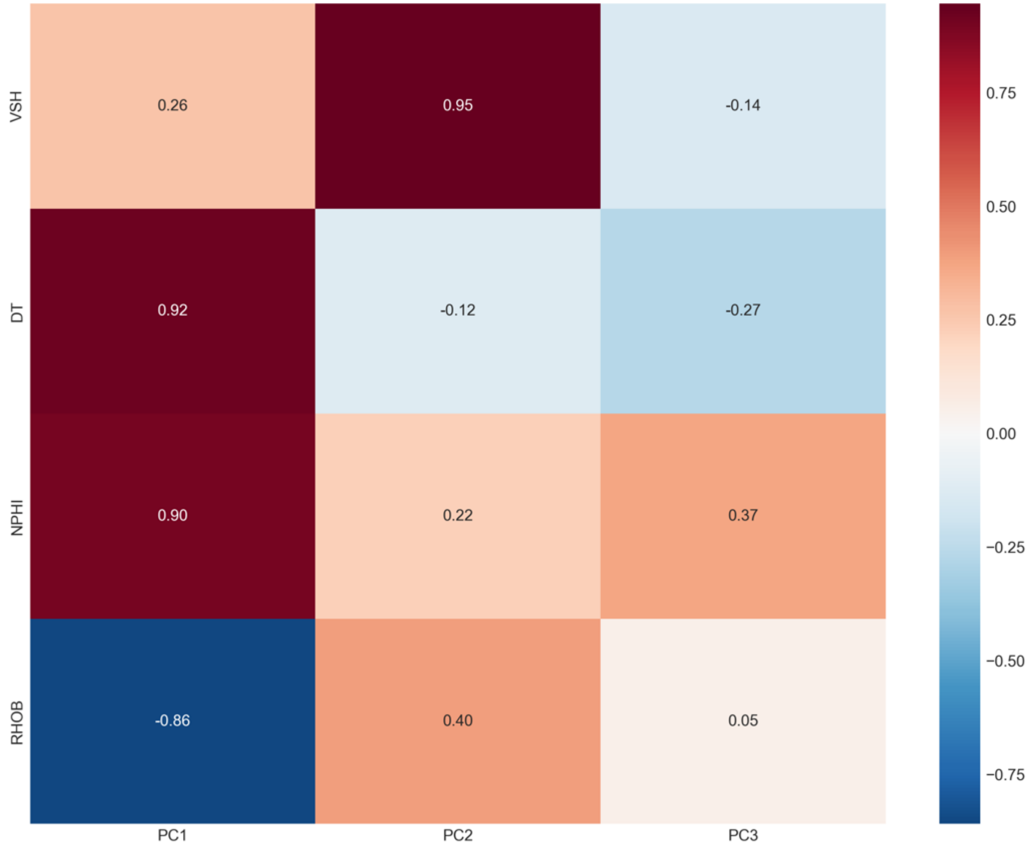
- Tackle imbalance and focus on confused pairs.
- Integrate core-derived features.
- Two-stage pipeline: impute missing logs (RF regressor) and use predictions as features.
- Tune engineered features.

BACK-UP SLIDES

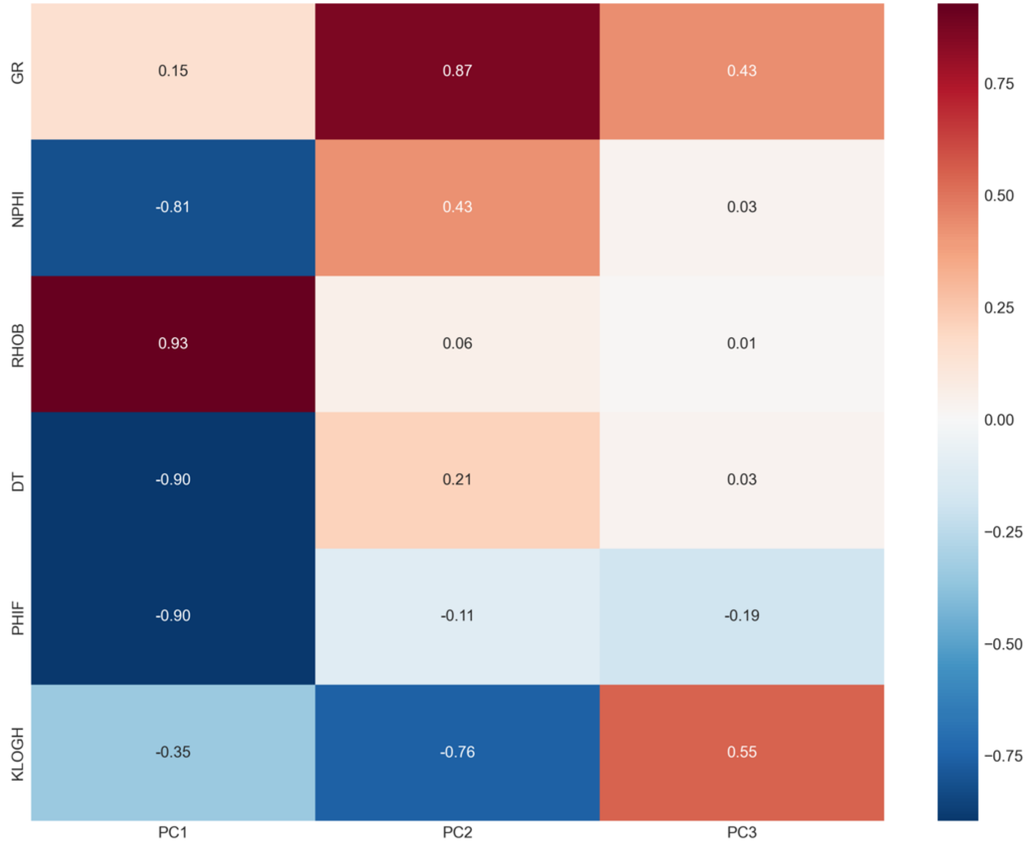
QUESTION 2 EXPERIMENTS RESULT

Two groups are tight and distinct in rock-property spaces.
PCA loadings plots show that Clustering at best separate porous & clean vs tight and clean vs shaly.
This is not enough for sand vs sand separation

VSH, DT, NPHI, RHOB



VSH, DT, NPHI, RHOB, KLOGH



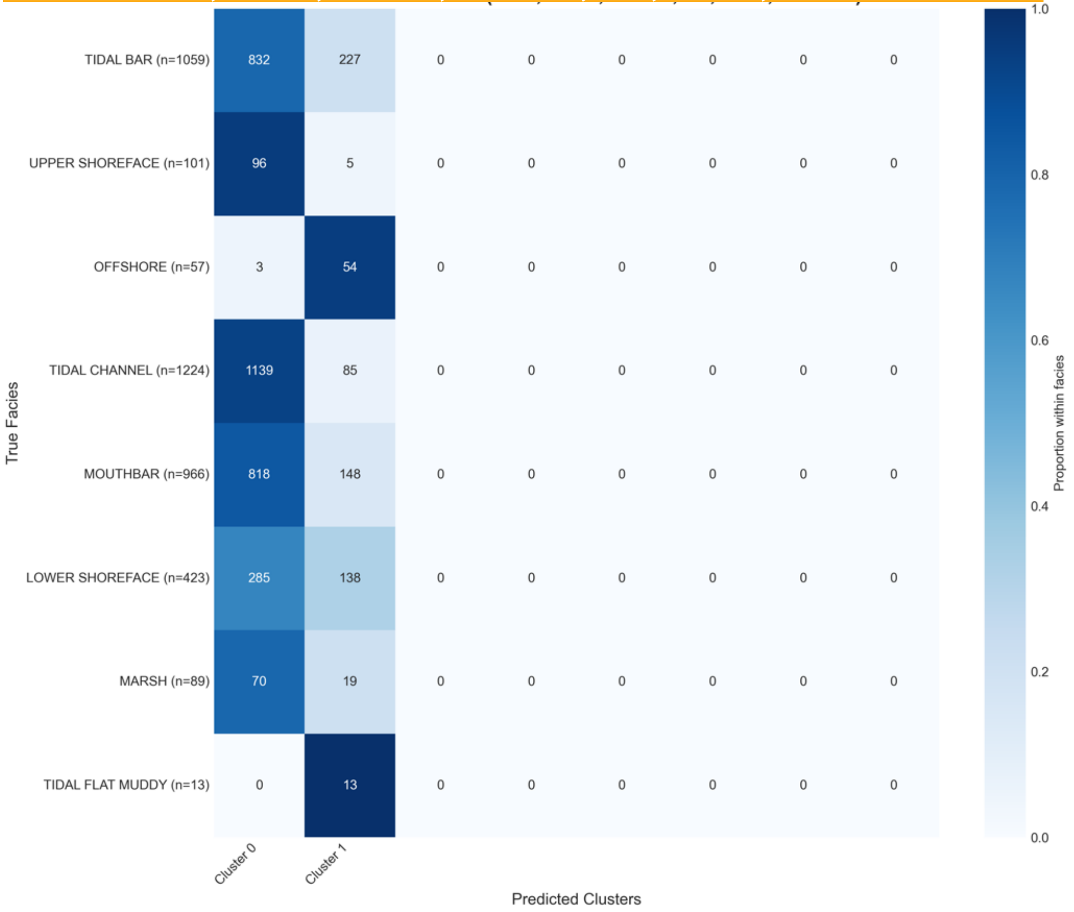
QUESTION 3 EXPERIMENTS RESULT

Clusters are geologically interpretable at a coarse level: clean sand vs shaly separation.

On the confusion matrices for k = 2 one cluster concentrates most sand depo. facies (Tidal Bar, Tidal Channel, Mouth Bar, Upper Shoreface), while other captures shaly depo. facies (Offshore, Tidal Flat Muddy).

k = 15 splits coarse sand/shale into many small clusters but does not separate sand depo. facies.

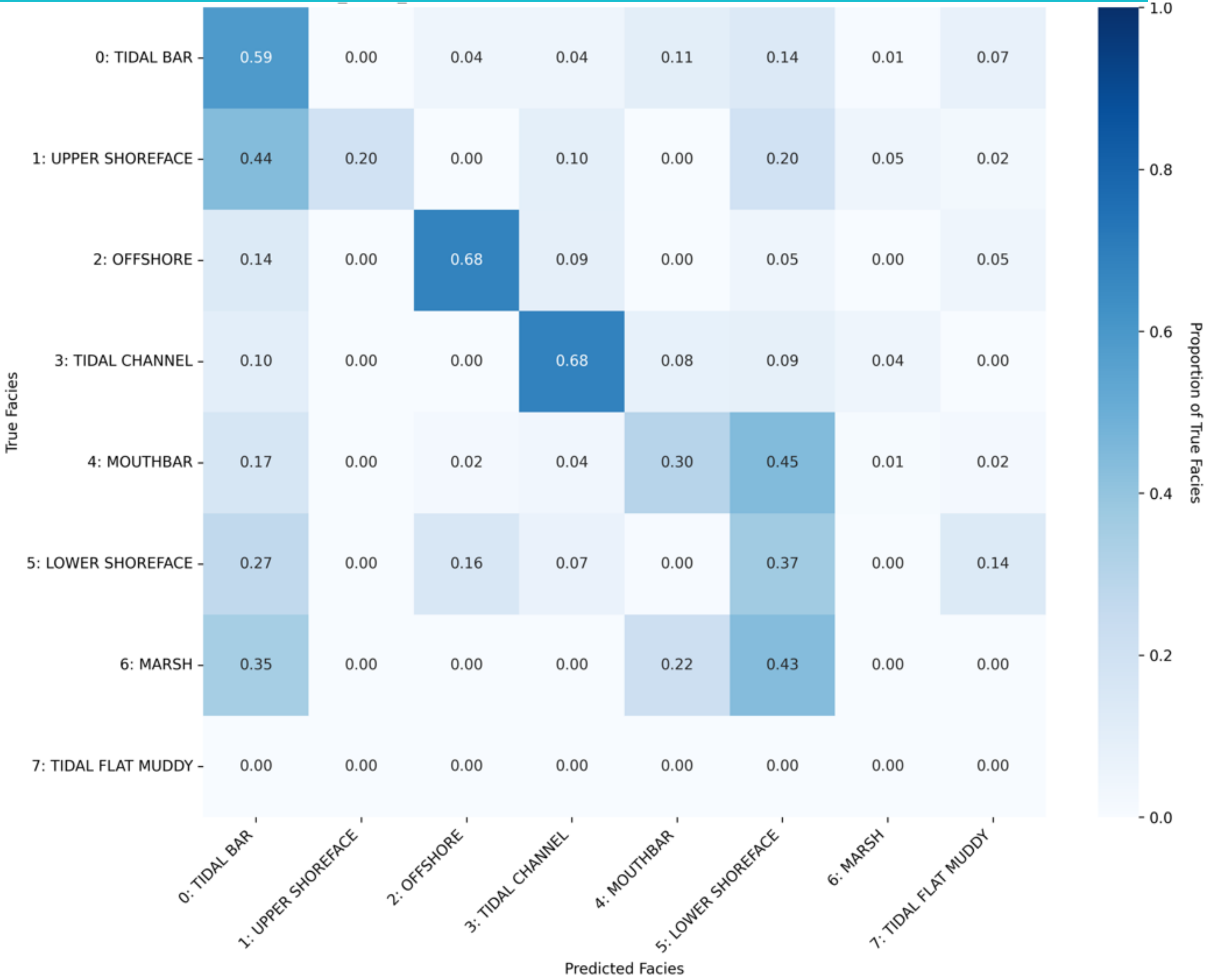
k = 2, VSH, NPHI, RHOB, DT, PHIF, KLOGH



k = 15, VSH, NPHI, RHOB, DT, PHIF, KLOGH

QUESTION 2 EXPERIMENTS RESULT

All logs set Confusion matrix + Lithology

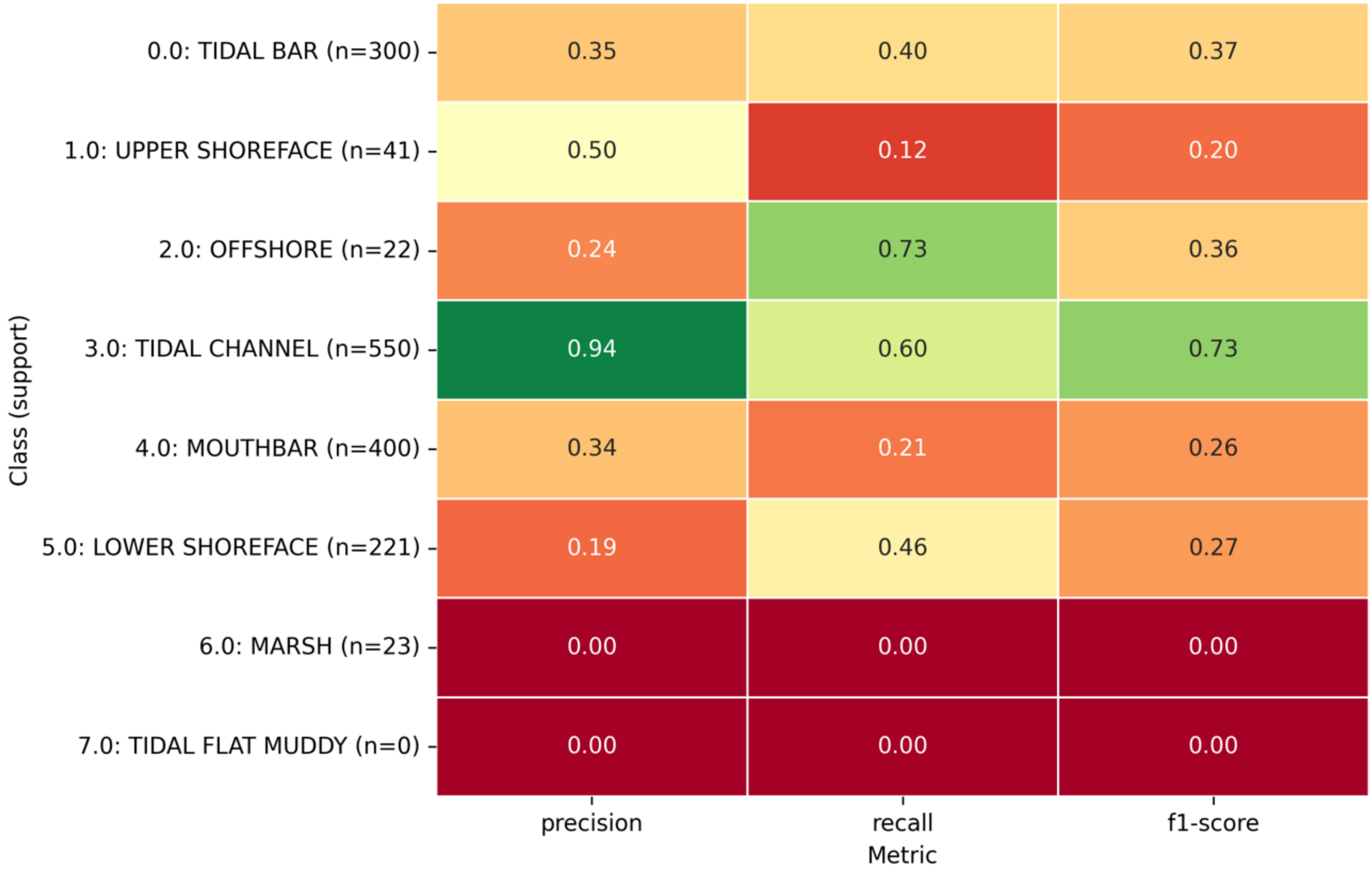


PHIK CORRELATION MATRIX

	MD	GR	DT	RT	RD	RS	NPHI	RHOB	VSH	PHIF	SW	KLOGH	Kg_hor	Kl_hor	So_core	Sw_core	Por	Depo. Facies	Lithology	Sand flag	Carb flag	Coal Flag
MD	1.00	0.37	0.38	0.13	0.12	0.31	0.24	0.38	0.40	0.42	0.59	0.26	0.32	0.31	0.42	0.45	0.36	0.38	0.38	0.18	0.09	0.09
GR	0.37	1.00	0.24	0.06	0.03	0.33	0.25	0.38	0.90	0.45	0.41	0.46	0.46	0.47	0.69	0.68	0.22	0.44	0.47	0.39	0.08	0.08
DT	0.38	0.24	1.00	0.08	0.07	0.07	0.45	0.64	0.30	0.74	0.54	0.38						0.34	0.34	0.46	0.43	0.23
RT	0.13	0.06	0.08	1.00	0.33		0.00	0.03	0.07	0.08	0.15	0.06							0.10	0.00	0.05	0.00
RD	0.12	0.03	0.07	0.33	1.00	0.16	0.00	0.03	0.07	0.05	0.11	0.14						0.20	0.00	0.00	0.00	0.00
RS	0.31	0.33	0.07		0.16	1.00	0.23	0.17	0.33	0.24	0.34	0.24						0.16	0.04	0.08	0.11	0.00
NPHI	0.24	0.25	0.45	0.00	0.00	0.23	1.00	0.26	0.29	0.56	0.21	0.12						0.12	0.12	0.16	0.25	0.02
RHOB	0.38	0.38	0.64	0.03	0.03	0.17	0.26	1.00	0.43	0.88	0.66	0.36						0.39	0.68	0.59	0.47	0.90
VSH	0.40	0.90	0.30	0.07	0.07	0.33	0.29	0.43	1.00	0.44	0.41	0.63	0.42	0.43	0.82	0.82	0.30	0.63	0.37	0.44	0.09	0.07
PHIF	0.42	0.45	0.74	0.08	0.05	0.24	0.56	0.88	0.44	1.00	0.54	0.32	0.27	0.27	0.78	0.80	0.64	0.34	0.63	0.62	0.41	0.24
SW	0.59	0.41	0.54	0.15	0.11	0.34	0.21	0.66	0.41	0.54	1.00	0.45	0.43	0.45	0.90	0.96	0.09	0.36	0.54	0.33	0.19	0.07
KLOGH	0.26	0.46	0.38	0.06	0.14	0.24	0.12	0.36	0.63	0.32	0.45	1.00	0.81	0.80	0.64	0.57	0.30	0.35	0.16	0.13	0.07	0.00
Kg_hor	0.32	0.46							0.42	0.27	0.43	0.81	1.00	1.00			0.36	0.33		0.13	0.09	
Kl_hor	0.31	0.47							0.43	0.27	0.45	0.80	1.00	1.00			0.37	0.34		0.13	0.08	
So_core	0.42	0.69							0.82	0.78	0.90	0.64			1.00	0.98		0.71		0.67	0.60	
Sw_core	0.45	0.68							0.82	0.80	0.96	0.57			0.98	1.00		0.75		0.79	0.71	
Por	0.36	0.22							0.30	0.64	0.09	0.30	0.36	0.37			1.00	0.36		0.50	0.30	
Depo. Facies	0.38	0.44	0.34		0.20	0.16	0.12	0.39	0.63	0.34	0.36	0.35	0.33	0.34	0.71	0.75	0.36	1.00	0.56	0.24	0.21	0.26
Lithology	0.38	0.47	0.34	0.10	0.00	0.04	0.12	0.68	0.37	0.63	0.54	0.16						0.56	1.00	0.45	0.36	0.47
Sand flag	0.18	0.39	0.46	0.00	0.00	0.08	0.16	0.59	0.44	0.62	0.33	0.13	0.13	0.13	0.67	0.79	0.50	0.24	0.45	1.00	0.98	0.10
Carb flag	0.09	0.08	0.43	0.05	0.00	0.11	0.25	0.47	0.09	0.41	0.19	0.07	0.09	0.08	0.60	0.71	0.30	0.21	0.36	0.98	1.00	0.00
Coal Flag	0.09	0.08	0.23	0.00	0.00	0.00	0.02	0.90	0.07	0.24	0.07	0.00						0.26	0.47	0.10	0.00	1.00

QUESTION 1

Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH

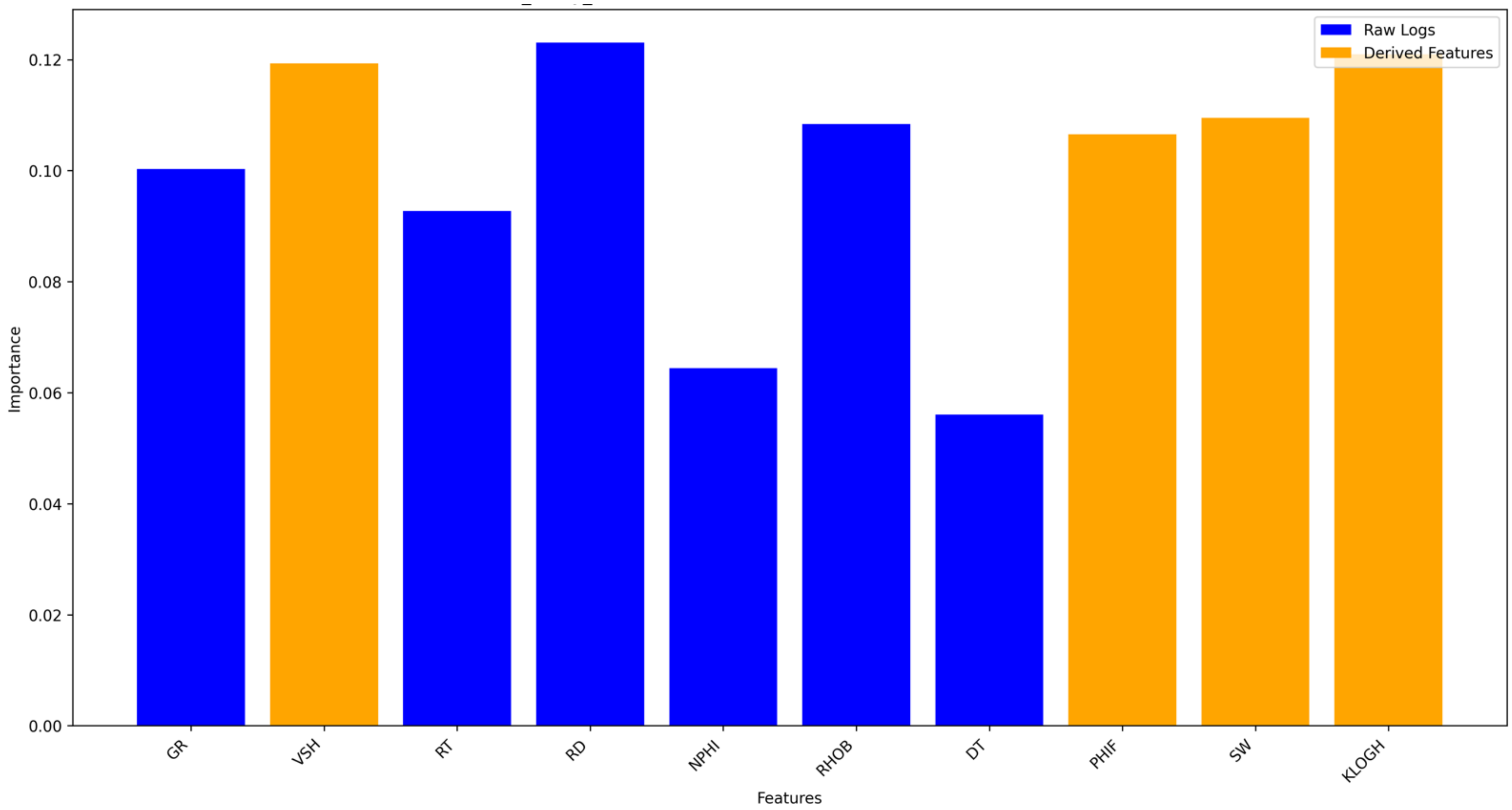


QUESTION 1

Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH

QUESTION 1

Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH

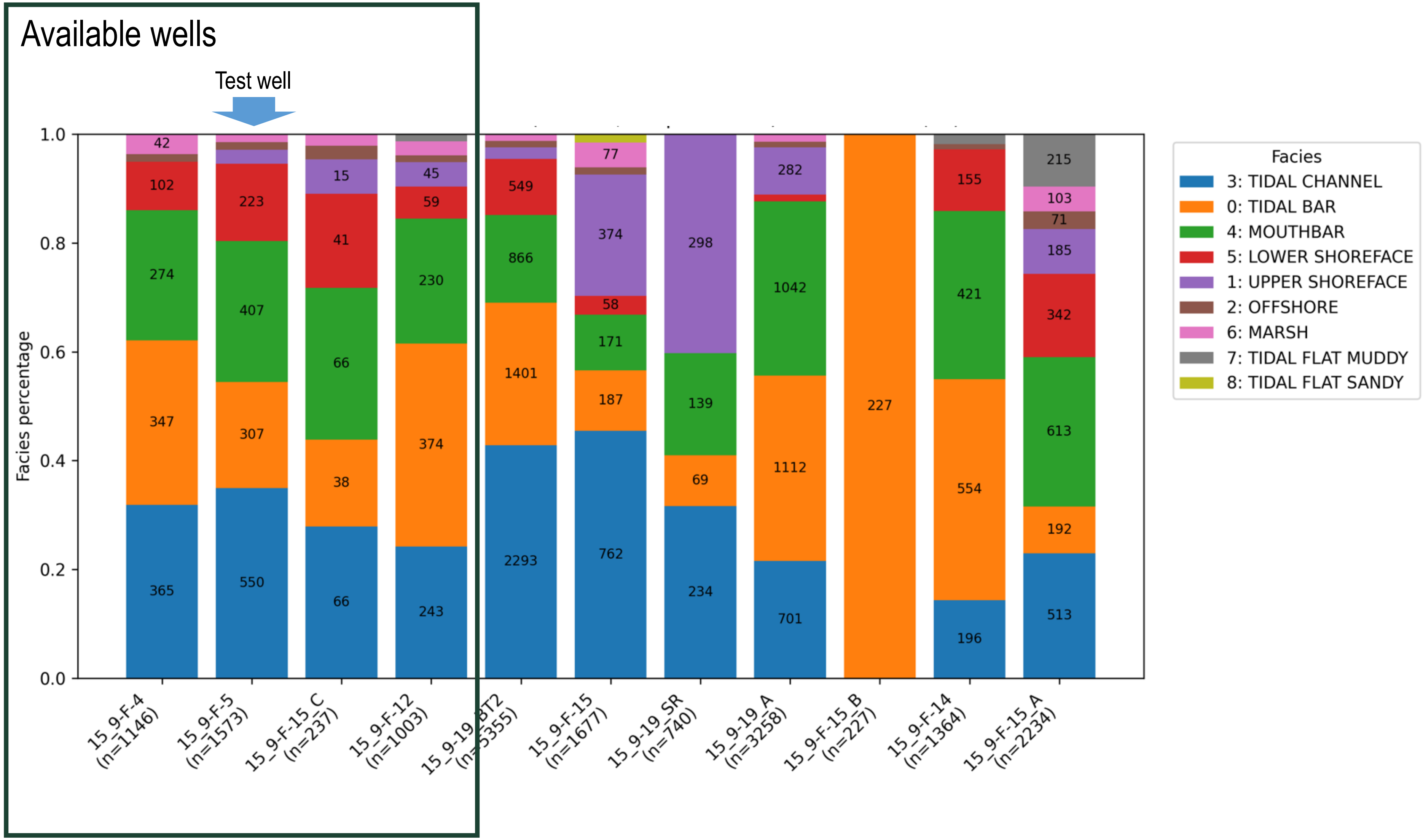


QUESTION 1

Feature set: VSH, NPHI, RHOB, DT, PHIF, KLOGH

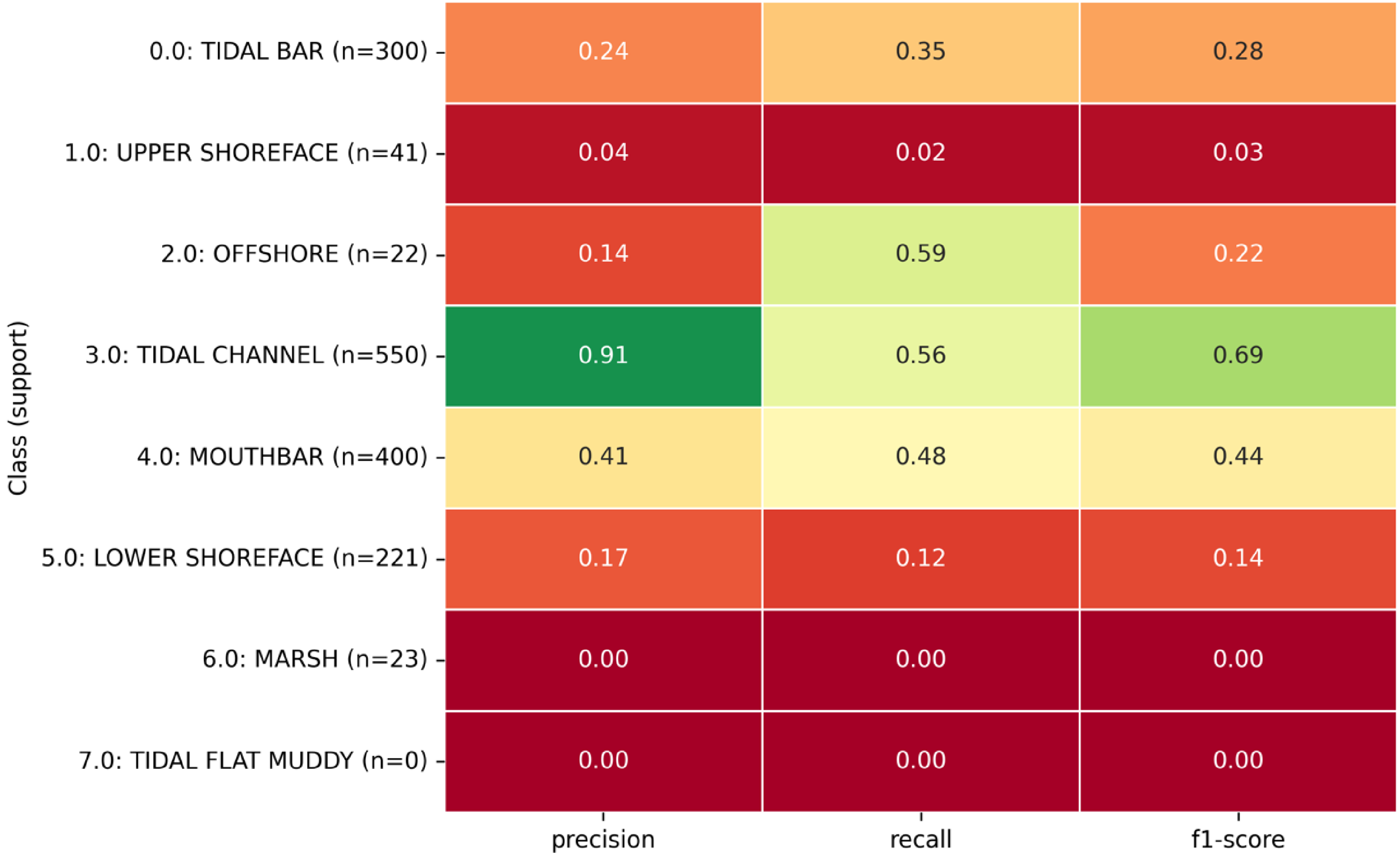
Four wells are available for this feature set

Well F-5 chosen as a test well because it is closest to the overall class mix



QUESTION 1

Feature set: VSH, NPHI, RHOB, DT, PHIF, KLOGH



QUESTION 1

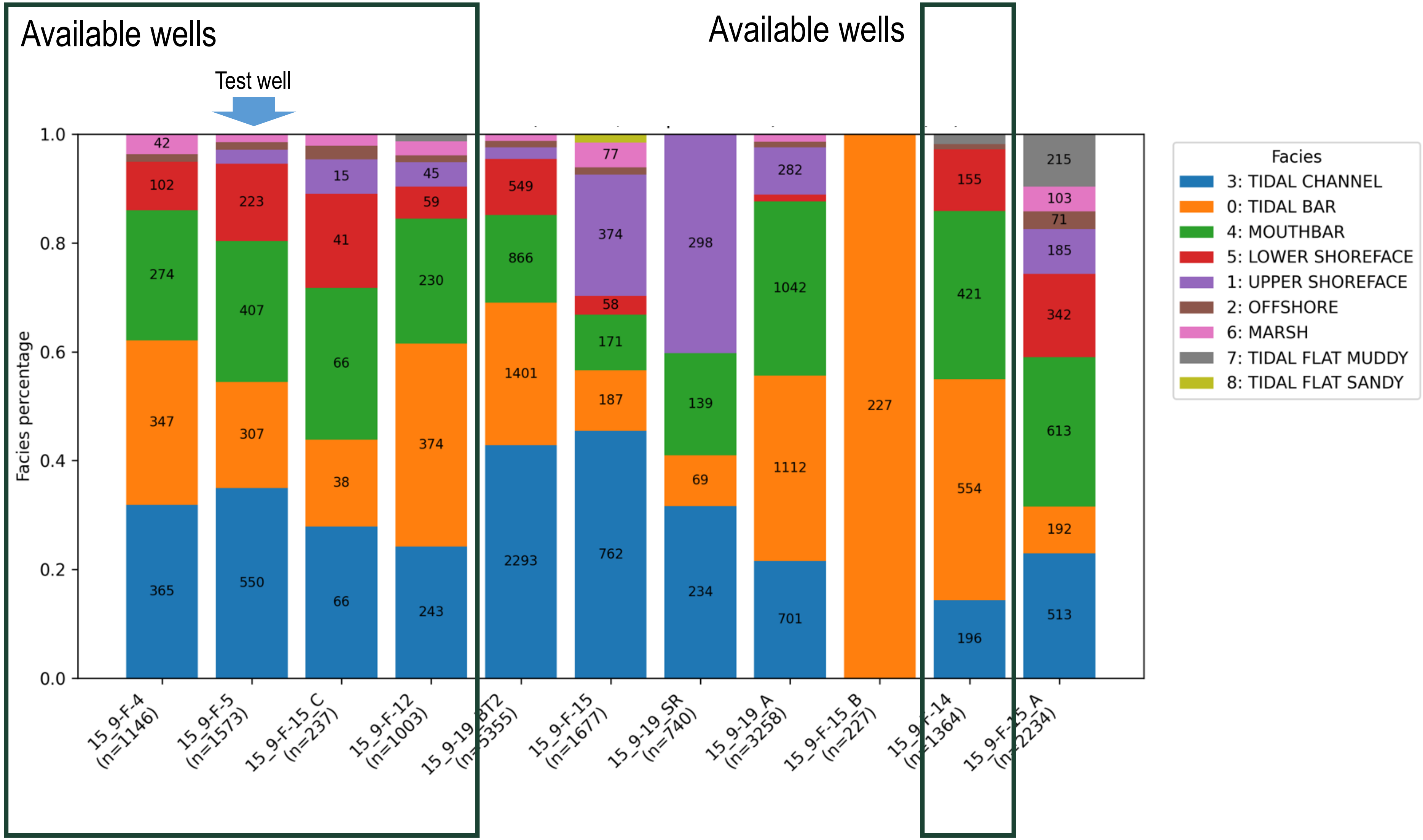
Feature set: VSH, NPHI, RHOB, DT, PHIF, KLOGH

QUESTION 1

Feature set: VSH, NPHI, RHOB, DT

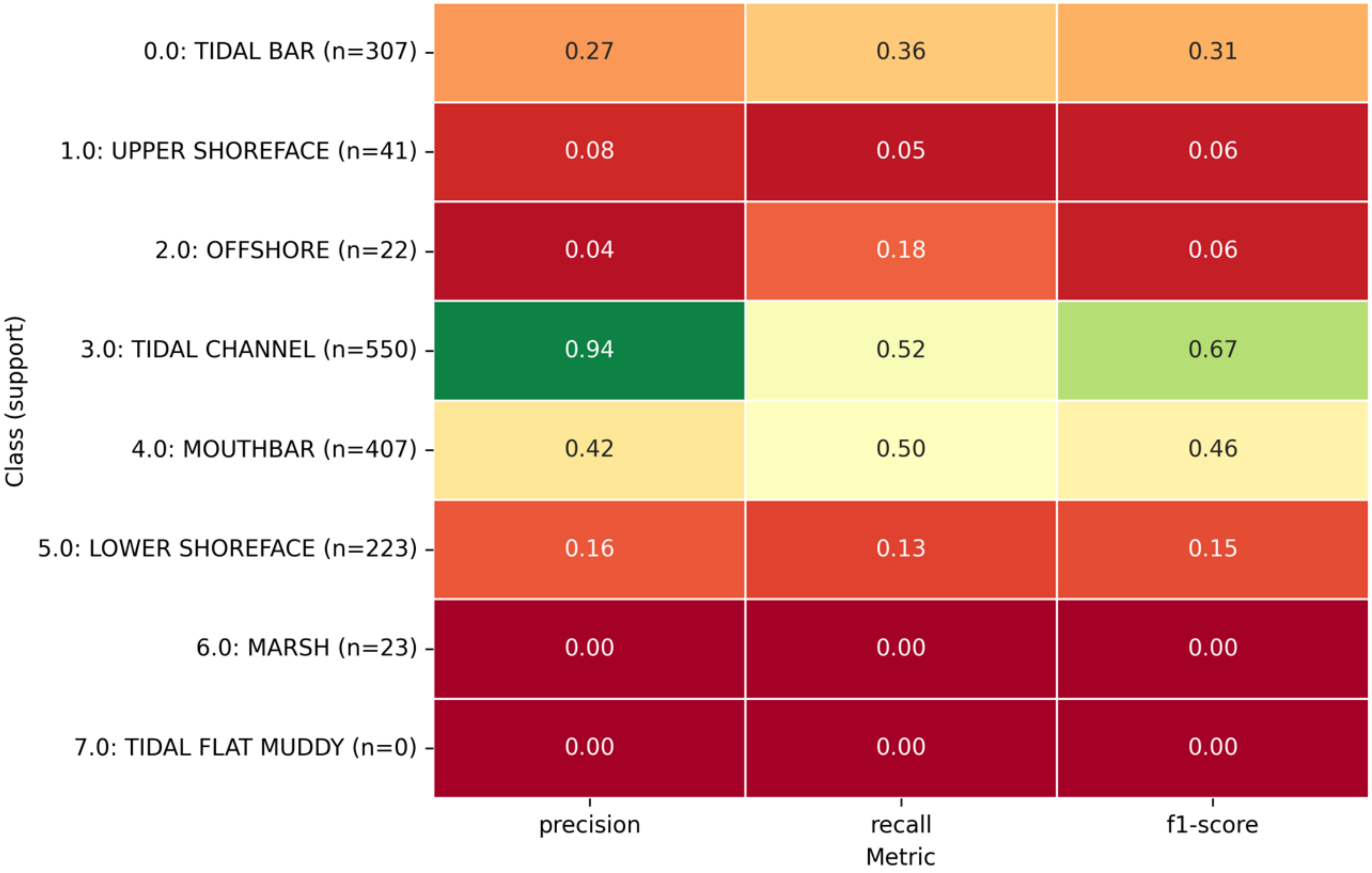
Five wells are available for this feature set

Well F-5 chosen as a test well because it is closest to the overall class mix



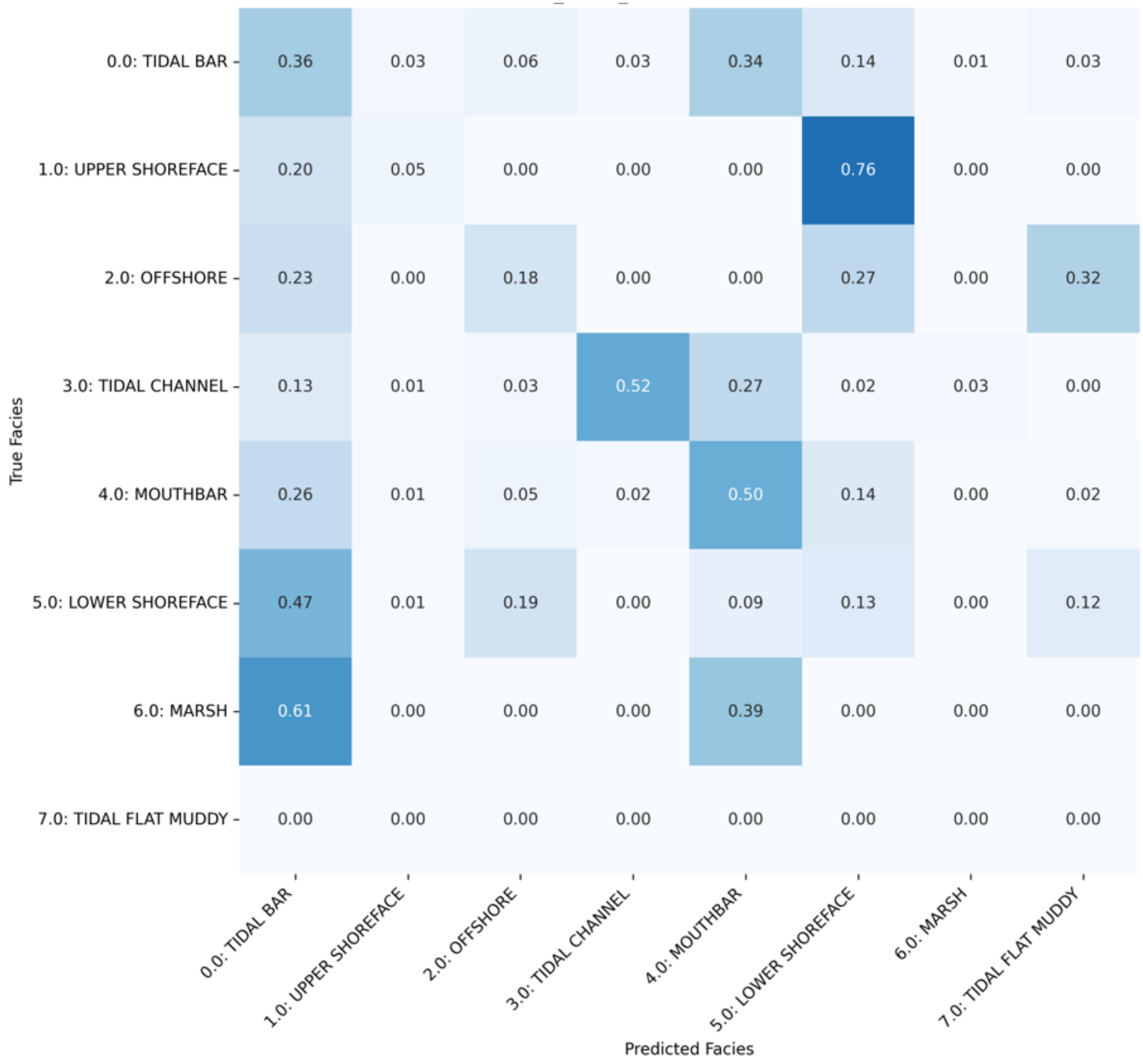
QUESTION 1

Feature set: VSH, NPHI, RHOB, DT



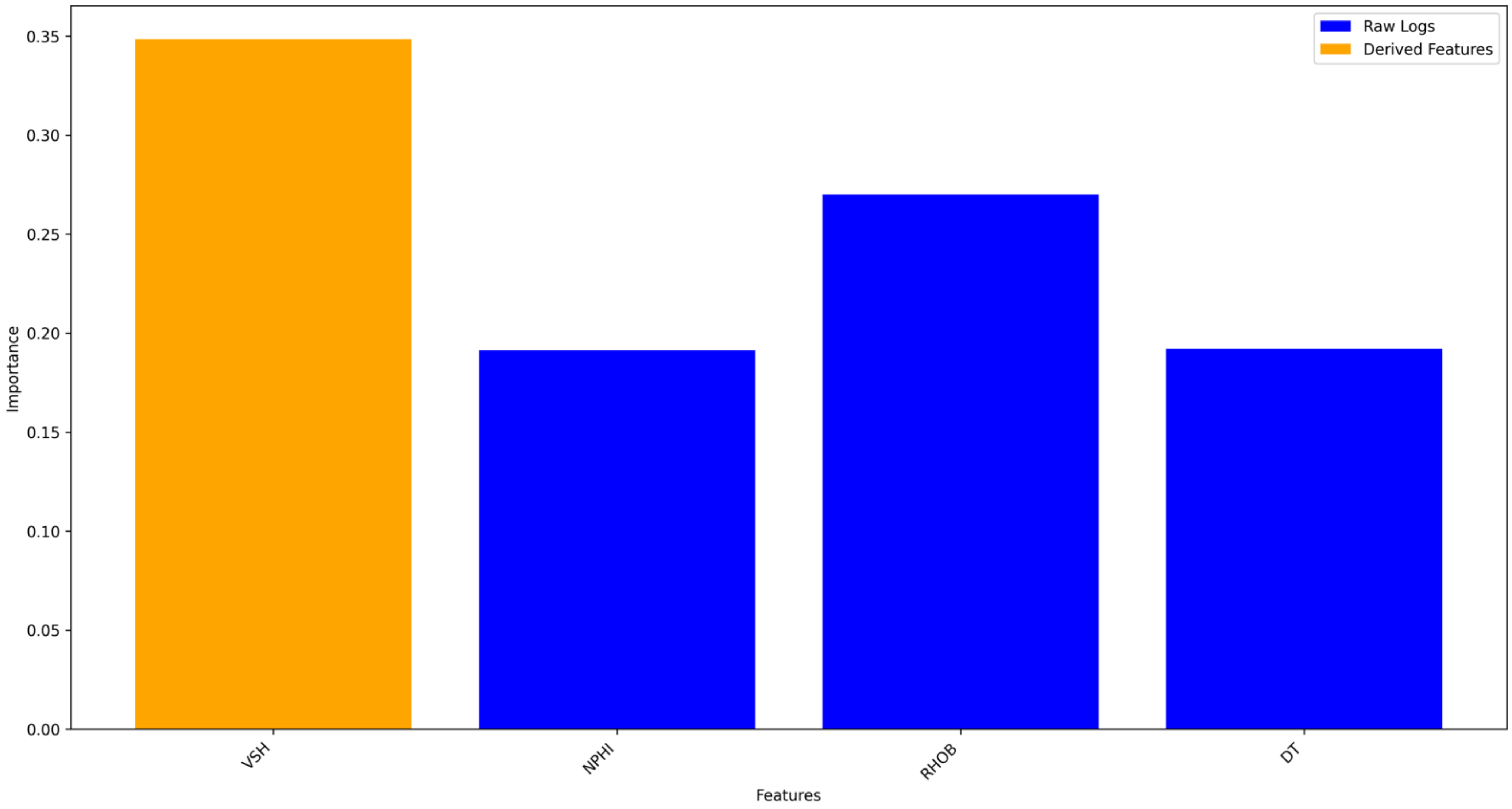
QUESTION 1

Feature set: VSH, NPHI, RHOB, DT



QUESTION 1

Feature set: VSH, NPHI, RHOB, DT

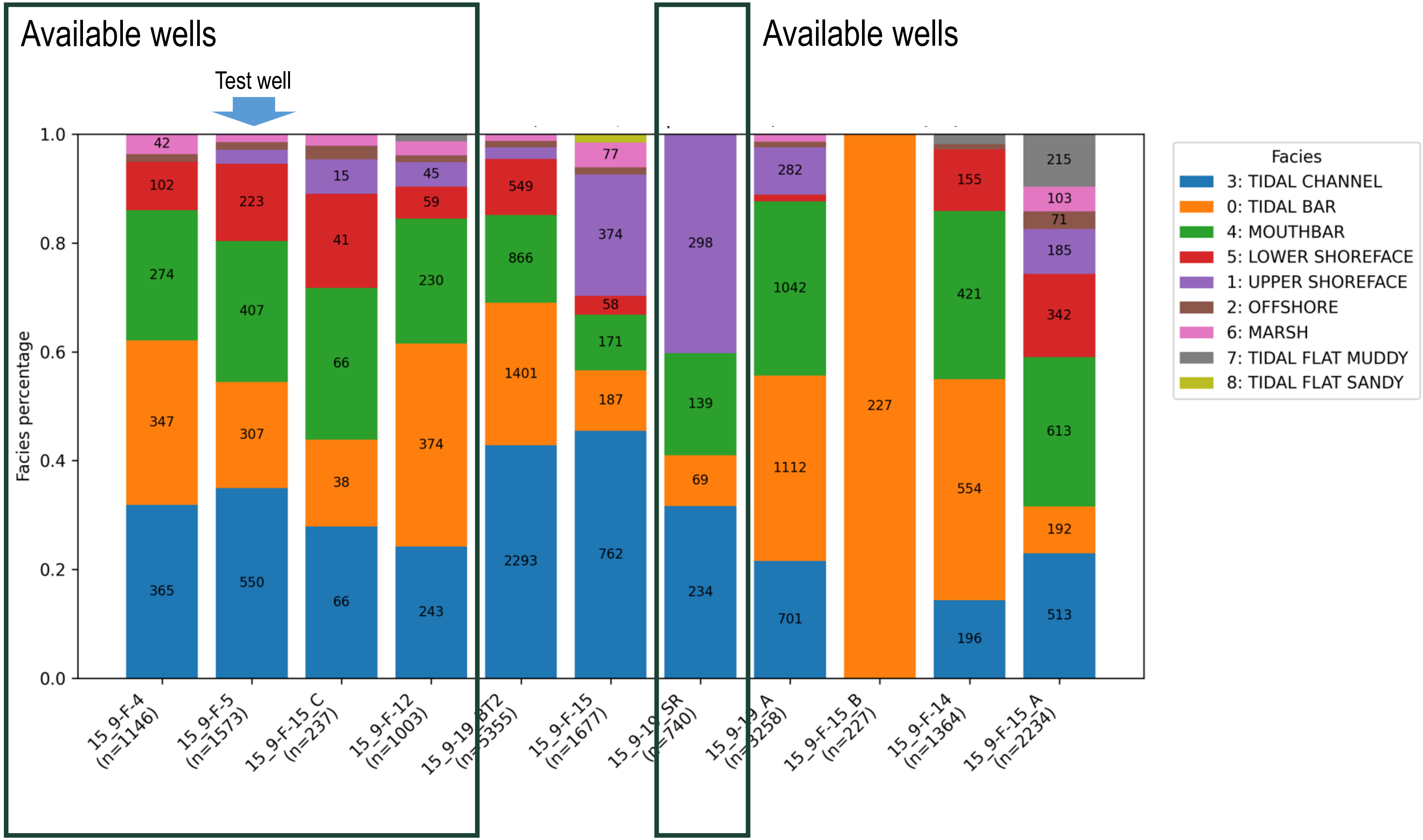


QUESTION 1

Feature set: VSH, PHIF, DT

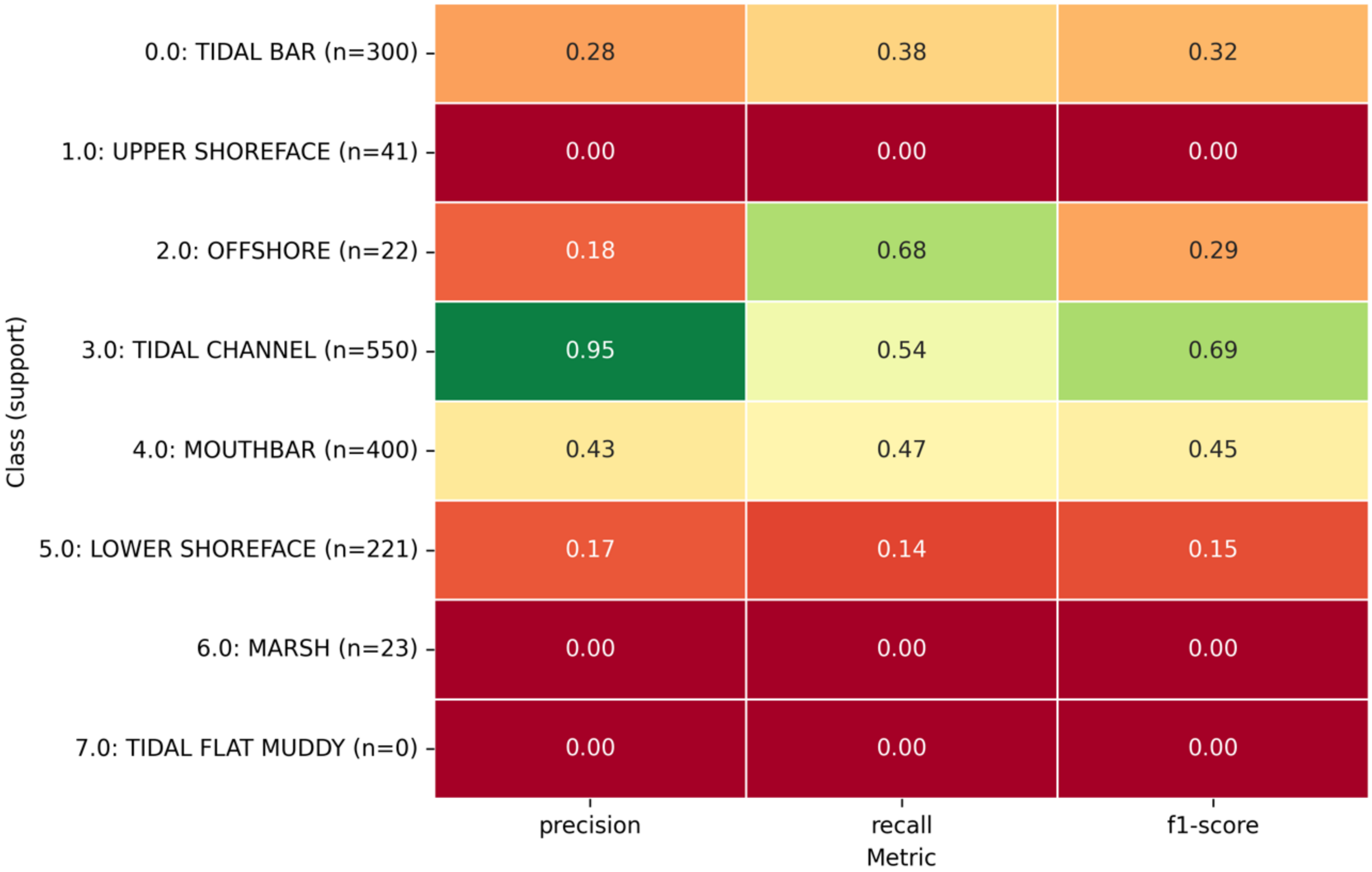
Five wells are available for this feature set

Well F-5 chosen as a test well because it is closest to the overall class mix



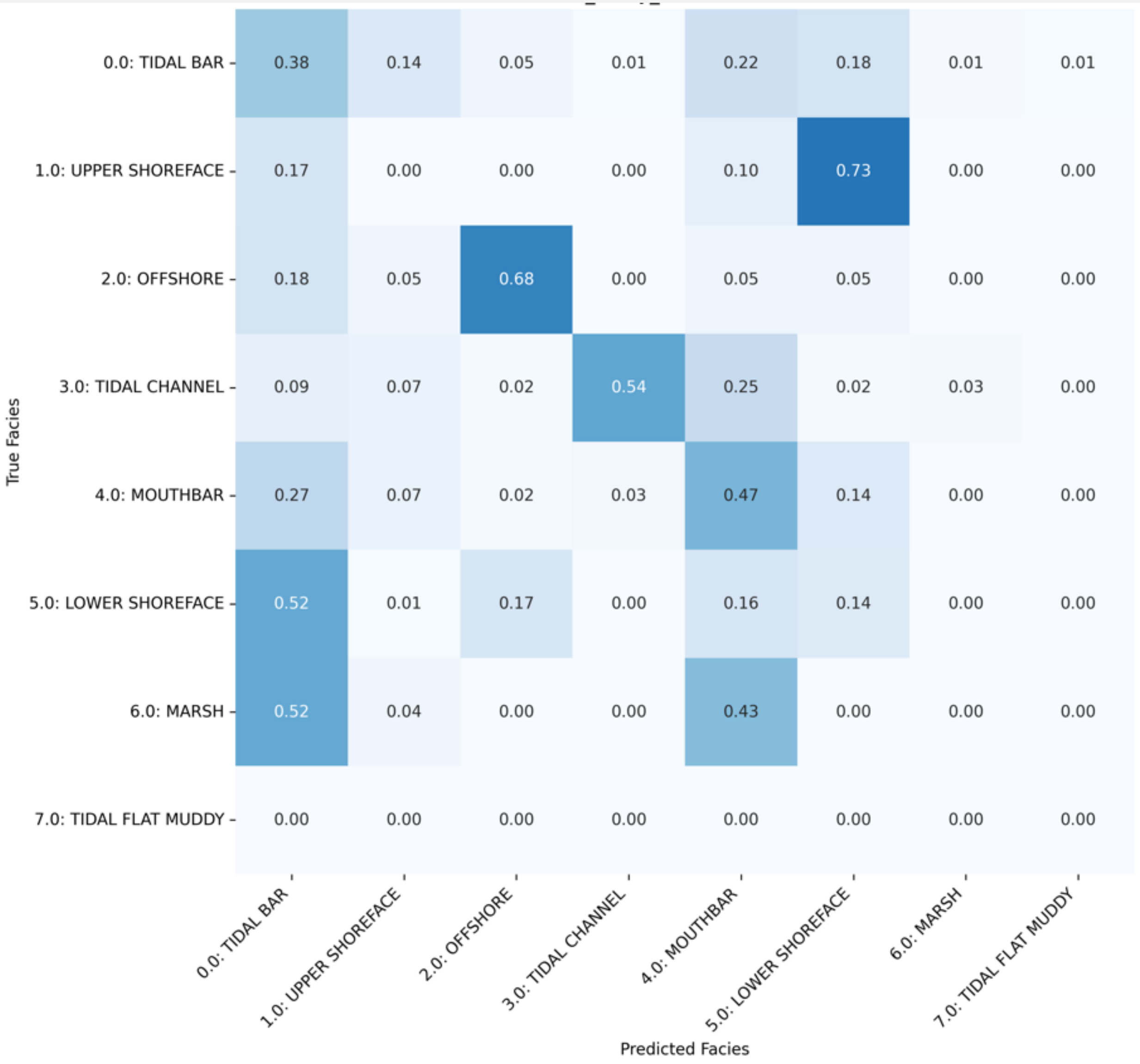
QUESTION 1

Feature set: VSH, PHIF, DT



QUESTION 1

Feature set: VSH, PHIF, DT

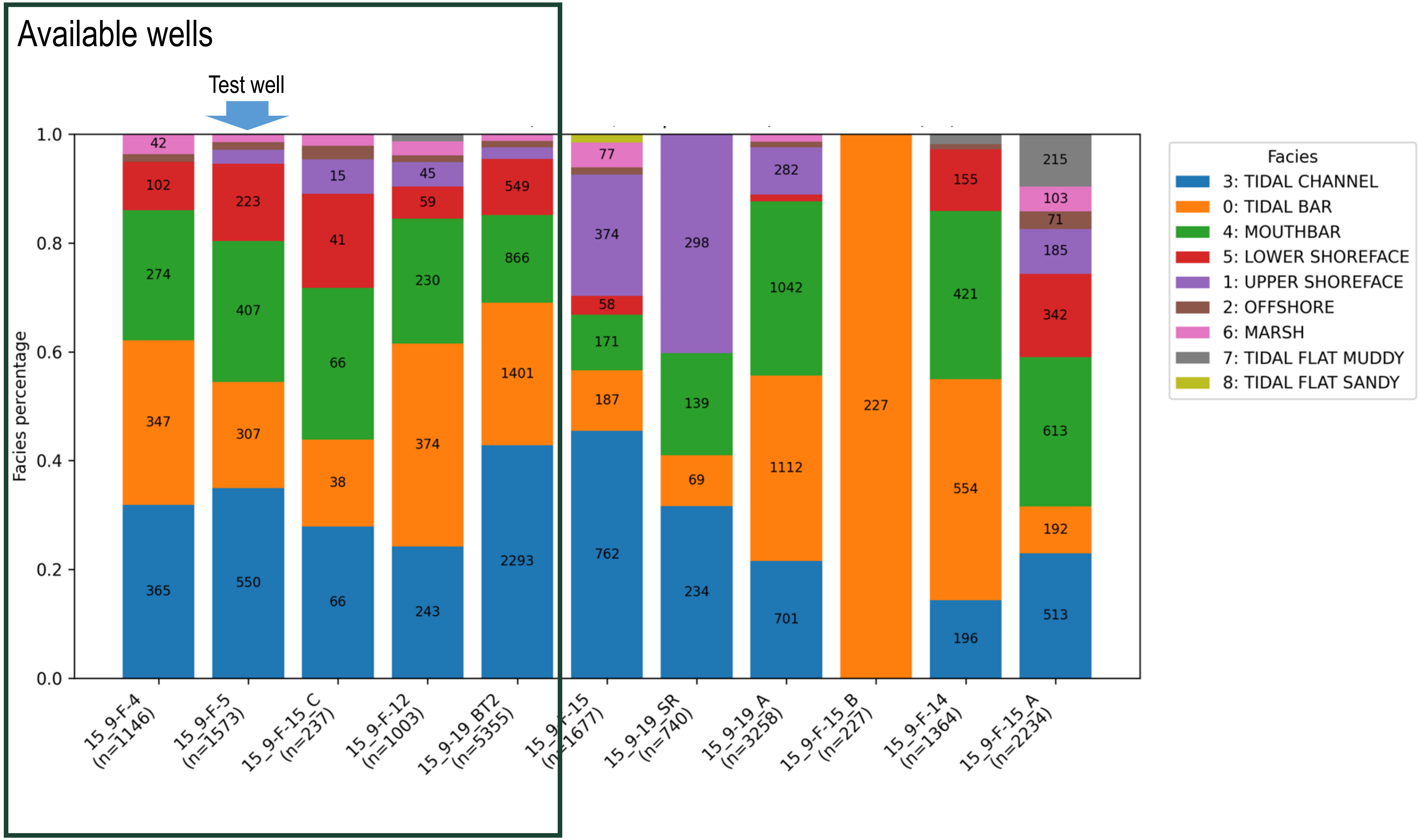


QUESTION 1

Feature set: VSH, PHIF, SW, KLOGH

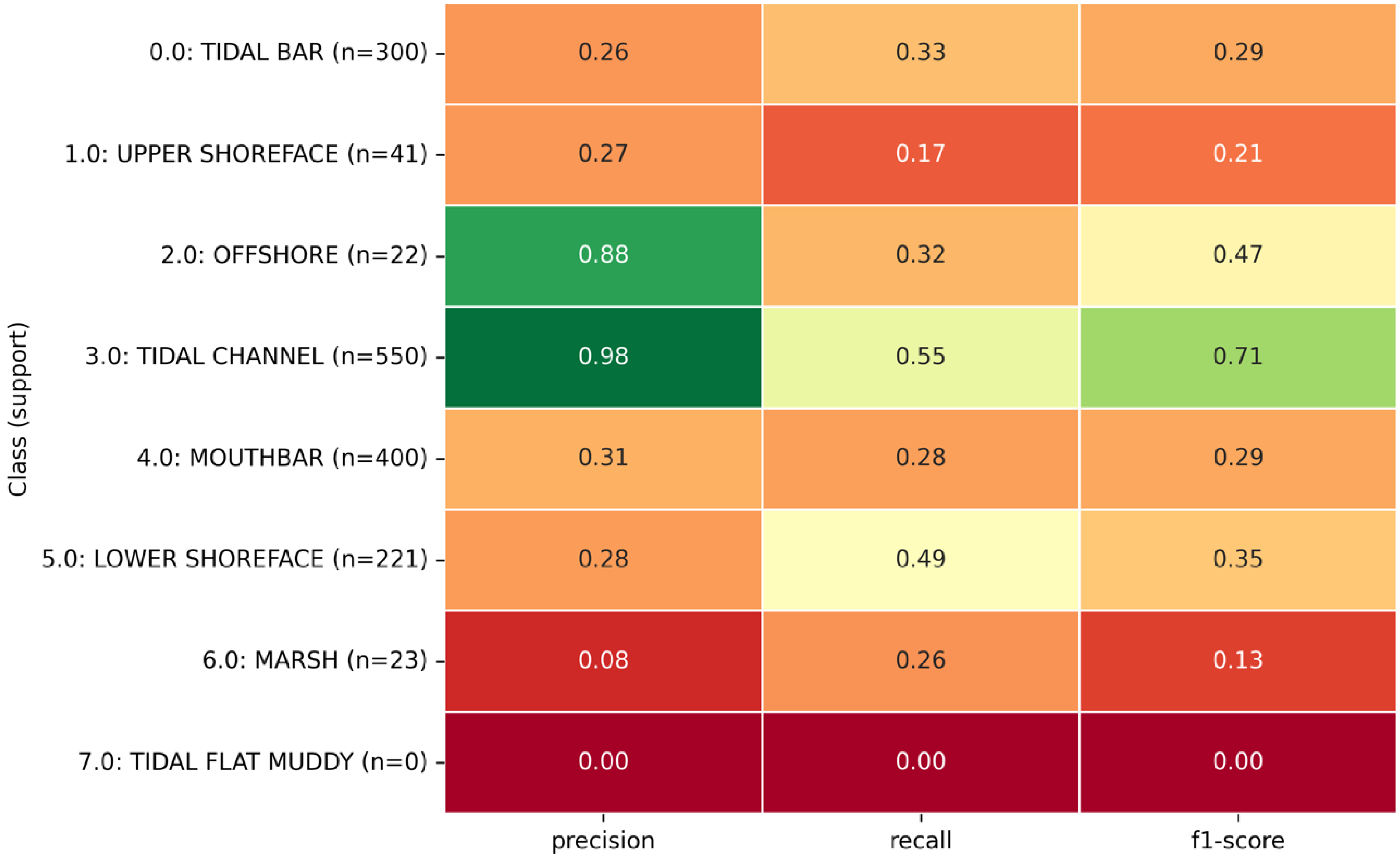
Five wells are available for this feature set

Well F-5 chosen as a test well because it is closest to the overall class mix (and a recorded problem with log distribution of F-12)



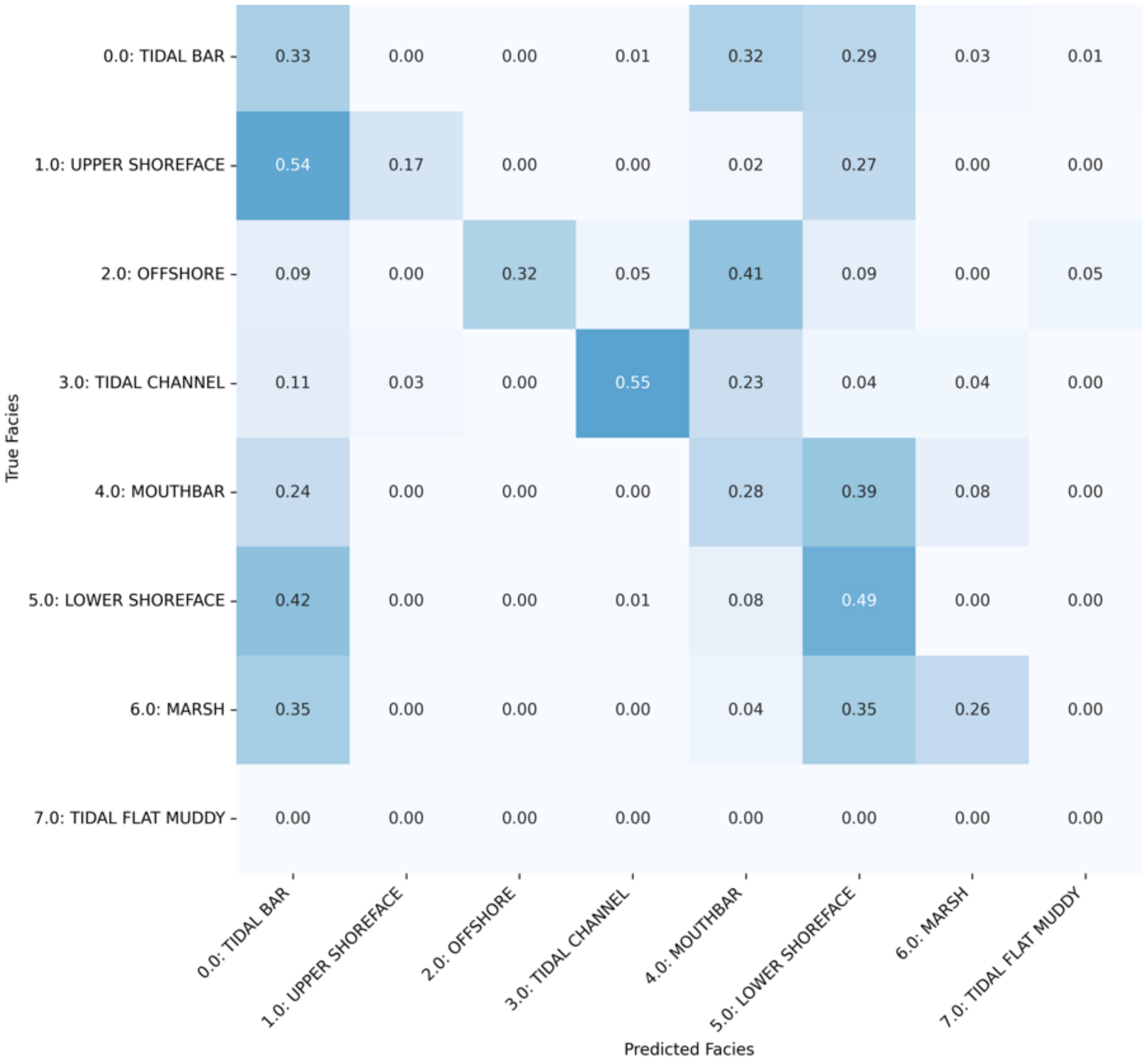
QUESTION 1

Feature set: VSH, PHIF, SW, KLOGH



QUESTION 1

Feature set: VSH, PHIF, SW, KLOGH

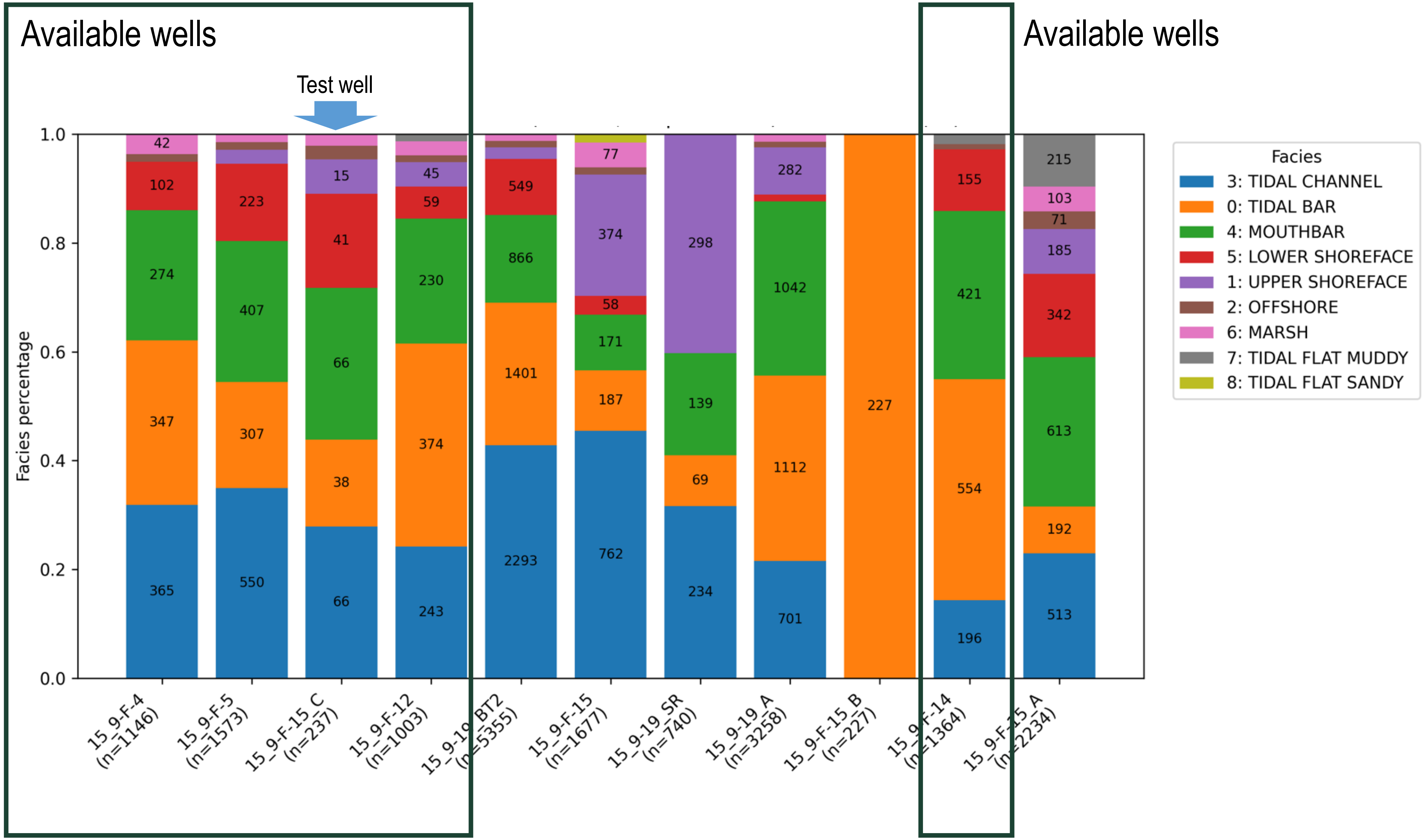


QUESTION 1

Feature set: GR, NPHI, RHOB, DT

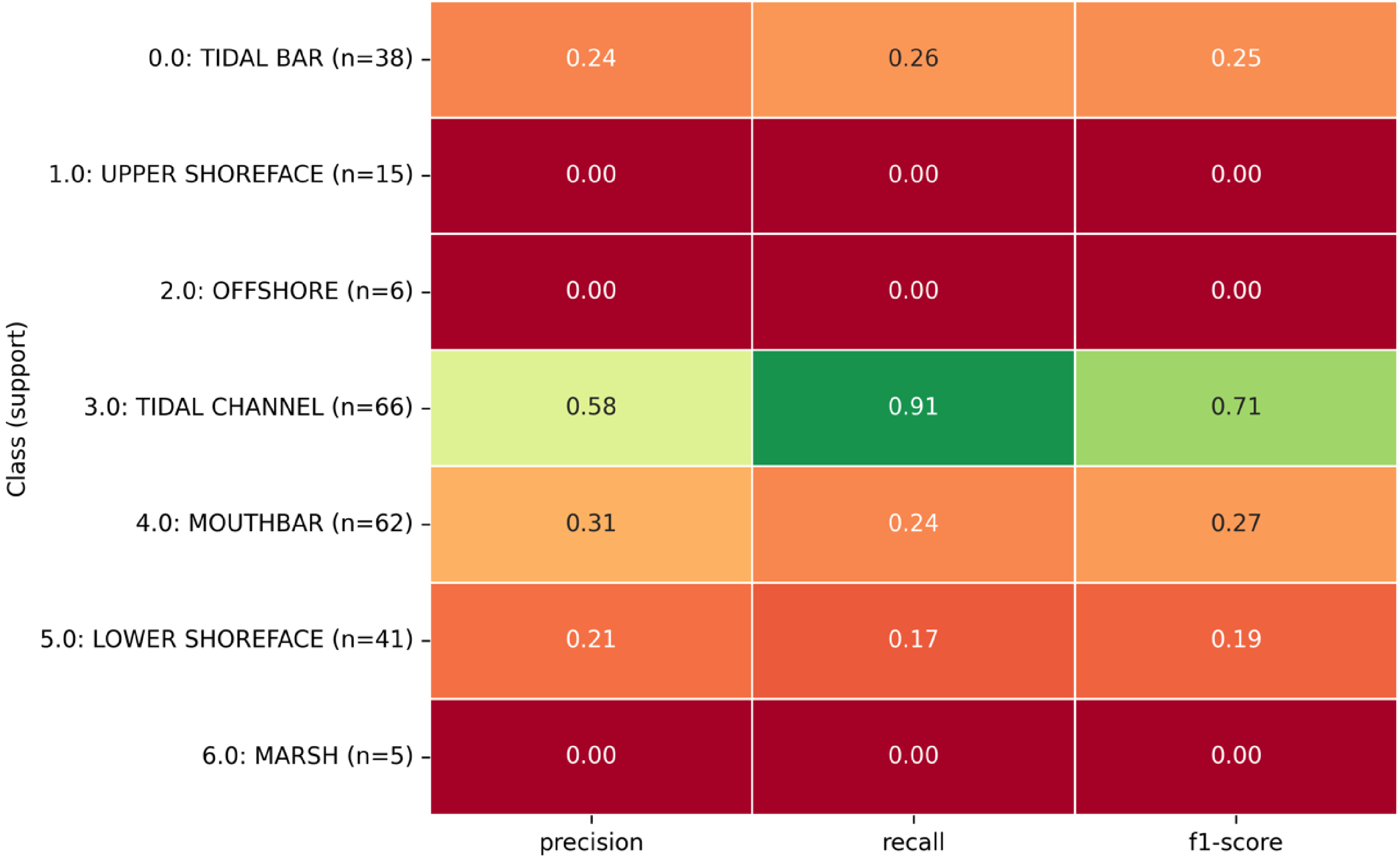
Five wells are available for this feature set

Well F-15C chosen as a test well because it is closest to the overall class mix (and a recorded problem with log distribution of F-12)



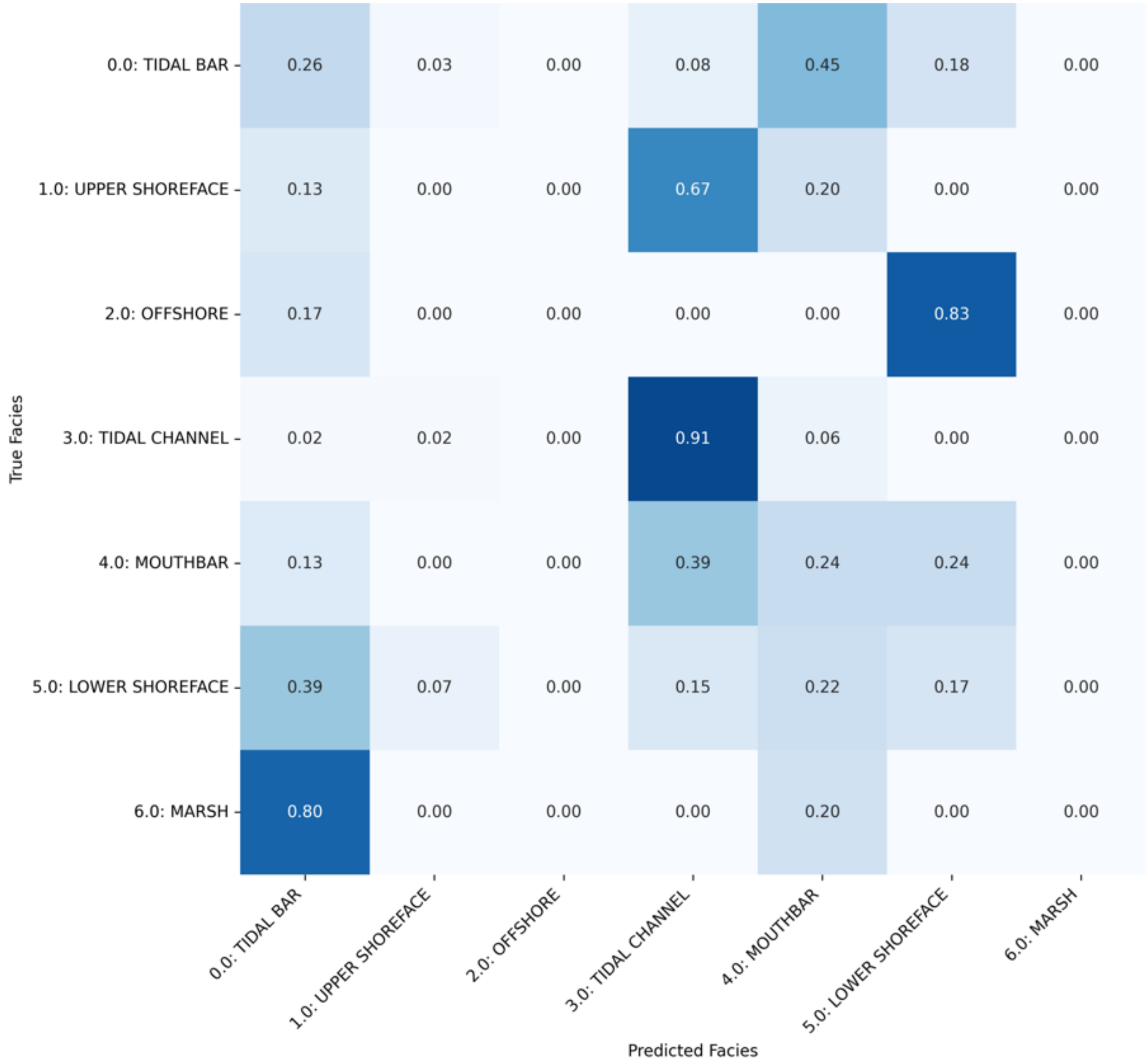
QUESTION 1

Feature set: GR, NPHI, RHOB, DT



QUESTION 1

Feature set: GR, NPHI, RHOB, DT

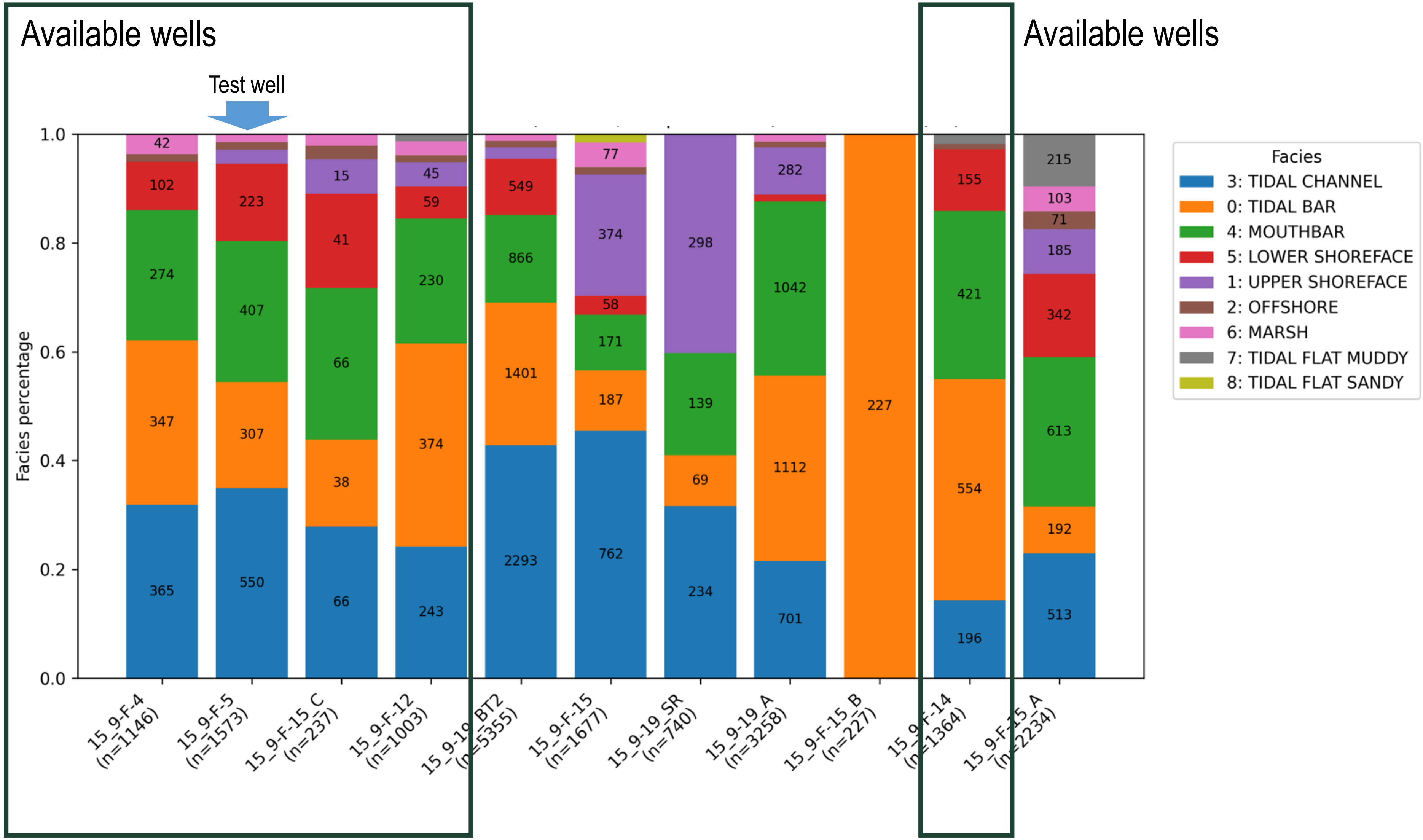


QUESTION 1

Feature set: GR, NPHI, RHOB, DT, RT, RD

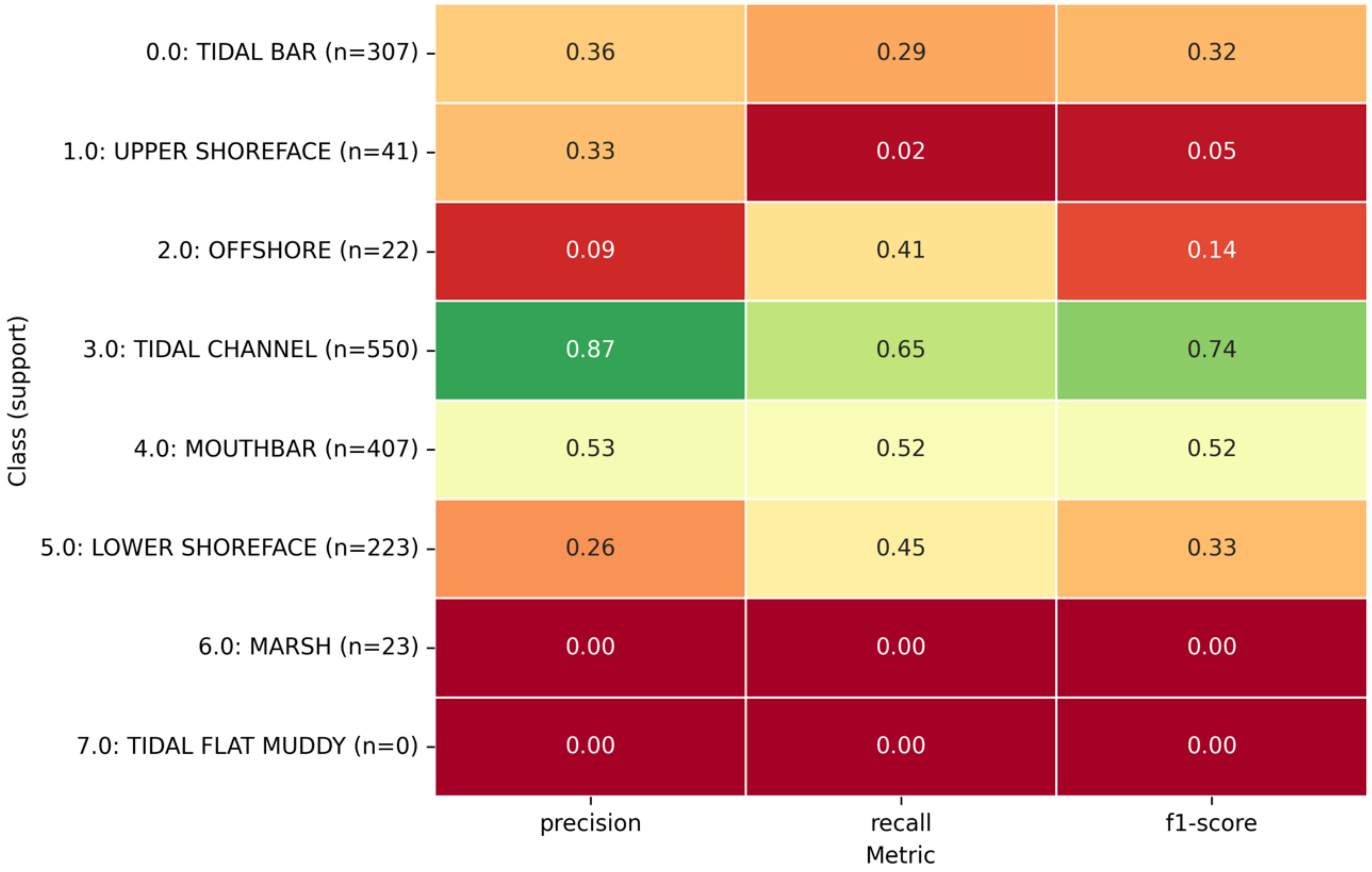
Five wells are available for this feature set

Well F-5 chosen as a test well because it is closest to the overall class mix (and a recorded problem with log distribution of F-12)



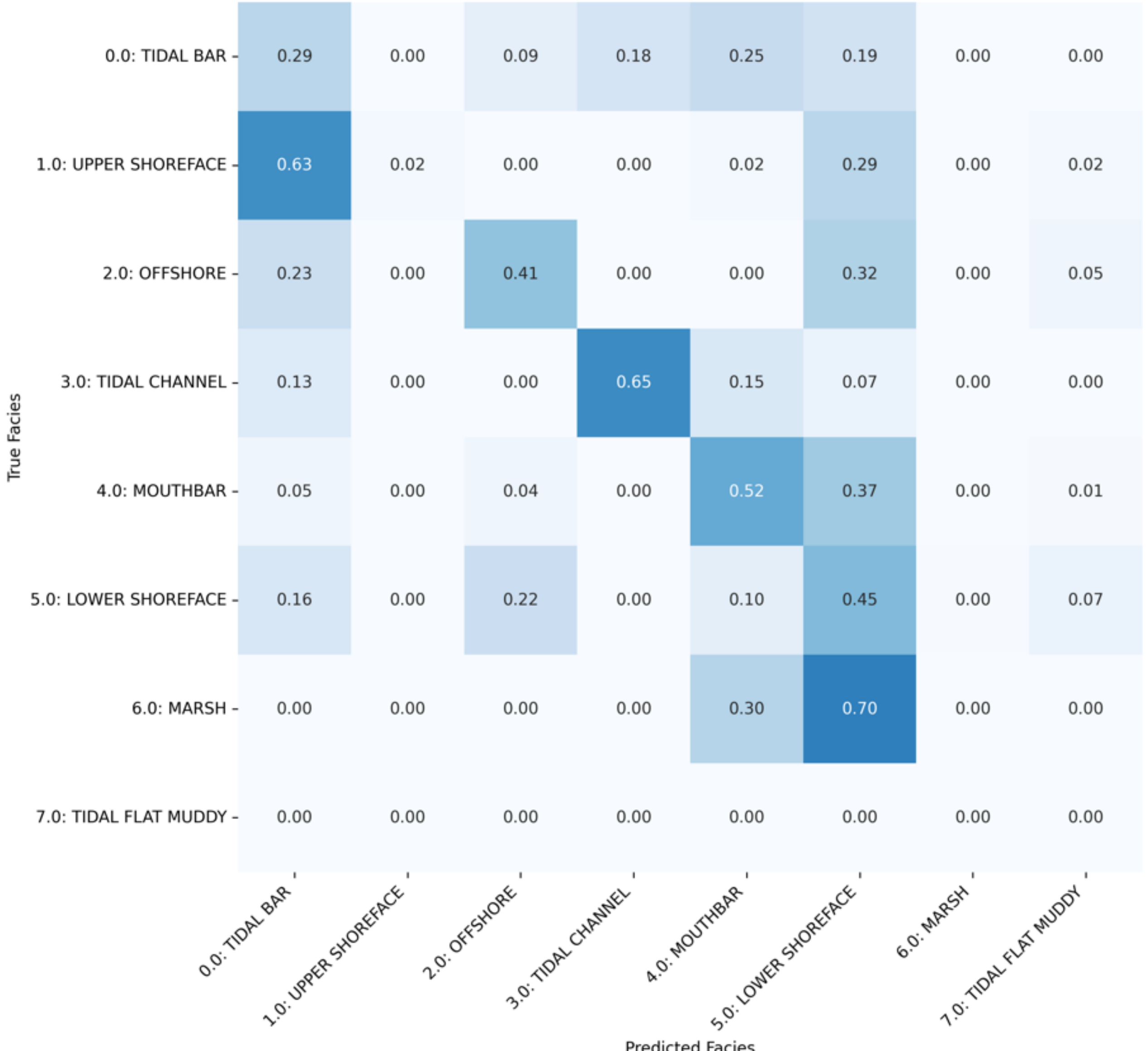
QUESTION 1

Feature set: GR, NPHI, RHOB, DT, RT, RD



QUESTION 1

Feature set: GR, NPHI, RHOB, DT, RT, RD



QUESTION 1

Feature set: DT, RHOB, NPHI

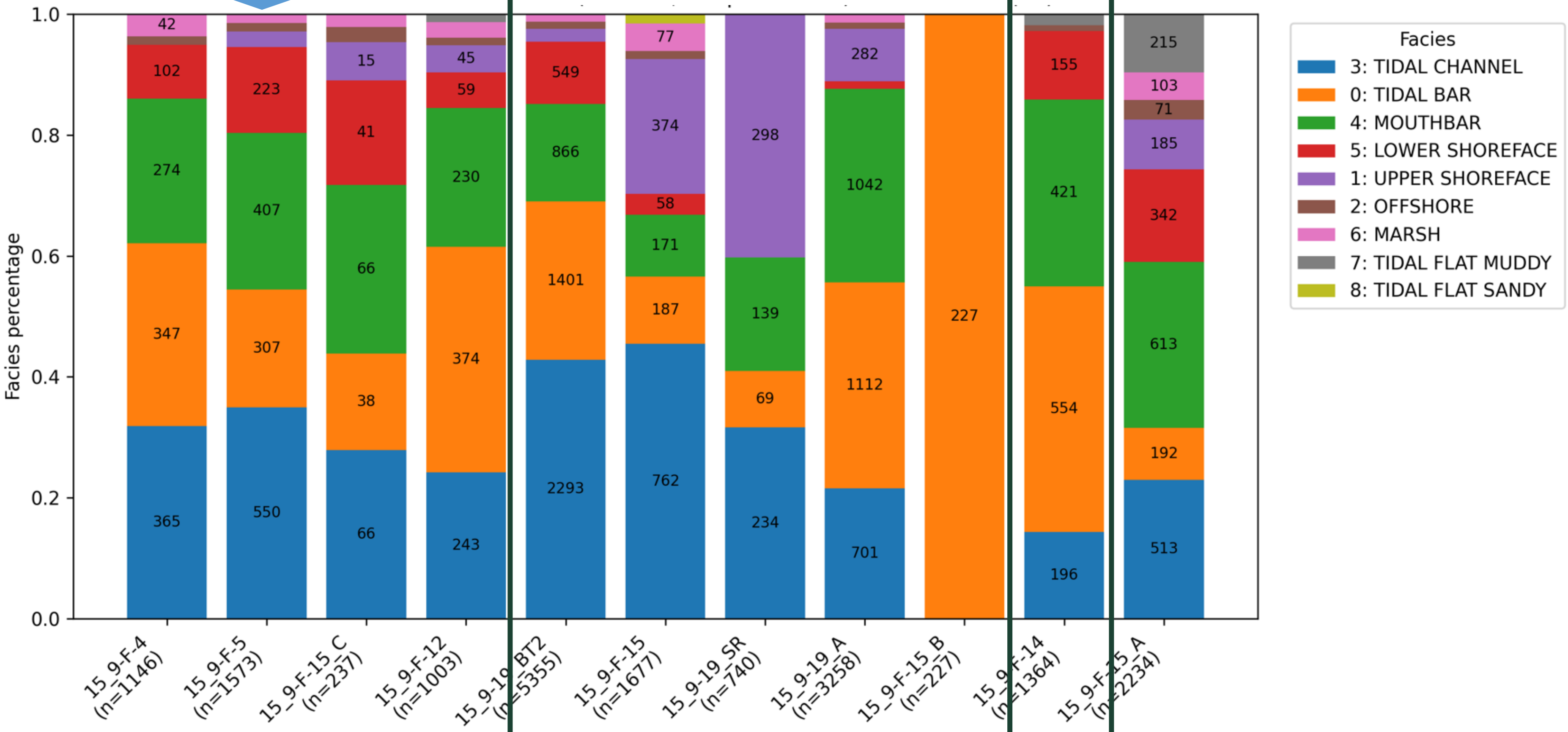
Five wells are available for this feature set

Well F-5 chosen as a test well because it is closest to the overall class mix (and a recorded problem with log distribution of F-12)

Available wells

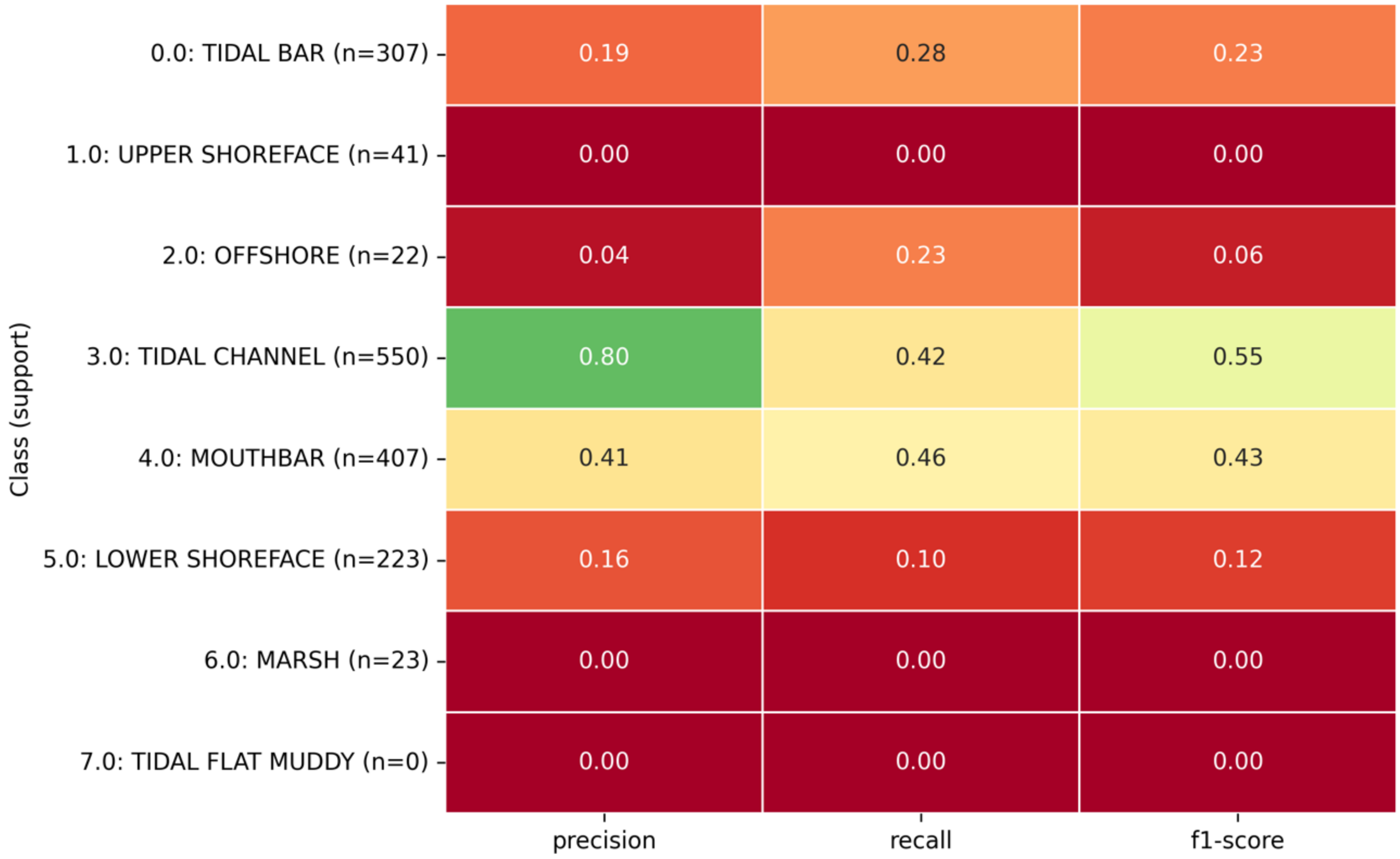
Available wells

Test well



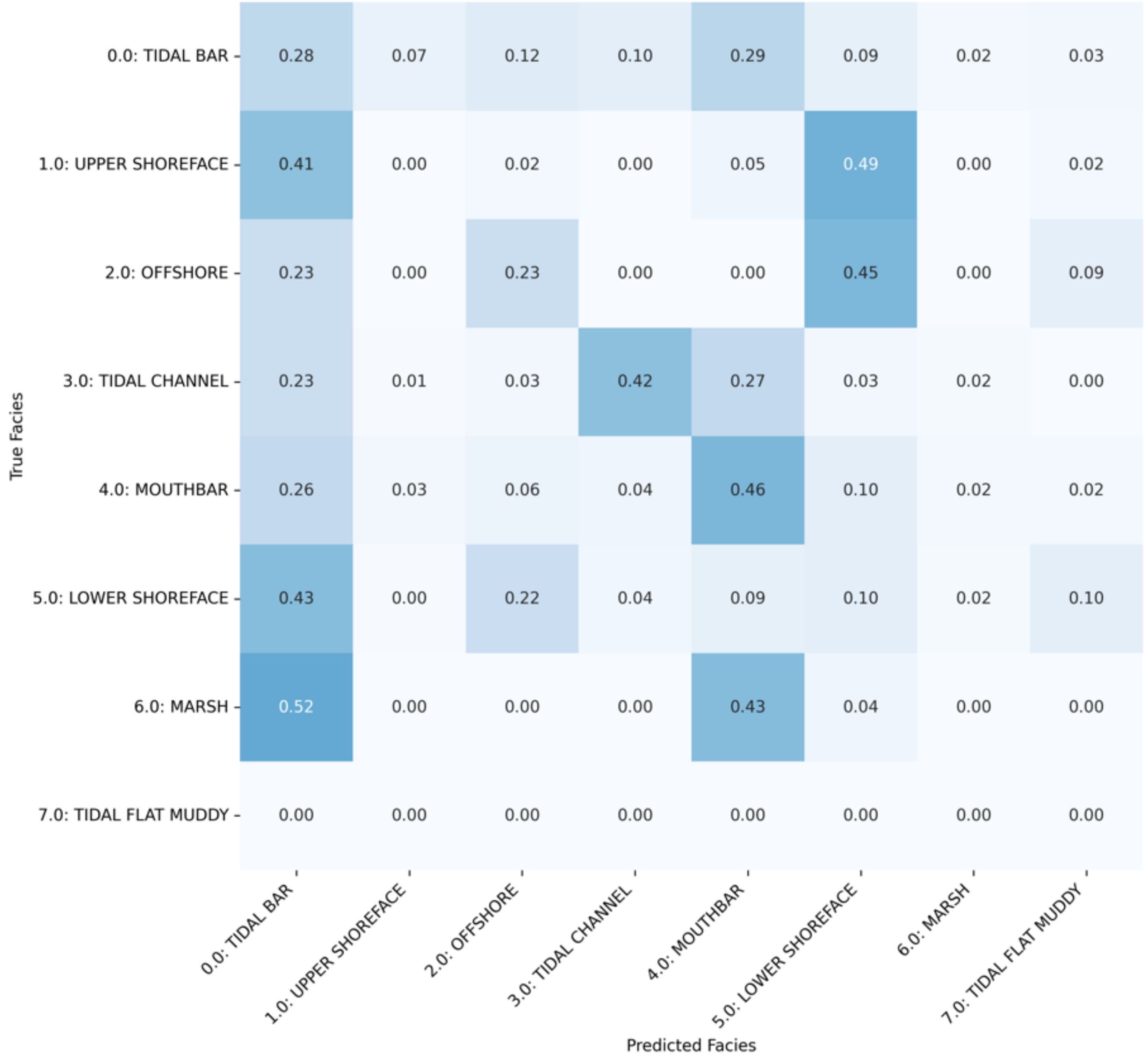
QUESTION 1

Feature set: DT, RHOB, NPHI



QUESTION 1

Feature set: DT, RHOB, NPHI

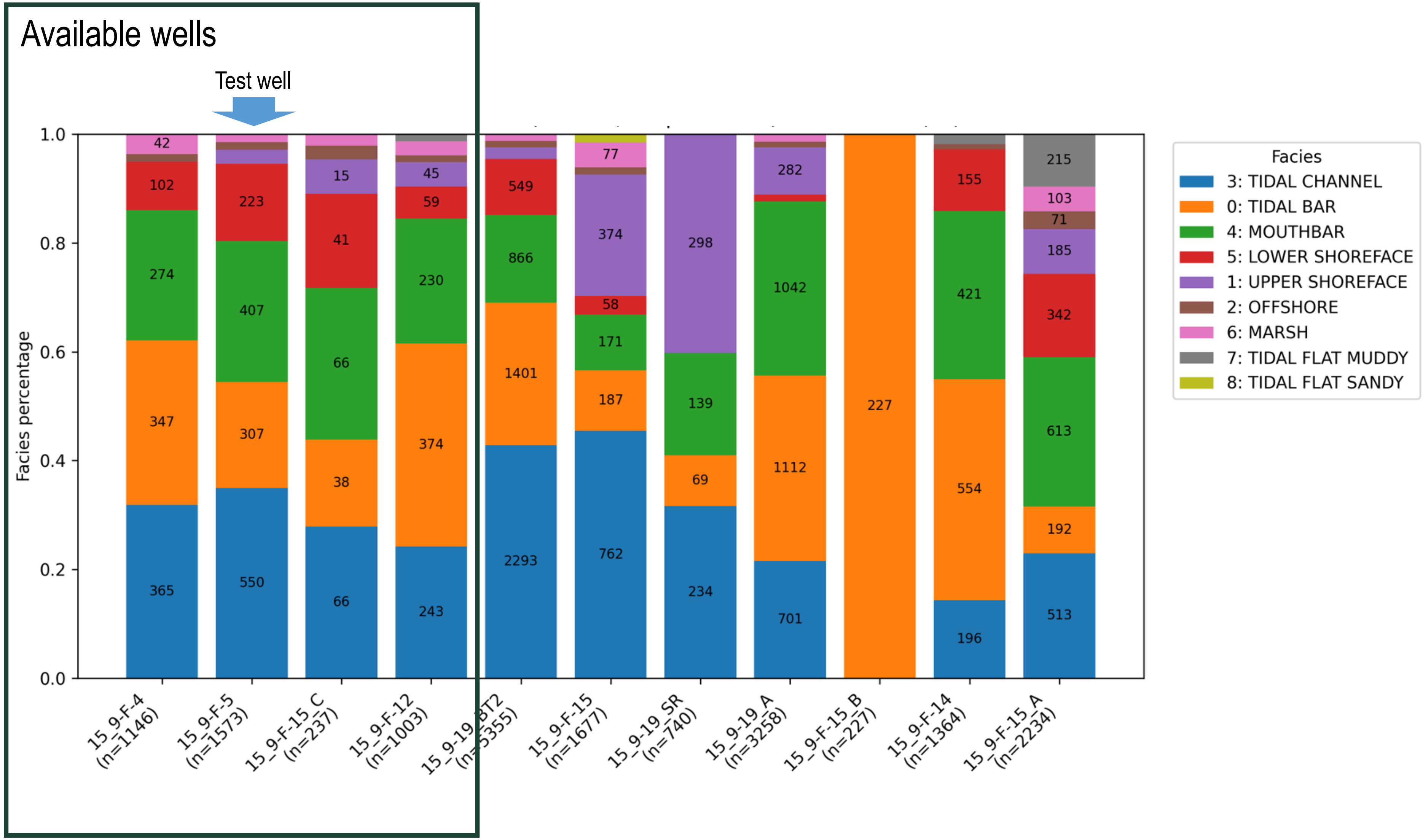


QUESTION 1

Feature set: GR, VSH, NPHI, RHOB, DT, PHIF, KLOGH

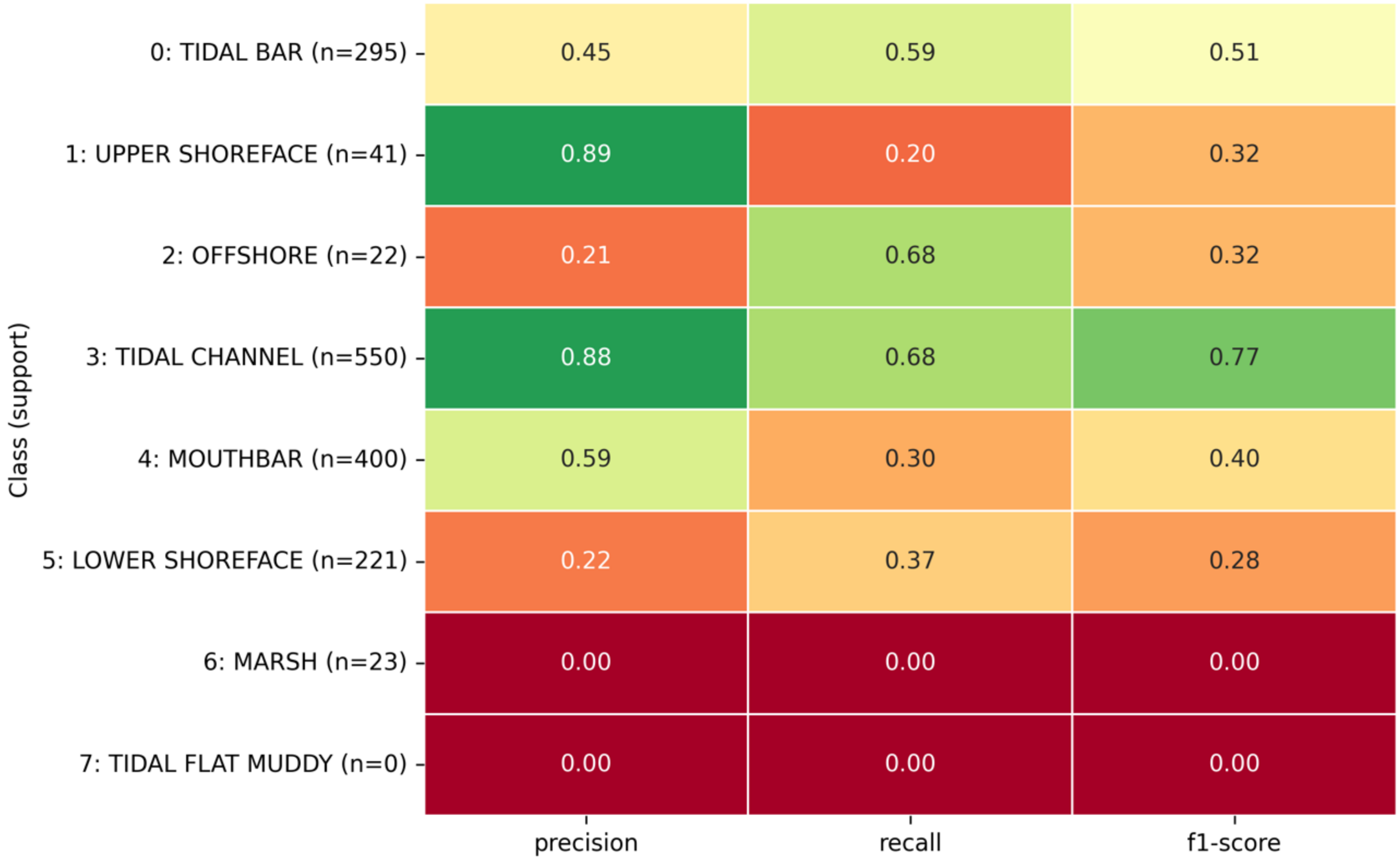
Four wells are available for this feature set

Well F-5 chosen as a test well because it is closest to the overall class mix



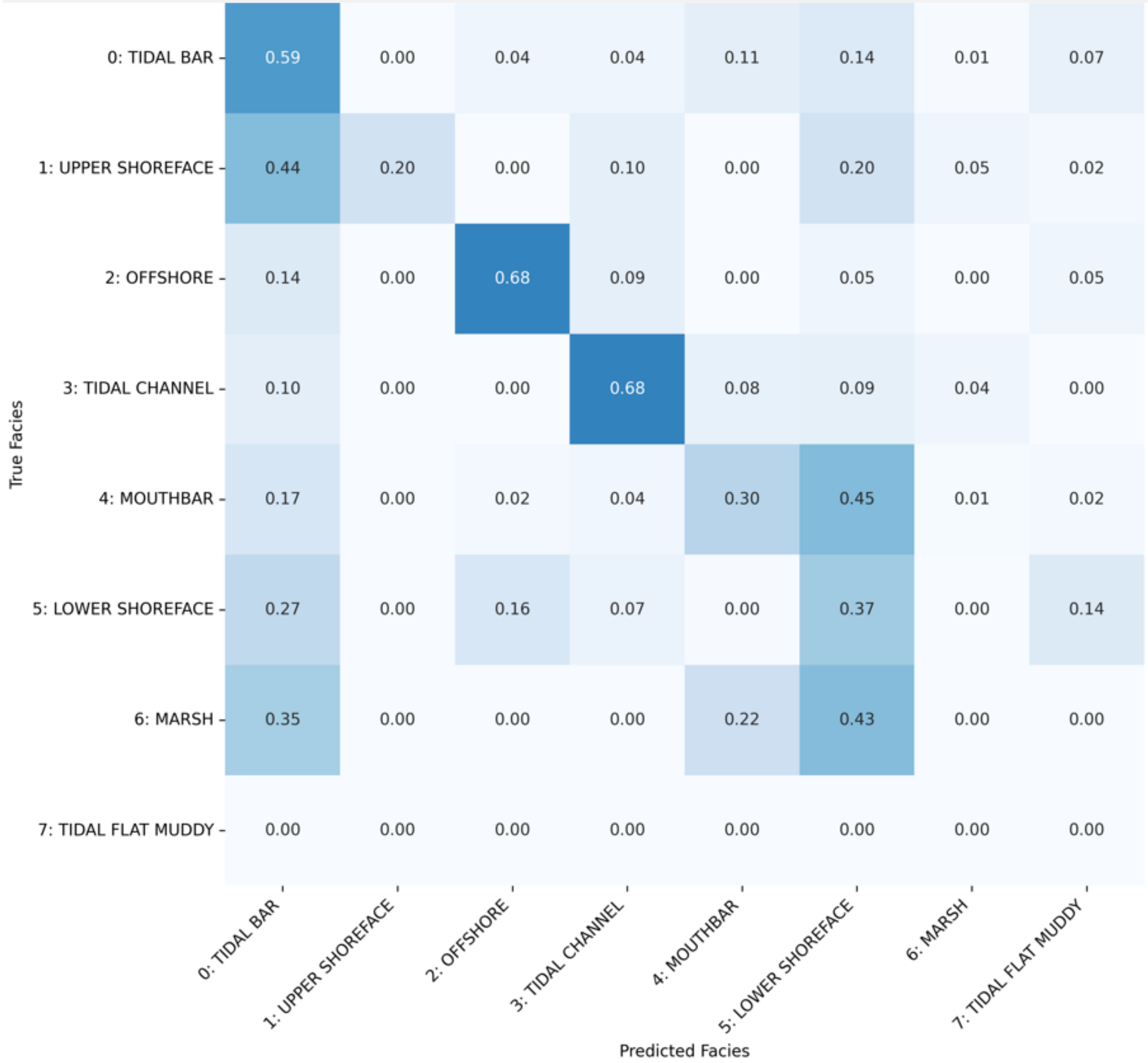
QUESTION 2

Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH + lithology



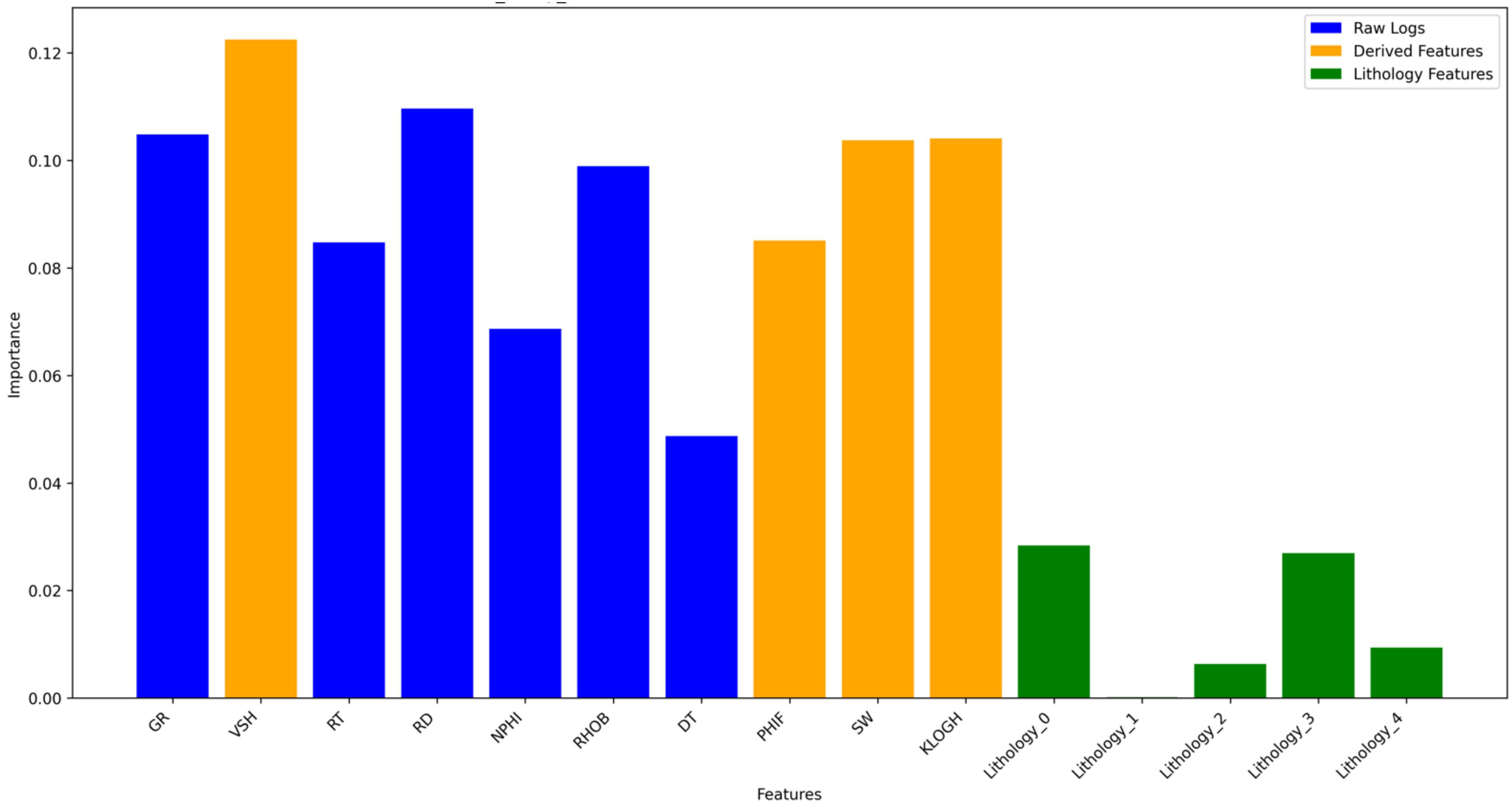
QUESTION 2

Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH + lithology



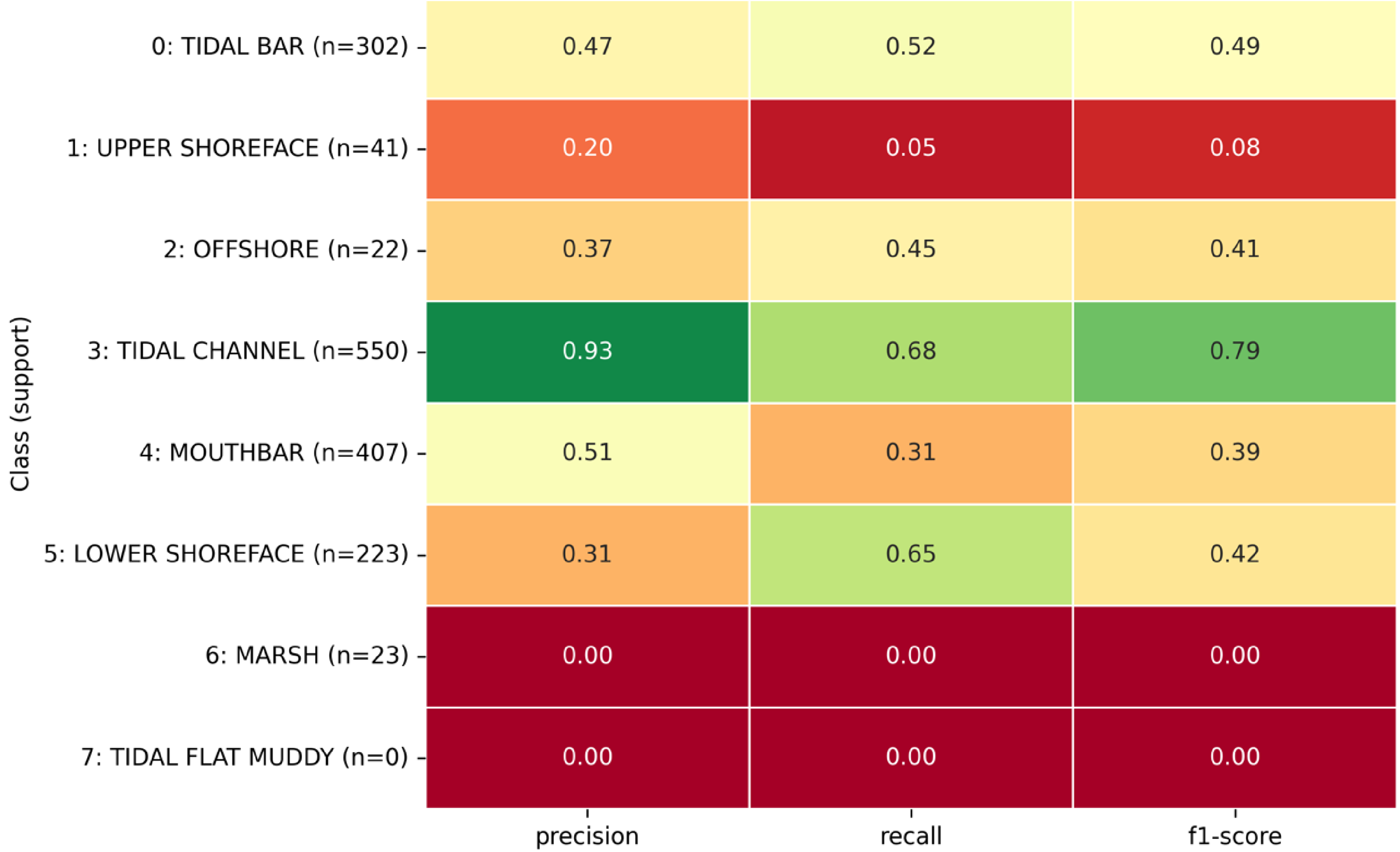
QUESTION 2

Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH + lithology



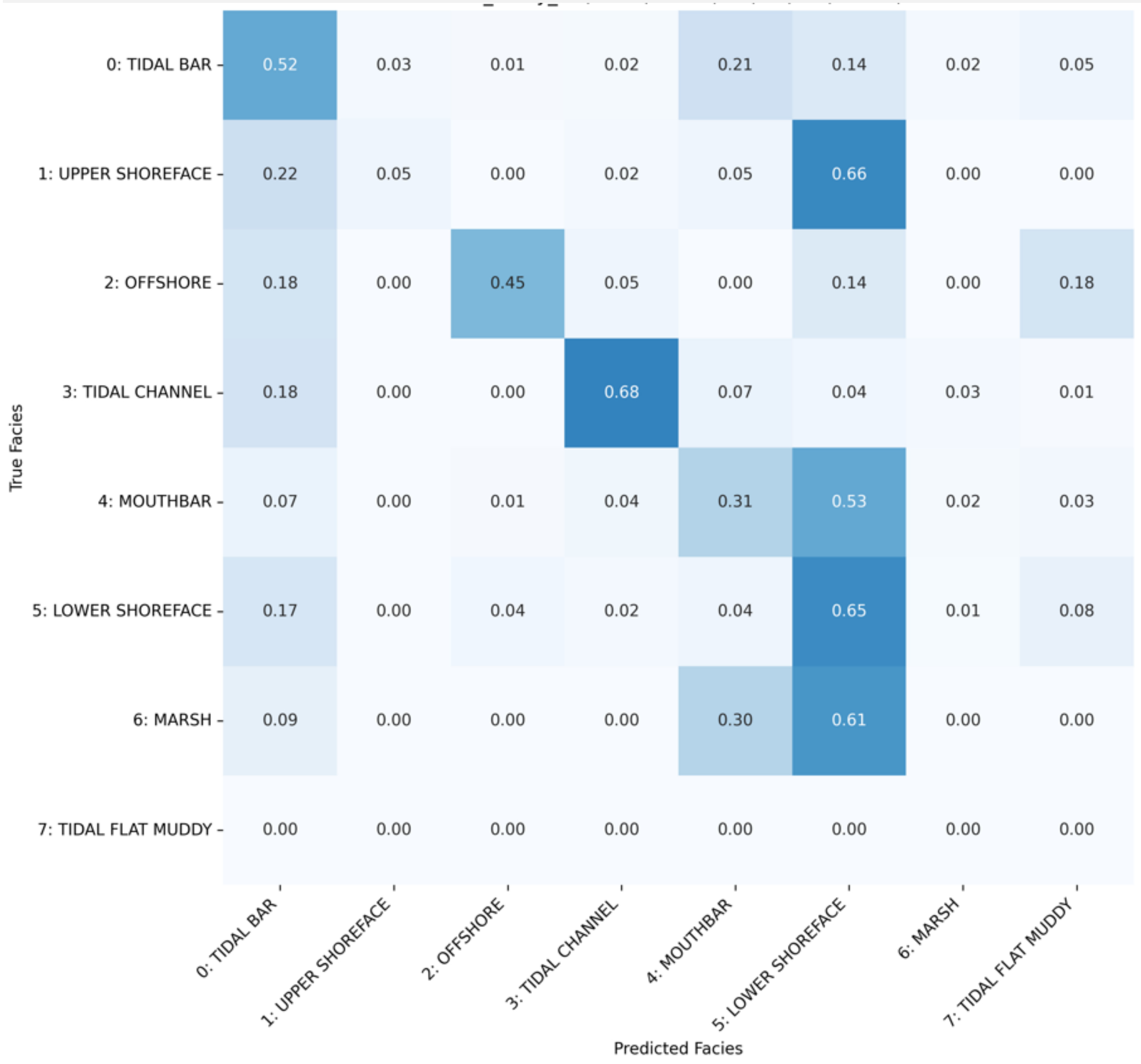
QUESTION 2

Feature set: GR, NPHI, RHOB, DT (plus RT, RD) + Lithology



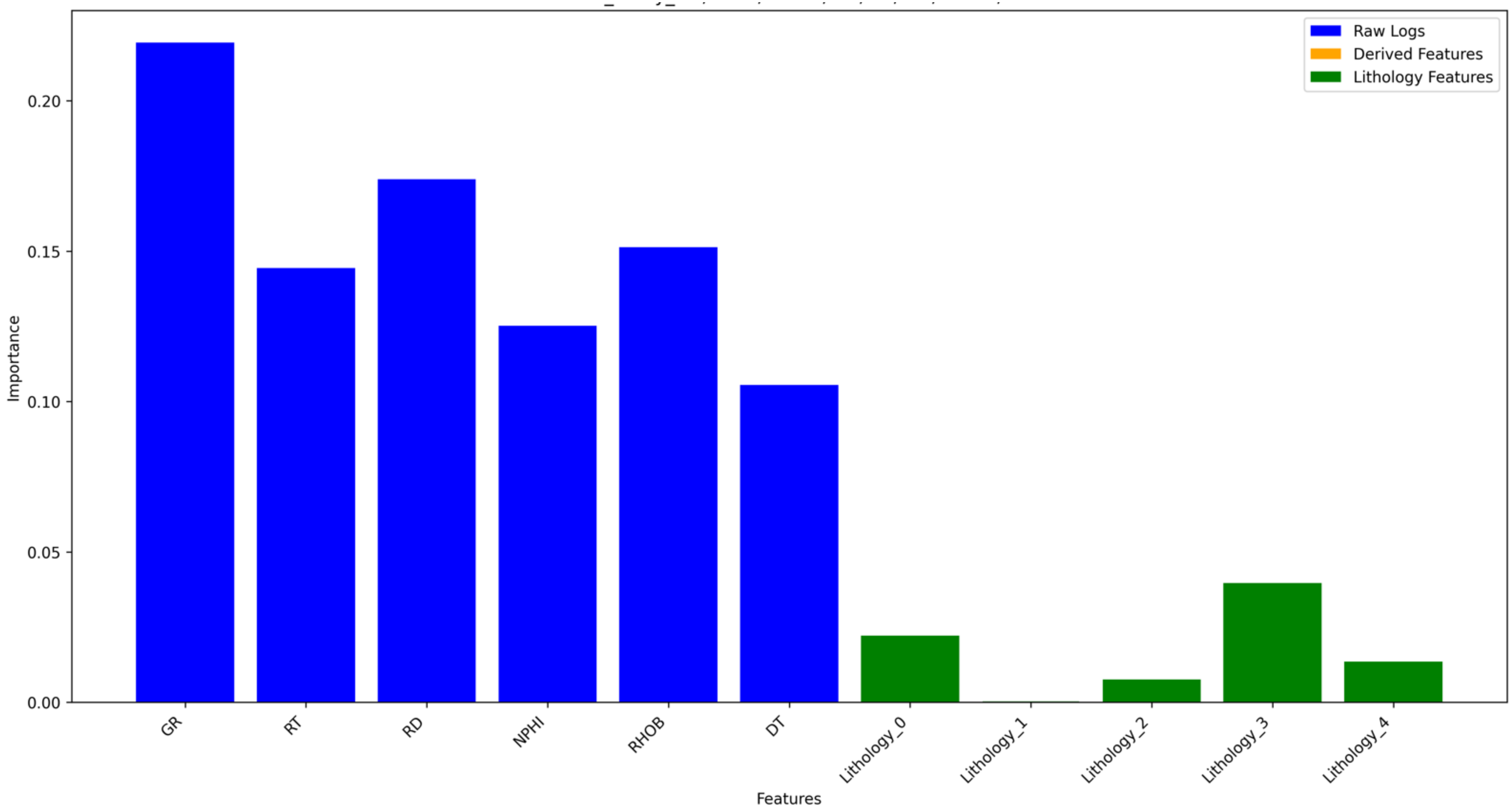
QUESTION 2

Feature set: GR, NPHI, RHOB, DT (plus RT, RD) + Lithology



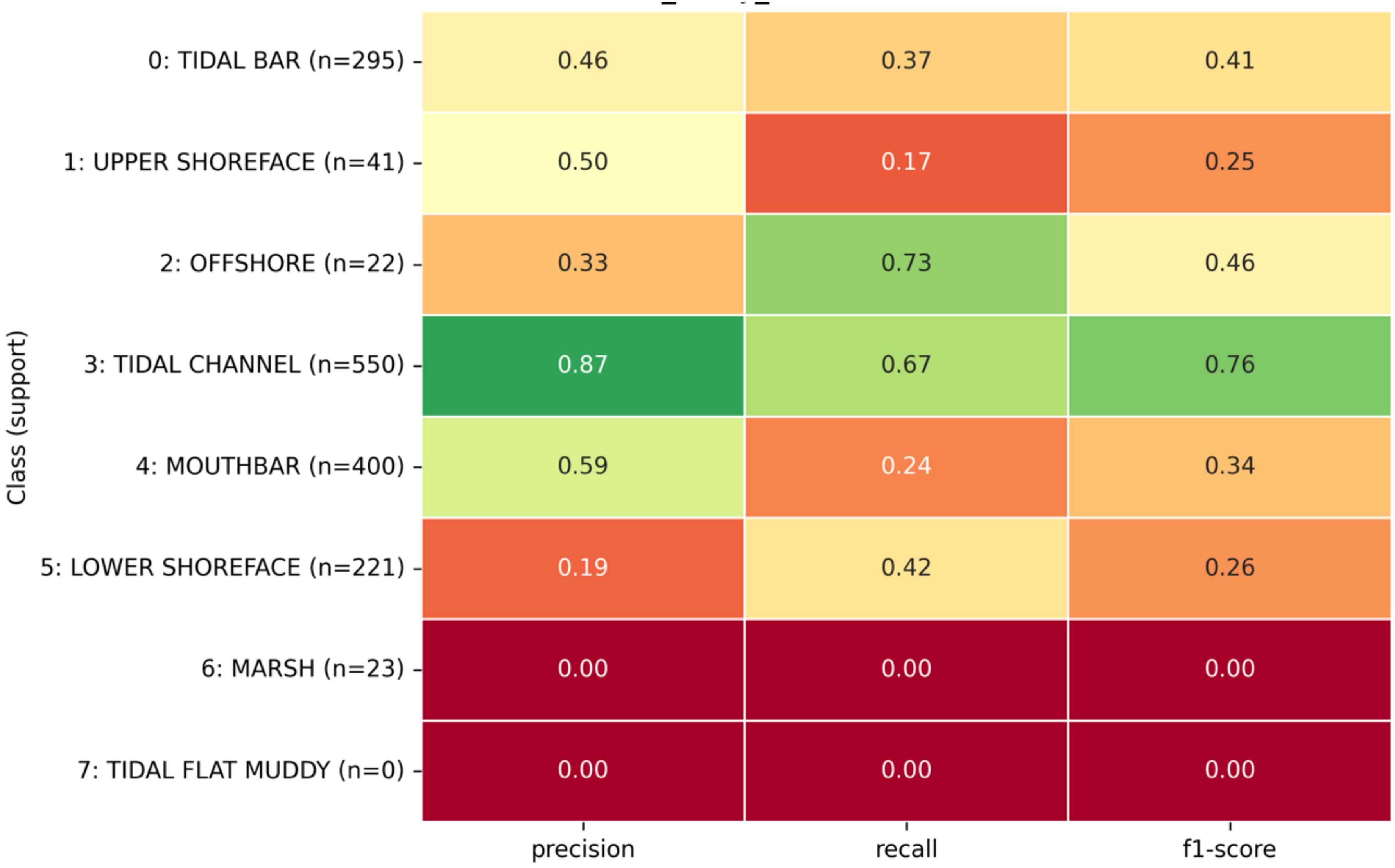
QUESTION 2

Feature set: GR, NPHI, RHOB, DT (plus RT, RD) + Lithology



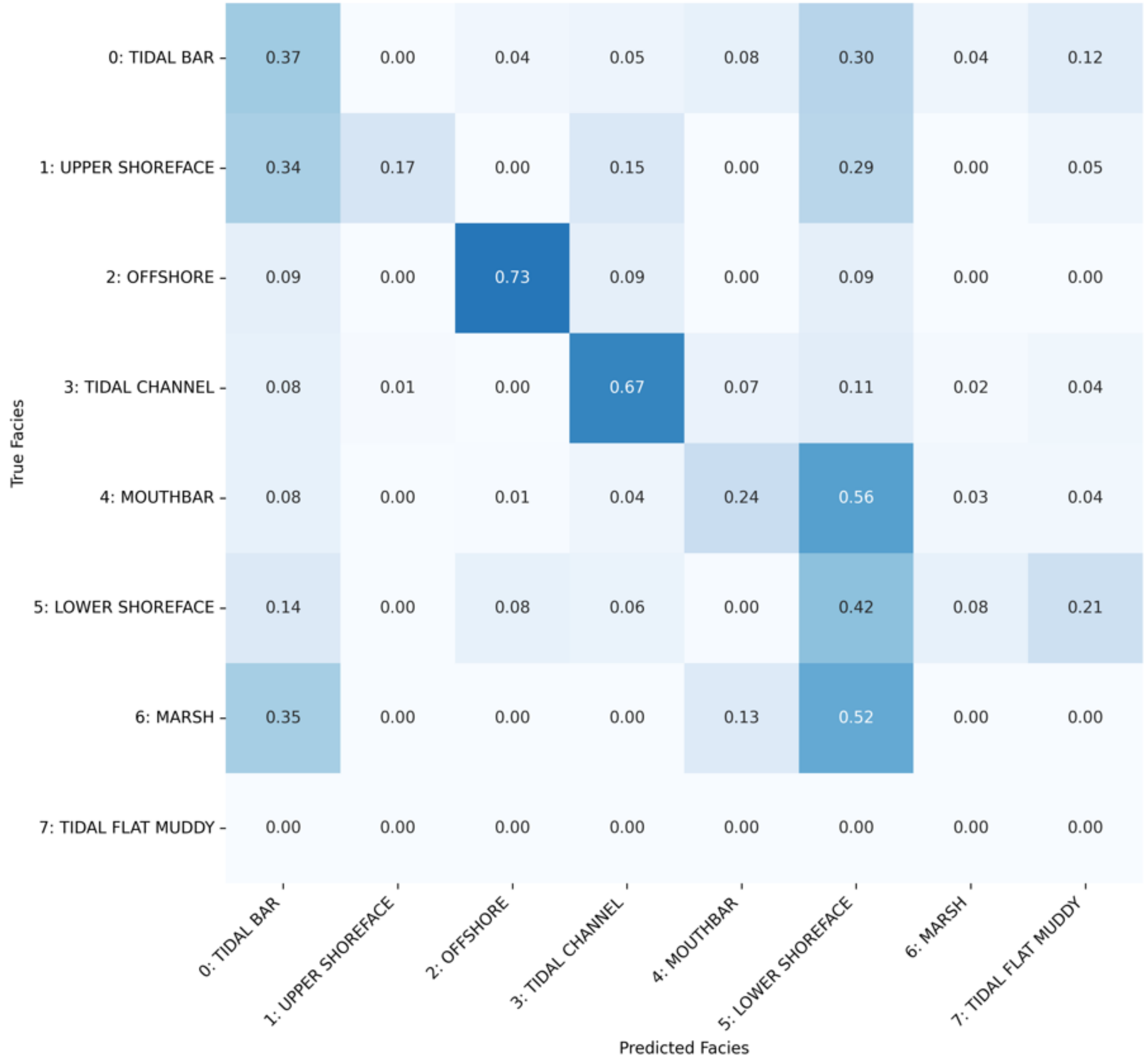
QUESTION 2

Feature set: VSH, PHIF, SW, KLOGH



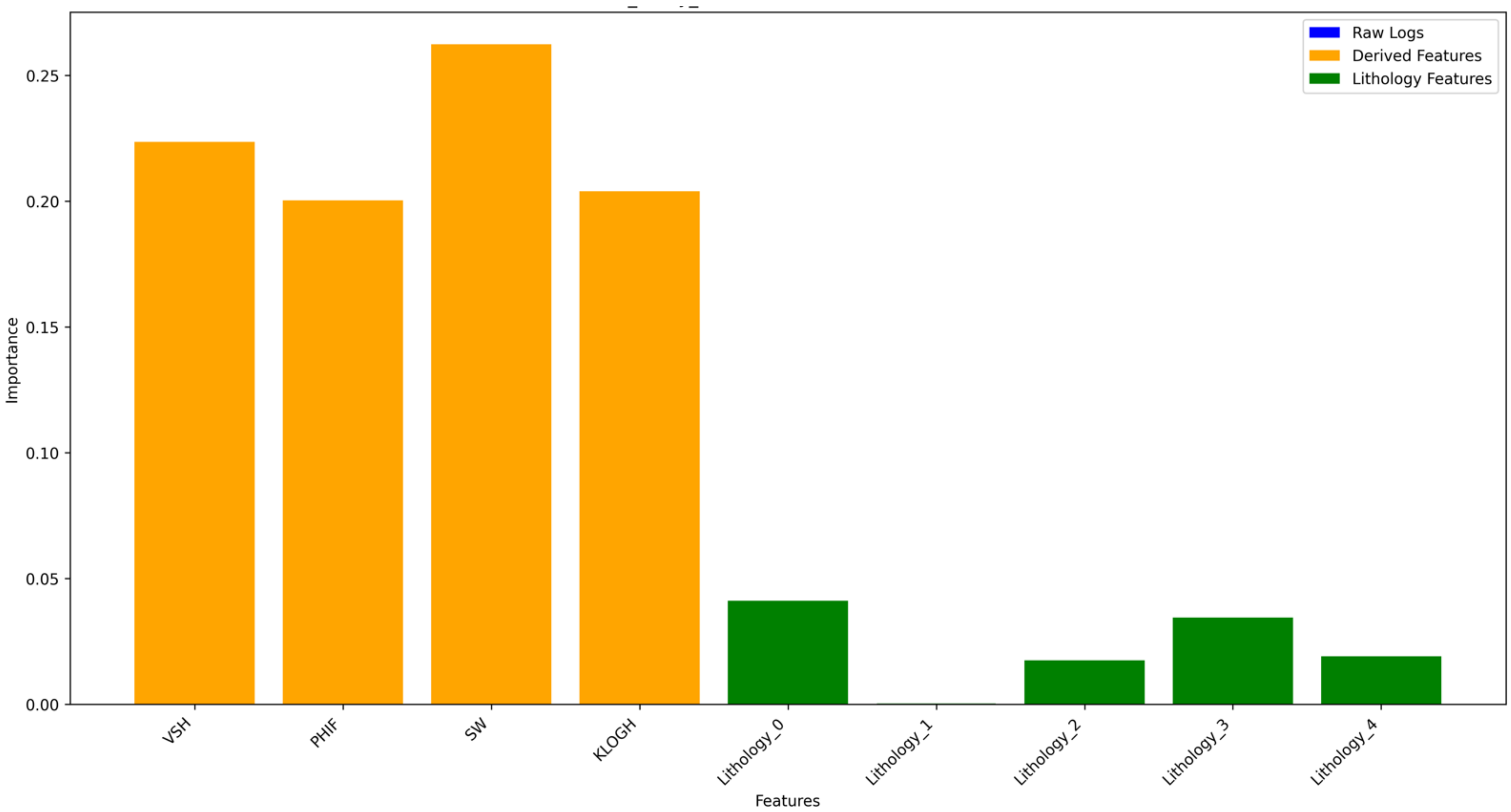
QUESTION 2

Feature set: VSH, PHIF, SW, KLOGH



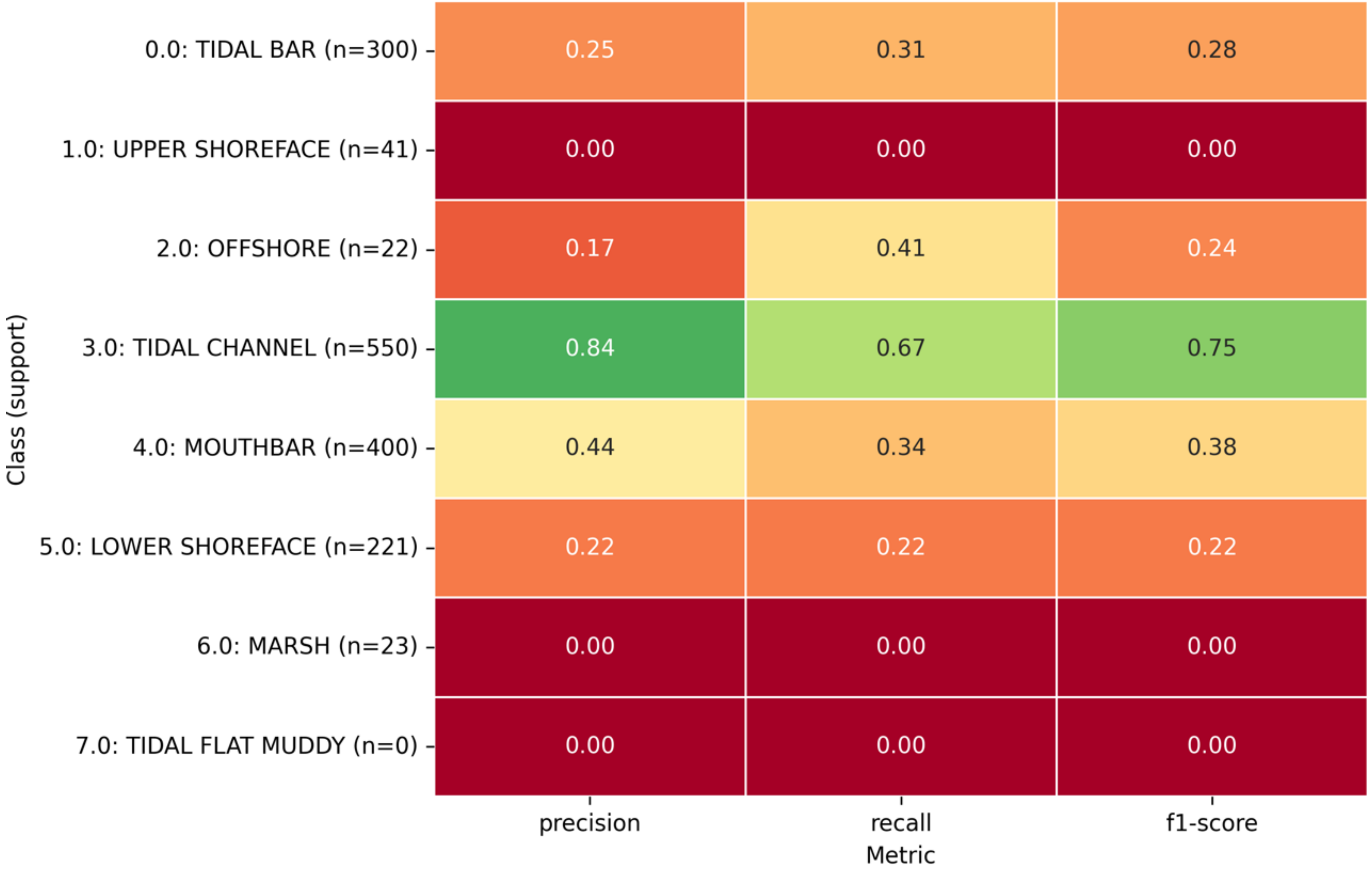
QUESTION 2

Feature set: VSH, PHIF, SW, KLOGH



QUESTION 2

Feature set: VSH, PHIF, DT



QUESTION 2

Feature set: VSH, PHIF, DT

QUESTION 3

Does adding engineered features help?

+ Engineered features:

GR slope: d(GR)/dz in 3.5 m window to capture coarsening up (Mouth Bar) vs fining up (Tidal Bar, Tidal Channel)

Script implementation

For each well:

- Sort data by depth

For each depth point:

- Define window boundaries (±1.75m from current point)
- Find all points within window

If ≥3 points available:

- Perform linear regression (least squares regression)
- Extract slope coefficient (d(GR)/dz)

QUESTION 3

Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH

+ Engineered features:

GR slope: d(GR)/dz in 3.5 m window to capture coarsening up (Mouth Bar) vs fining up (Tidal Bar, Tidal Channel)

Without the GR slope feature

Class (support)	precision	recall	f1-score
0: TIDAL BAR (n=295)	0.45	0.59	0.51
1: UPPER SHOREFACE (n=41)	0.89	0.20	0.32
2: OFFSHORE (n=22)	0.21	0.68	0.32
3: TIDAL CHANNEL (n=550)	0.88	0.68	0.77
4: MOUTHBAR (n=400)	0.59	0.30	0.40
5: LOWER SHOREFACE (n=221)	0.22	0.37	0.28
6: MARSH (n=23)	0.00	0.00	0.00
7: TIDAL FLAT MUDDY (n=0)	0.00	0.00	0.00

With GR slope feature

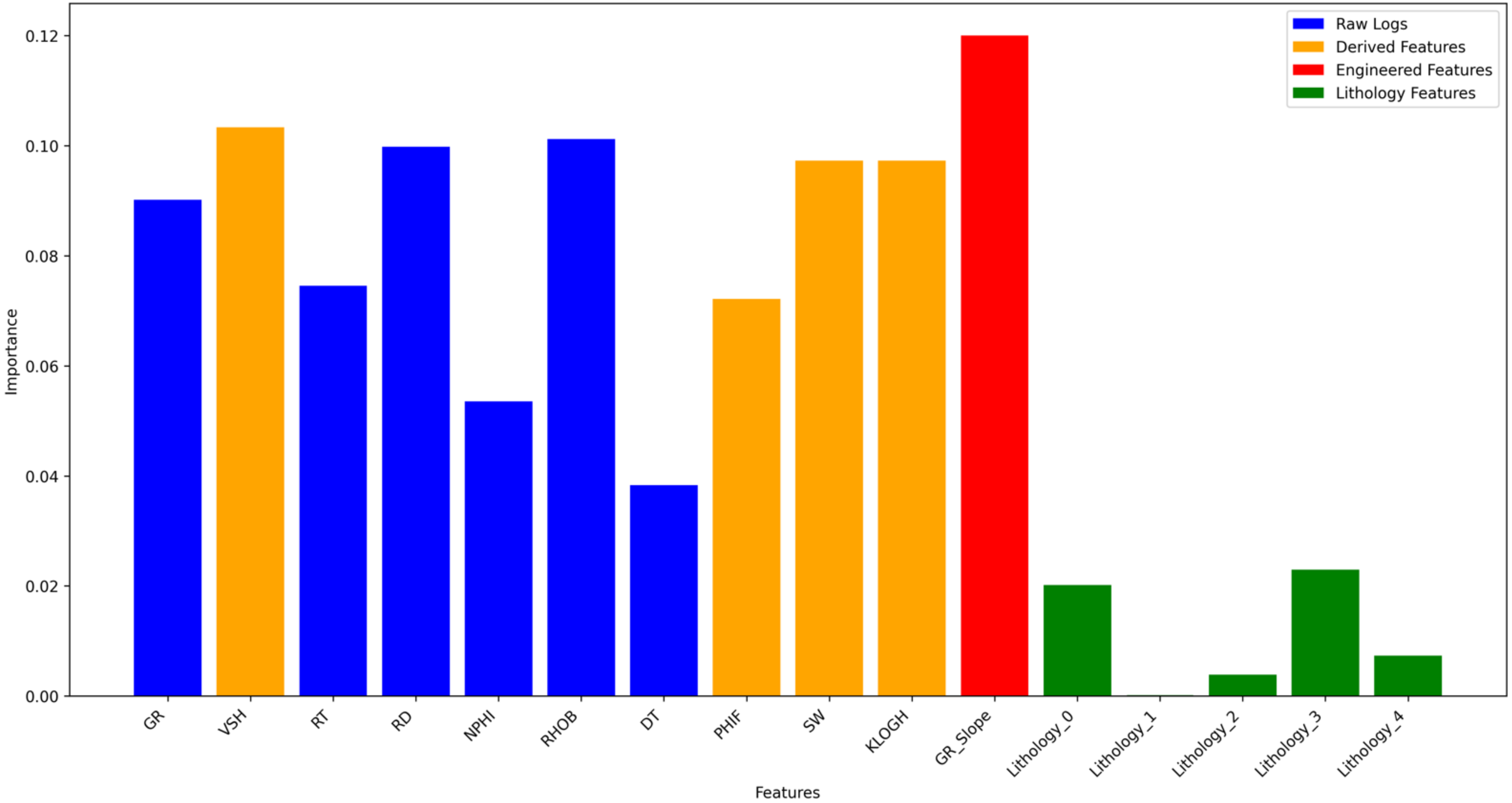
Class (support)	precision	recall	f1-score
0.0: TIDAL BAR (n=295)	0.46	0.63	0.53
1.0: UPPER SHOREFACE (n=41)	1.00	0.15	0.26
2.0: OFFSHORE (n=22)	0.20	0.59	0.30
3.0: TIDAL CHANNEL (n=550)	0.89	0.68	0.77
4.0: MOUTHBAR (n=400)	0.55	0.27	0.36
5.0: LOWER SHOREFACE (n=221)	0.21	0.35	0.26
6.0: MARSH (n=23)	0.00	0.00	0.00
7.0: TIDAL FLAT MUDDY (n=0)	0.00	0.00	0.00

QUESTION 3

Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH

+ Engineered features:

GR slope: d(GR)/dz in 3.5 m window to capture coarsening up (Mouth Bar) vs fining up (Tidal Bar, Tidal Channel)



QUESTION 3

Does adding engineered features help?

+ Engineered features:

GR serration: rolling std (GR) in 3m window because it is higher in heteroliths/mud-draped tidal bars, lower in blocky channels

Script implementation

For each well:

- Sort data by depth

For each depth point:

- Define window boundaries (±1.5m from current point)
- Find all points within window
- Create paired arrays
- Calculate means
- Calculate slope with least squares formula

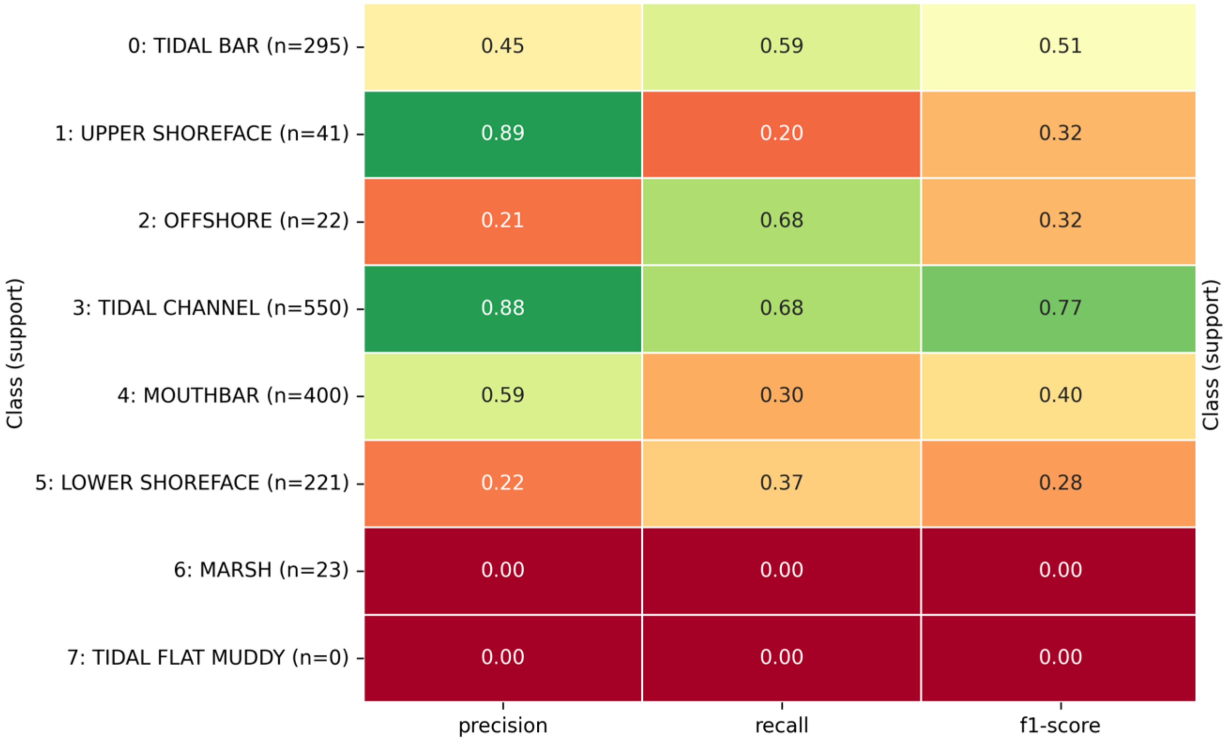
QUESTION 3

Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH

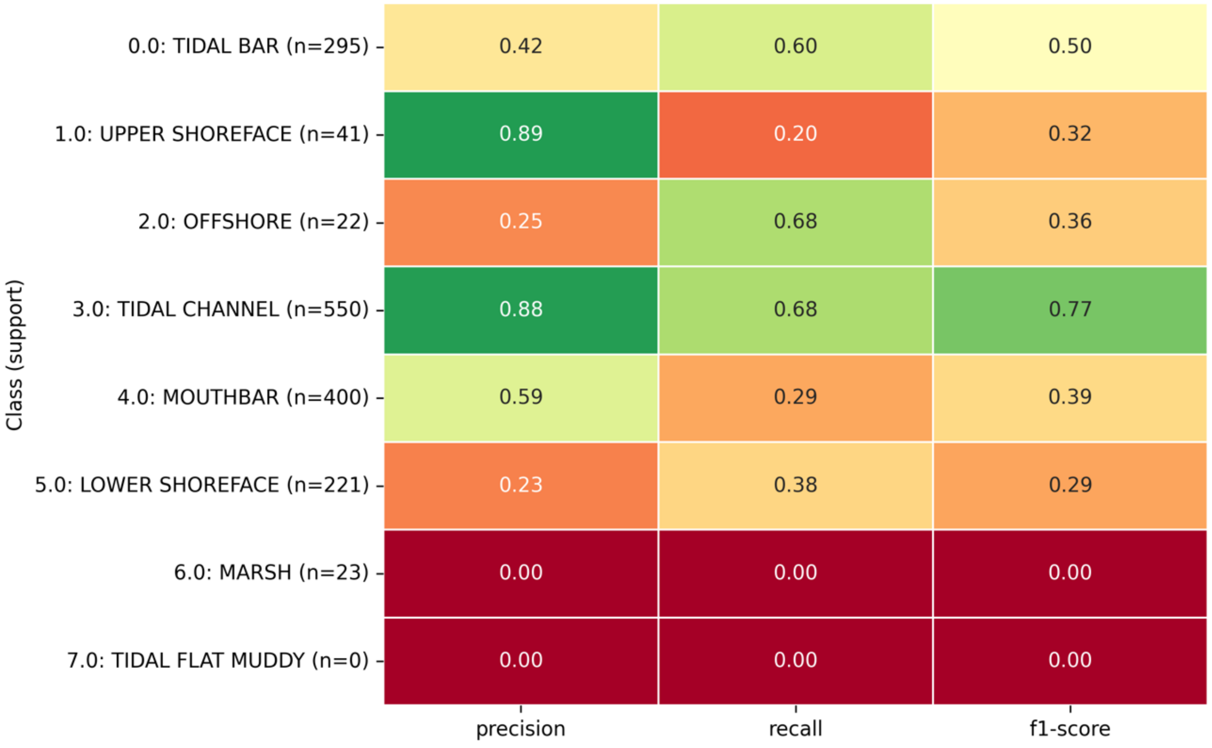
+ Engineered features:

GR serration: rolling std (GR) in 3m window because it is higher in heteroliths/mud-draped tidal bars, lower in blocky channels

Without the GR serration feature



With GR serration feature

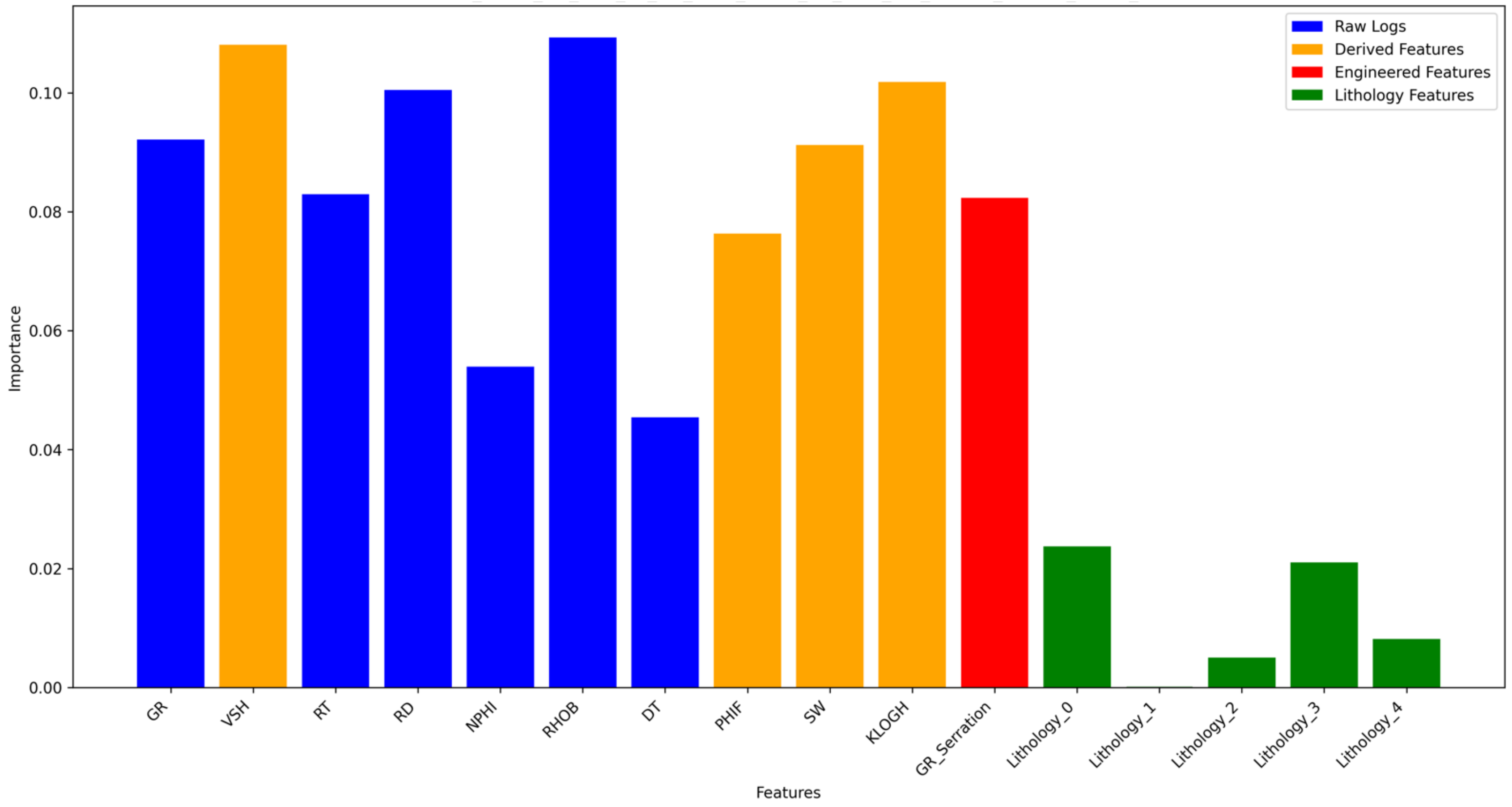


QUESTION 3

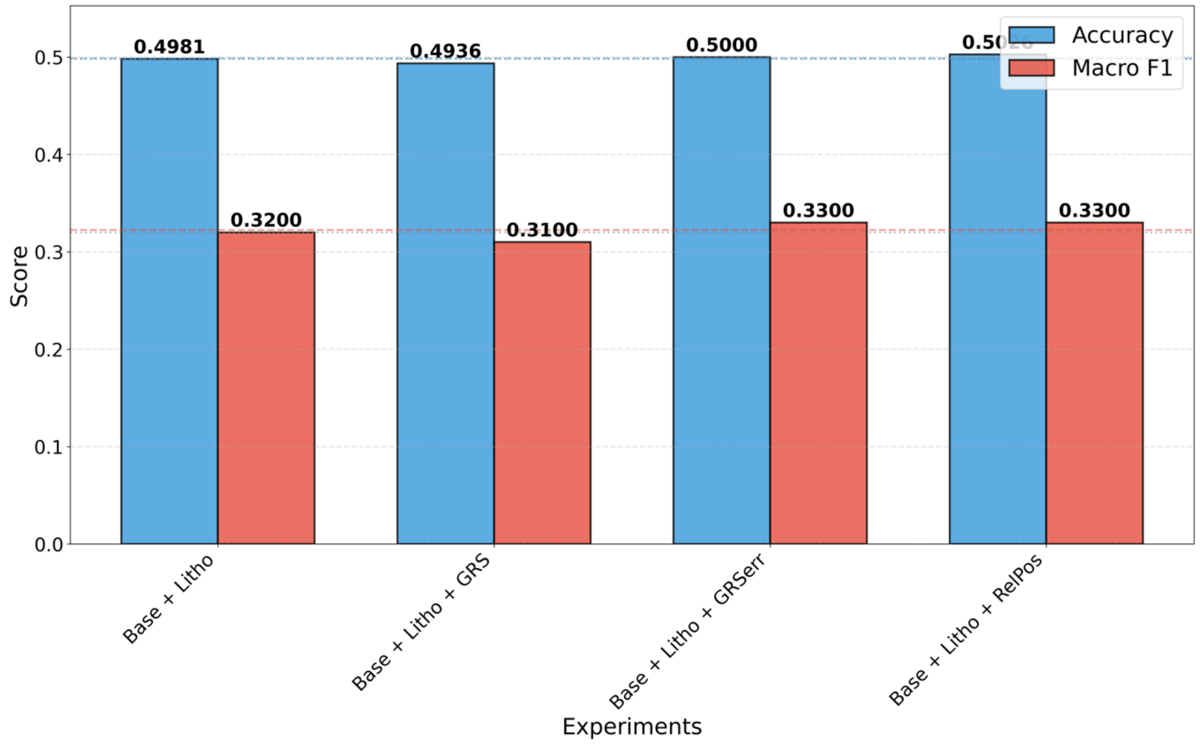
Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH

+ Engineered features:

GR serration: rolling std (GR) in 3m window because it is higher in heteroliths/mud-draped tidal bars, lower in blocky channels



QUESTION 3



Accuracy
Macro F1
0.4981
0.4936
0.5000
0.50
0.3200
0.3100
0.3300
0.3300
Score
0.0
0.1
0.2
0.3
0.4
0.5
Base + Litho
Base + Litho + GRS
Base + Litho + GRSerr
Base + Litho + RelPos
Experiments

QUESTION 3

Does adding engineered features help?

+ Engineered features:

Relpos: relative position in GR-cycle to help distinguish sands in regressive-transgressive cycles

Script implementation

For each well:

- Sort data by depth
- Extract depth and GR arrays
- Apply GR curve smoothing
- Find turning points
- Define cycles
- Calculate RelPos for each cycle: RelPos = (GR_start - GR_current) / (GR_start - GR_end)

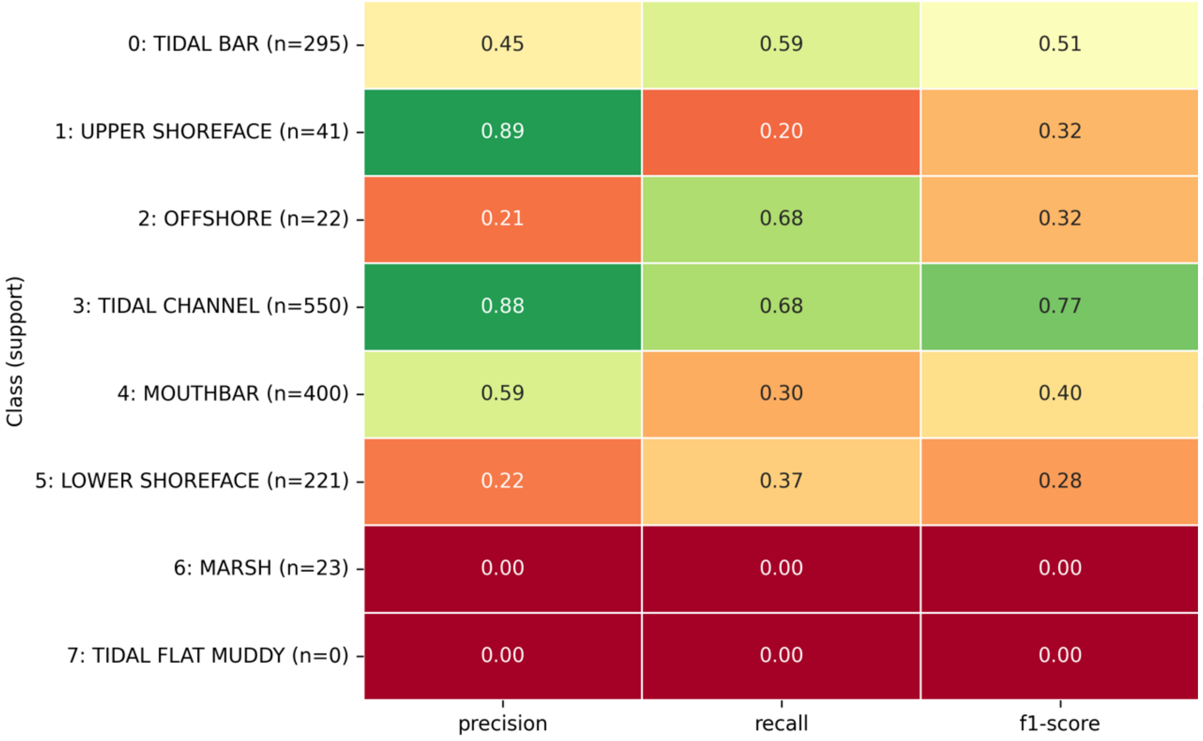
QUESTION 3

Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH

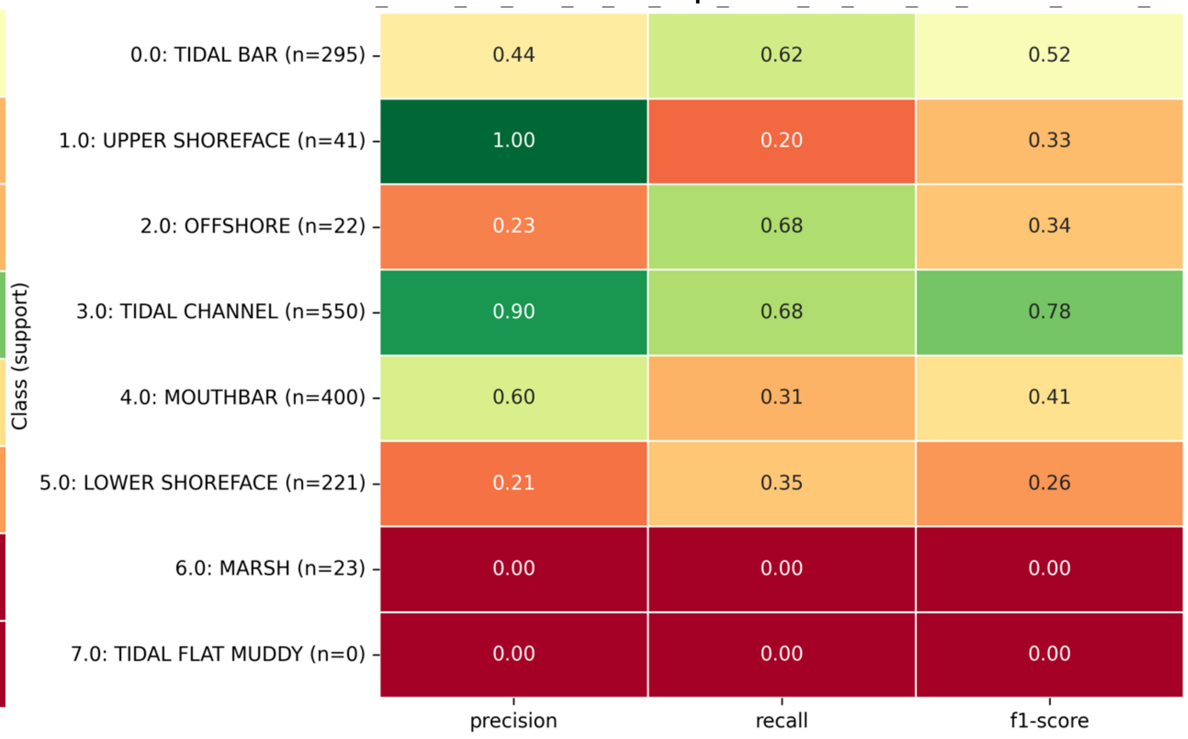
+ Engineered features:

Relpos: relative position in GR-cycle to help distinguish sands in regressive-transgressive cycles

Without Relpos



With Relpos

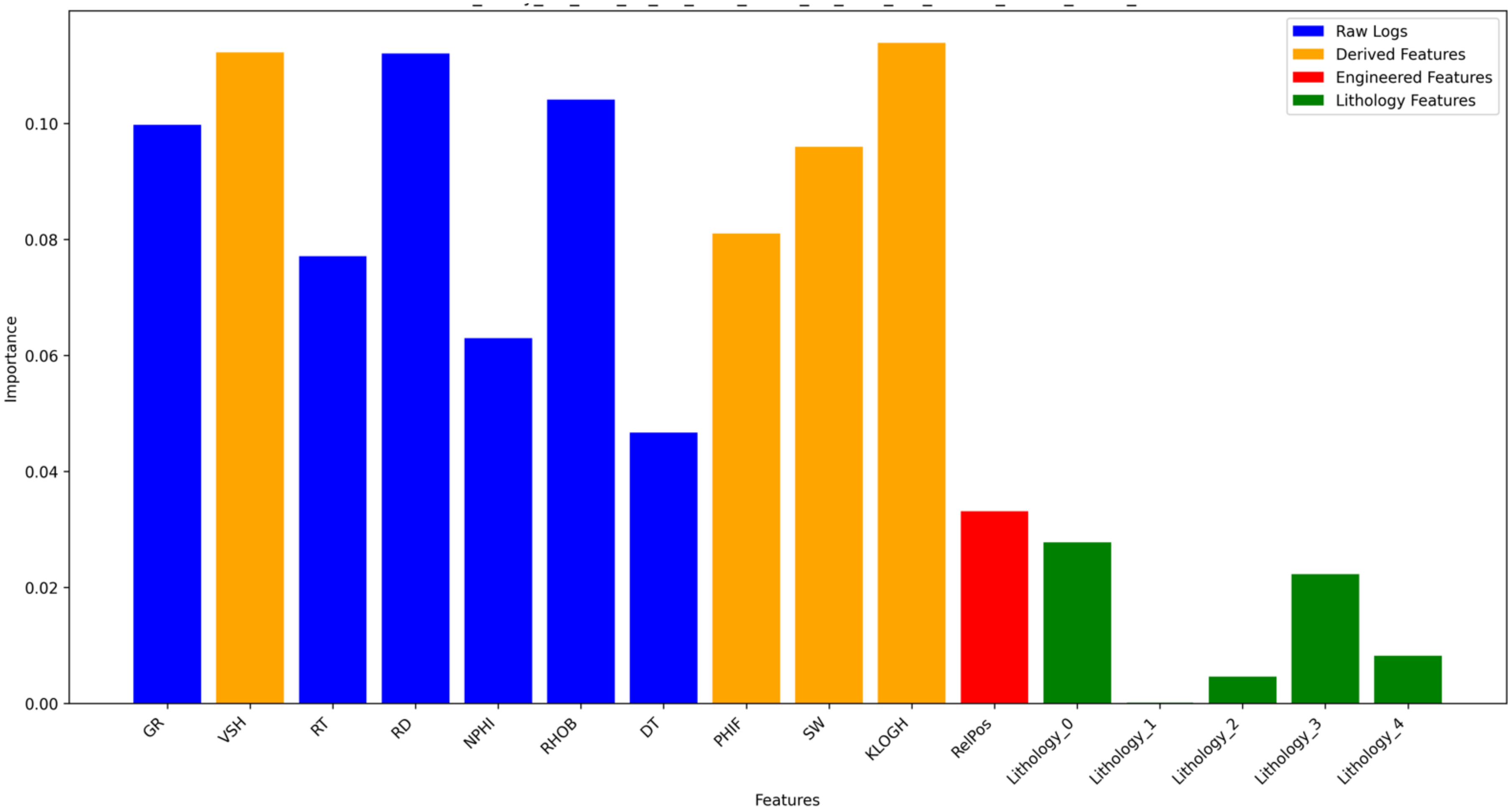


QUESTION 3

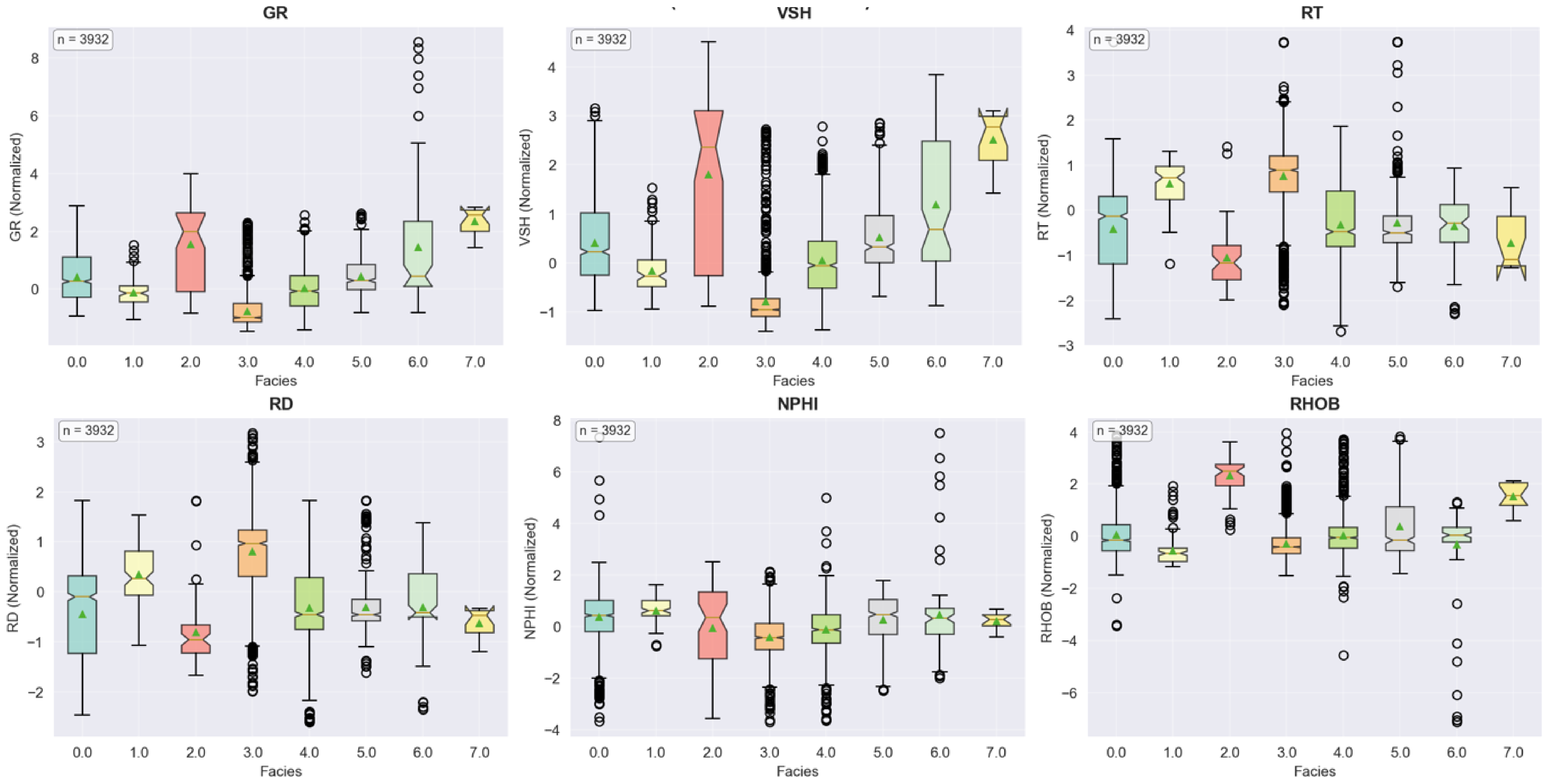
Feature set: GR, VSH, RT, RD, NPHI, RHOB, DT, PHIF, SW, KLOGH

+ Engineered features:

Relpos: relative position in GR-cycle to help distinguish sands in regressive-transgressive cycles



NORMALIZED VALUES



NORMALIZED VALUES