

# Outlier Identification in Semi-Automated Bedrock Surface Dataset

Riley Balikian

Assistant Research Scientist

Illinois State Geological Survey

Prairie Research Institute

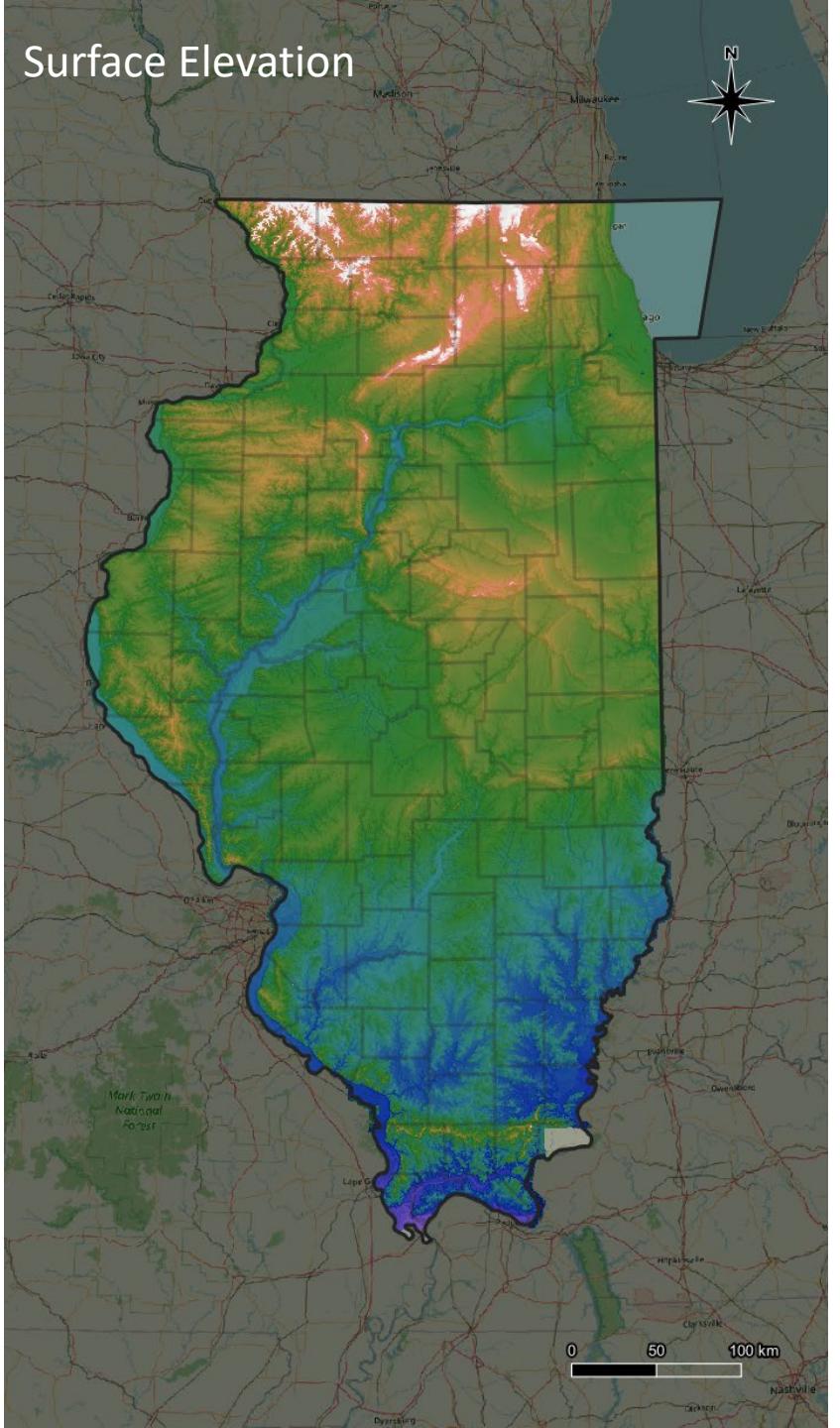
University of Illinois



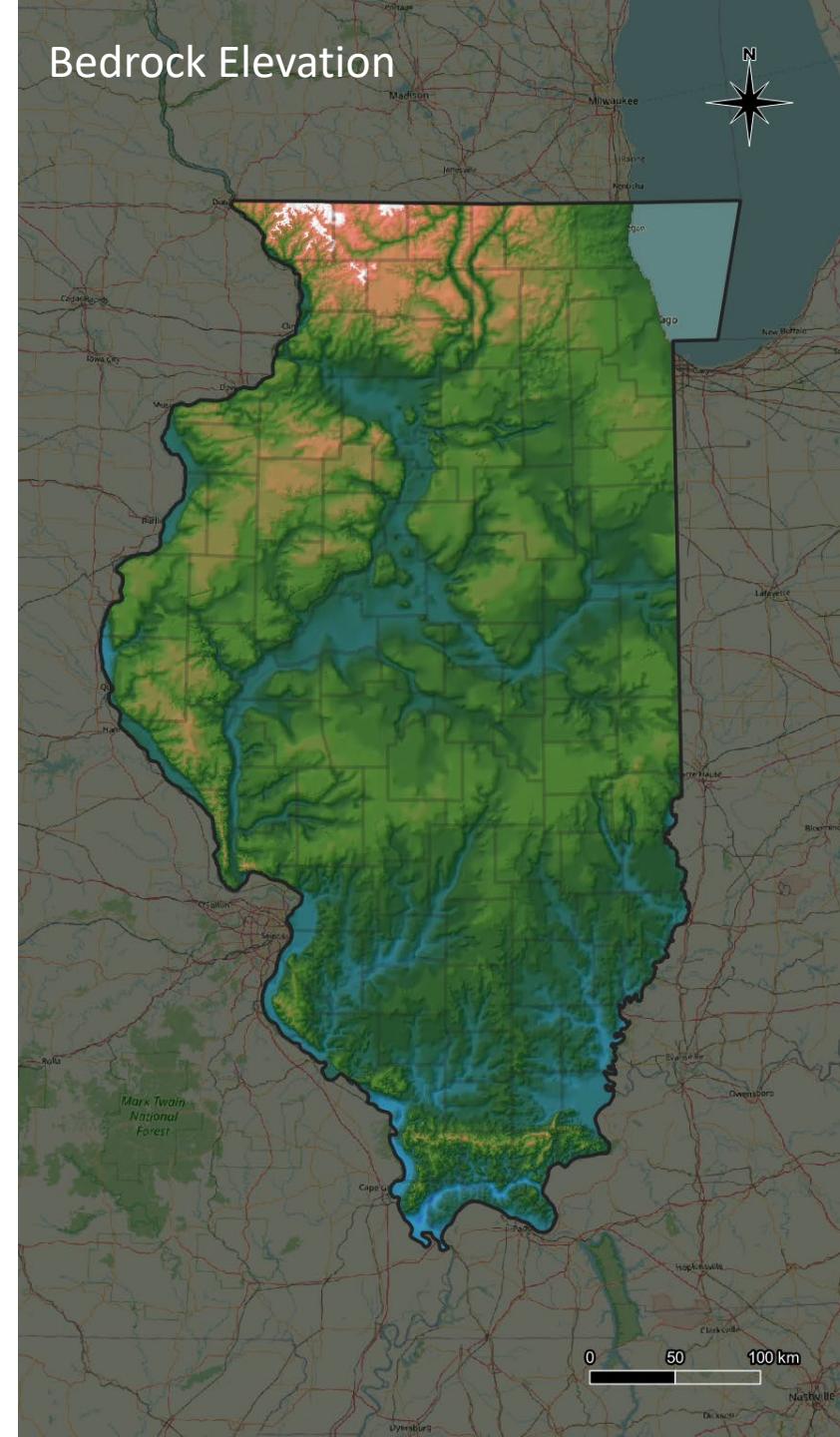
Illinois State Geological Survey

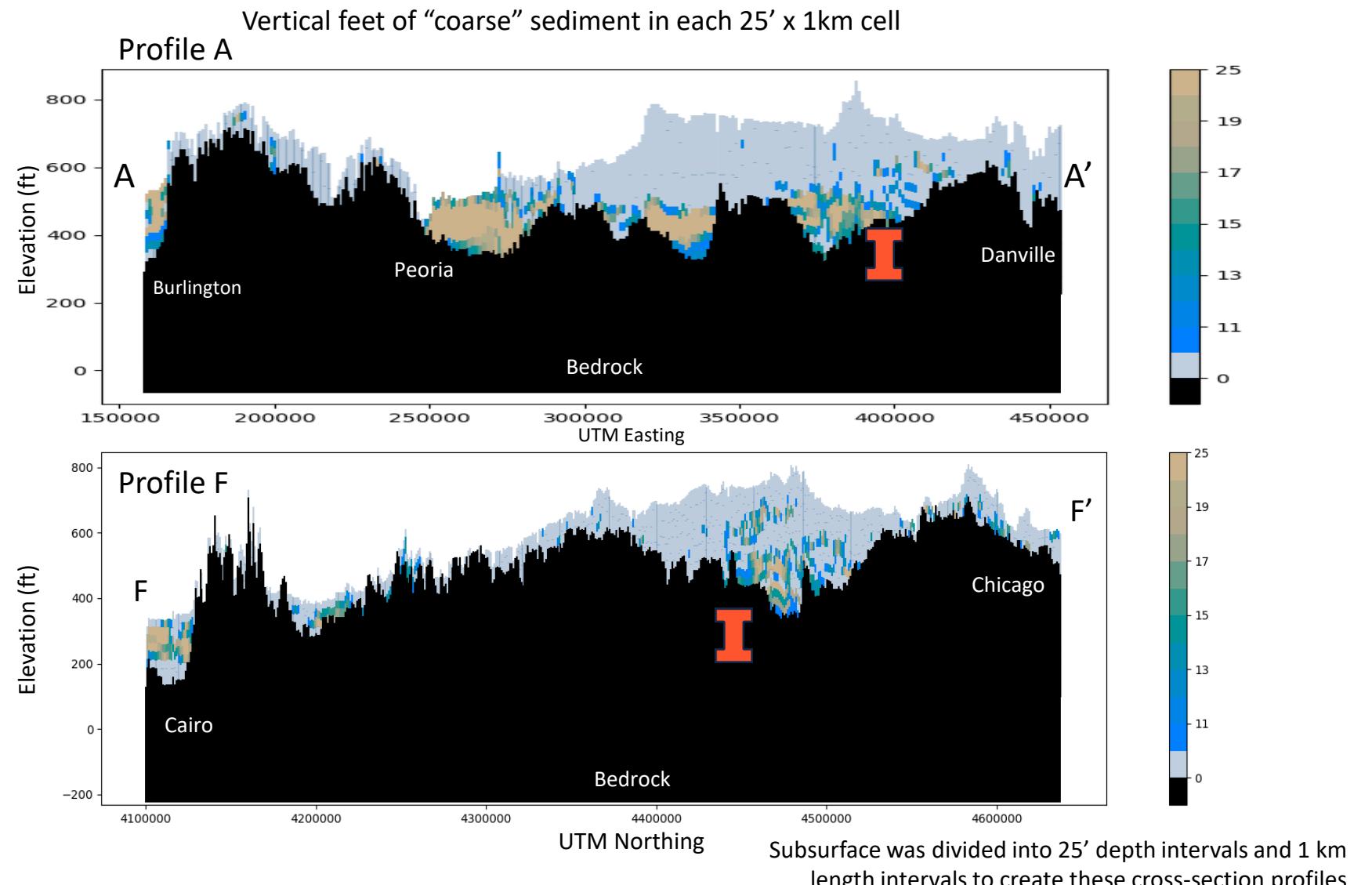
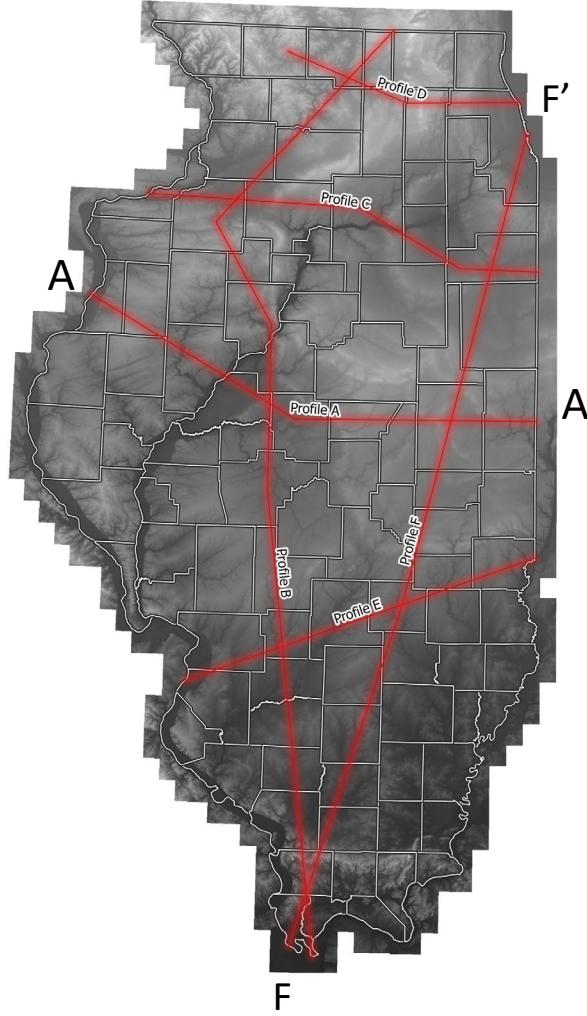
PRAIRIE RESEARCH INSTITUTE

## Surface Elevation

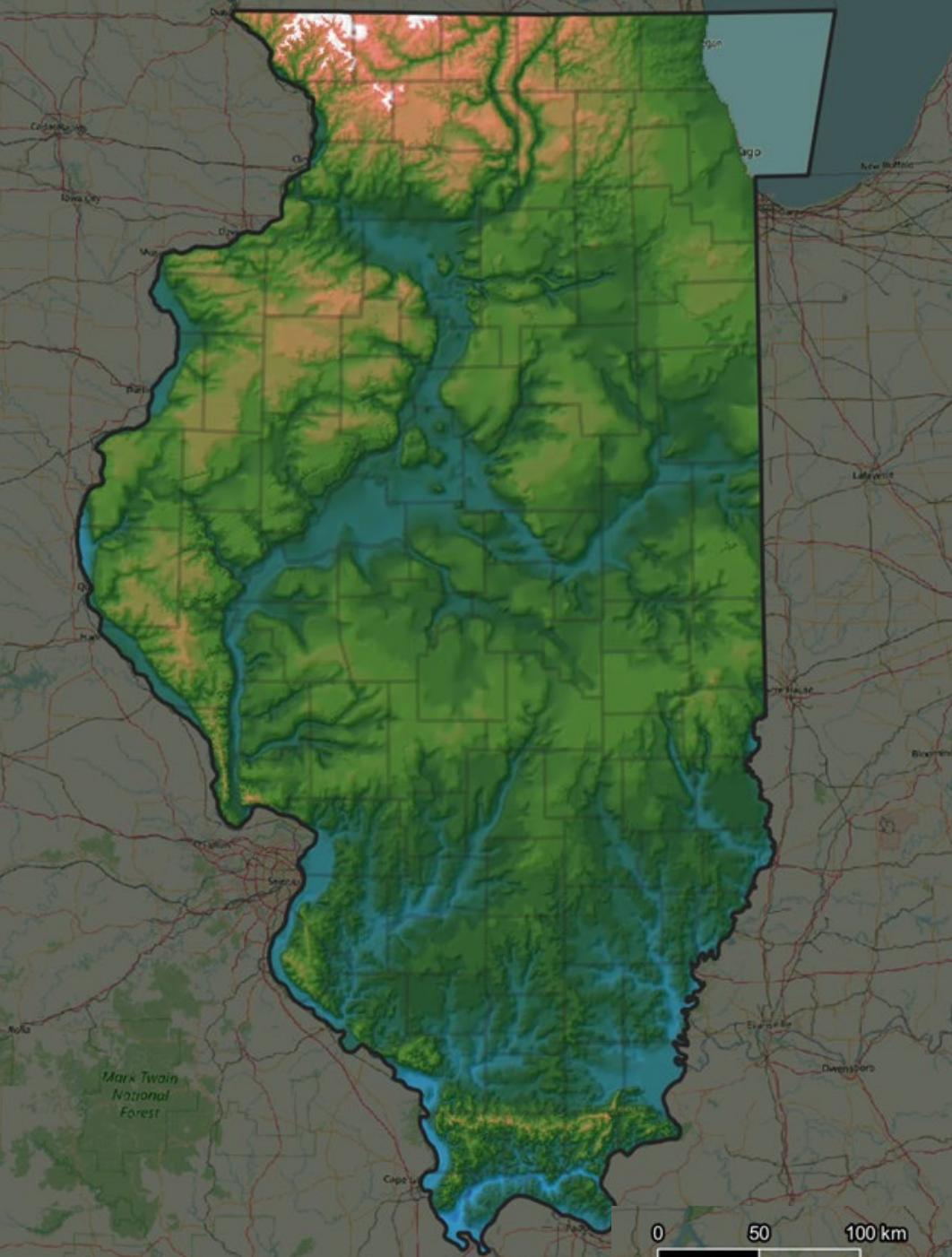


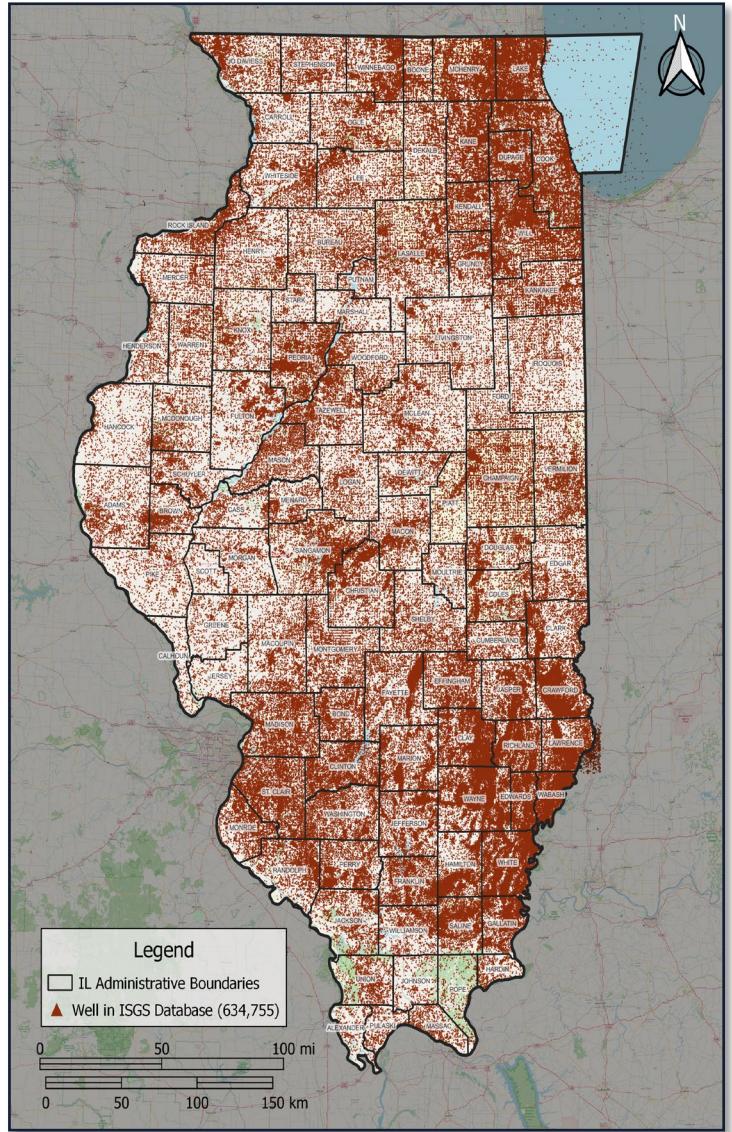
## Bedrock Elevation





# Why bedrock surface?





- More than ~635,000 unique wells
- Over 3 million individual well records

Well  
12001001090  
(6+ records/intervals)

Well  
120010011000  
(7 records/intervals)

Well  
120010011100  
(7 records/intervals)

Well  
120010011300  
(11 records/intervals)

API_NUMBE	TABLE_NAM	FORMATION	THICKNESS	TOP	BOTTOM
120010010900	FORMATION_T	Devonian		660	
120010010900	FORMATION_T	Galena		860	
120010010900	FORMATION_T	Keokuk		260	
120010010900	FORMATION_T	Kinderhook		480	
120010010900	FORMATION_T	Maquoketa		673	
120010010900	FORMATION_T	Silurian		662	
120010011000	WFORMATION	coal & shale,bl	1	56	57
120010011000	WFORMATION	s,whsh,med,w	13	25	38
120010011000	WFORMATION	sand,yellow,m	9	16	25
120010011000	WFORMATION	soil,dark brown	5	0	5
120010011000	WFORMATION	ss #22323	0	0	0
120010011000	WFORMATION	till,s,ox,some s	18	38	56
120010011000	WFORMATION	till,yl to bf belo	11	5	16
120010011100	WFORMATION	loess,gry,cly,ho	5	30	35
120010011100	WFORMATION	s.s. #22324	0	0	0
120010011100	WFORMATION	soil on till,yl,sa	5	35	40
120010011100	WFORMATION	soil,dark brown	5	0	5
120010011100	WFORMATION	till,bright oran	10	40	50
120010011100	WFORMATION	till,yl & bf/gry,	25	5	30
120010011200	WFORMATION	dirt & clay	36	0	36
120010011200	WFORMATION	lime	100	380	480
120010011200	WFORMATION	shale	340	40	380
120010011200	WFORMATION	shell rock w/ a	4	36	40
120010011300	WFORMATION	clay, blue	30	30	60
120010011300	WFORMATION	clay, yellow	30	0	30
120010011300	WFORMATION	rock, blue	25	150	175
120010011300	WFORMATION	rock, blue	40	60	100
120010011300	WFORMATION	rock, blue turn	25	230	255
120010011300	WFORMATION	rock, blue with	30	175	205
120010011300	WFORMATION	rock, blue, fine	25	205	230
120010011300	WFORMATION	soapstone, blu	50	100	150
120010011301	WFORMATION	"blue clay"	30	30	60

# Data Cleaning

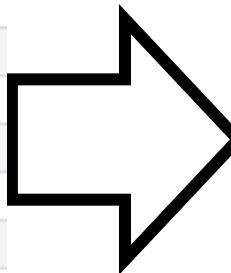
## Cleaning Steps:

- Remove data with missing depths
- Remove data with no locations
- Remove data with no description
- Remove data where top is below bottom
- Remove data from outside state

API_NUMBE	TABLE_NAM	FORMATION	THICKNESS	TOP	BOTTOM
120010010900	FORMATION_T	Devonian		660	
120010010900	FORMATION_T	Galena		860	
120010010900	FORMATION_T	Keokuk		260	
120010010900	FORMATION_T	Kinderhook		480	
120010010900	FORMATION_T	Maquoketa		673	
120010010900	FORMATION_T	Silurian		662	
120010011000	WFORMATION	coal & shale,bl	1	56	57
120010011000	WFORMATION	s,whsh,med,w	13	25	38
120010011000	WFORMATION	sand,yellow,m	9	16	25
120010011000	WFORMATION	soil,dark brown	5	0	5
120010011000	WFORMATION	ss #22323	0	0	0
120010011000	WFORMATION	till,s,ox,some s	18	38	56
120010011000	WFORMATION	till,yl to bf belo	11	5	16
120010011100	WFORMATION	loess,gry,cly,hc	5	30	35
120010011100	WFORMATION	s.s. #22324	0	0	0
120010011100	WFORMATION	soil on till,yl,sa	5	35	40
120010011100	WFORMATION	soil,dark brown	5	0	5
120010011100	WFORMATION	till,bright orang	10	40	50
120010011100	WFORMATION	till,yl & bf/gry,	25	5	30
120010011200	WFORMATION	dirt & clay	36	0	36
120010011200	WFORMATION	lime	100	380	480
120010011200	WFORMATION	shale	340	40	380
120010011200	WFORMATION	shell rock w/ a	4	36	40
120010011300	WFORMATION	clay, blue	30	30	60
120010011300	WFORMATION	clay, yellow	30	0	30
120010011300	WFORMATION	rock, blue	25	150	175
120010011300	WFORMATION	rock, blue	40	60	100
120010011300	WFORMATION	rock, blue turn	25	230	255
120010011300	WFORMATION	rock, blue with	30	175	205
120010011300	WFORMATION	rock, blue, fine	25	205	230
120010011300	WFORMATION	soapstone, blu	50	100	150
120010011301	WFORMATION	"blue clay"	30	30	60

# Data Classification

API_NUMBE	FORMATION
120010004200	Devonian
120010004200	Trenton
120010004300	Silurian
120010004500	"gray clay"
120010004500	"gray clay"
120010004500	"yellow clay"
120010004500	SS #26463
120010004500	chert white,limestone,dolomitic
120010004500	chert,white
120010004500	chert white

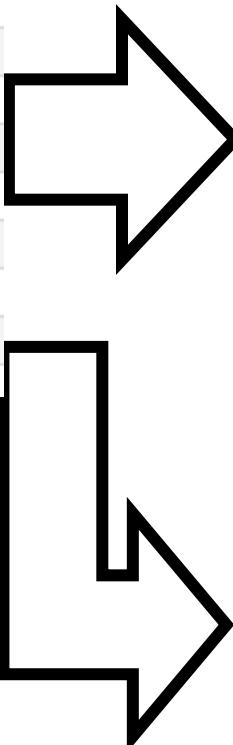


## Exact Definitions

ID	SearchTerm	InterpUpdate
0	sand, fine gr	SAND
1	sand, fine-large gr	SAND
2	sand, some large gr	SAND
3	toop soil	SOIL
4	grav in muck	GRAVEL
5	sand & clay	SAND AND CLAY
6	sandrock	BEDROCK
7	yellow fine loose limestone	BEDROCK
8	limestone, broken grav	DIRT AND BEDROCK
9	limestone, broken clay	DIRT AND BEDROCK
10	slot sand	SAND
11	clay & gravel loam	CLAY AND GRAVEL

# Data Classification

API_NUMBE	FORMATION
120010004200	Devonian
120010004200	Trenton
120010004300	Silurian
120010004500	"gray clay"
120010004500	"gray clay"
120010004500	"yellow clay"
120010004500	SS #26463
120010004500	chert white,limestone,dolomitic
120010004500	chert,white
120010004500	chert white



## Exact Definitions

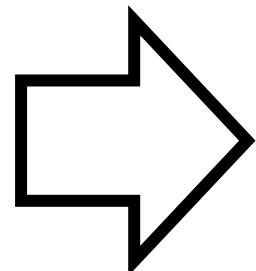
ID	SearchTerm	InterpUpdate
0	sand, fine gr	SAND
1	sand, fine-large gr	SAND
2	sand, some large gr	SAND
3	toop soil	SOIL
4	grav in muck	GRAVEL
5	sand & clay	SAND AND CLAY
6	sandrock	BEDROCK
7	yellow fine loose limestone	BEDROCK
8	limestone, broken grav	DIRT AND BEDROCK
9	limestone, broken clay	DIRT AND BEDROCK
10	slot sand	SAND
11	clay & gravel loam	CLAY AND GRAVEL

## "Starts with..." Definitions (i.e., wildcard search: [searchterm]\*)

ID	StartTerm	InterpUpdate
1	s,med,brn;	SAND
2	s,med,calc,lgt gry brn,	SAND
3	s,crs,brn,	SAND
4	cly st,calc,	SILT
5	sty cl,lea,gry,	CLAY
6	broken lime (clay seams)	BEDROCK AND OTHER
7	gravel,sandy,some water	GRAVEL WITH SAND
9	st,lgt gry	SILT
10	st,slgty calc,	SILT
11	st,sy,cly,grngry,	SILT AND SAND MIX
12	st,sy,brn-grn	SILT AND SAND MIX

# Data Classification

API_NUMBE	FORMATION
120010004200	Devonian
120010004200	Trenton
120010004300	Silurian
120010004500	"gray clay"
120010004500	"gray clay"
120010004500	"yellow clay"
120010004500	SS #26463
120010004500	chert white,limestone,dolomitic
120010004500	chert,white
120010004500	chert white



## Known Bedrock Depth (550')

- No bedrock valley in Illinois is deeper than 550'
- All well records with a top depth > 550' = bedrock

# Bedrock Picking

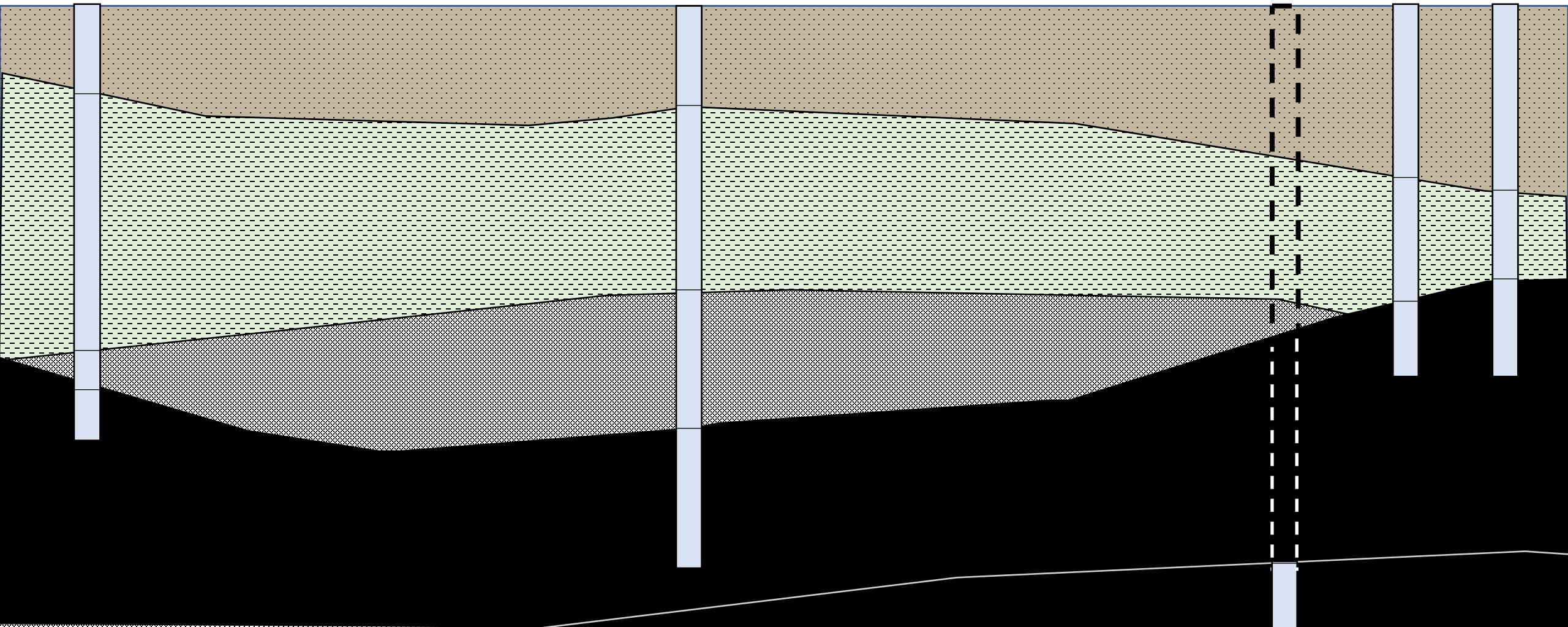
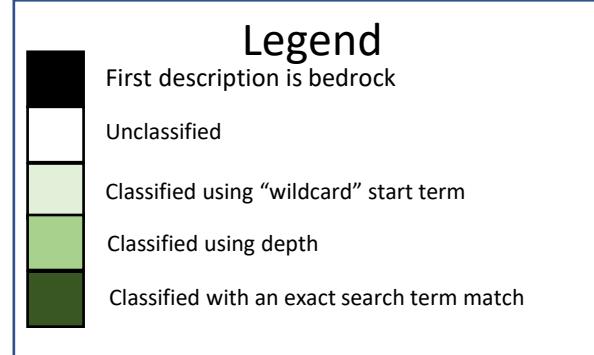
Well A

Well B

Well C

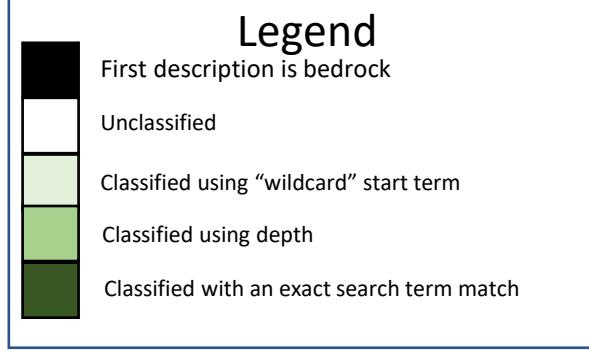
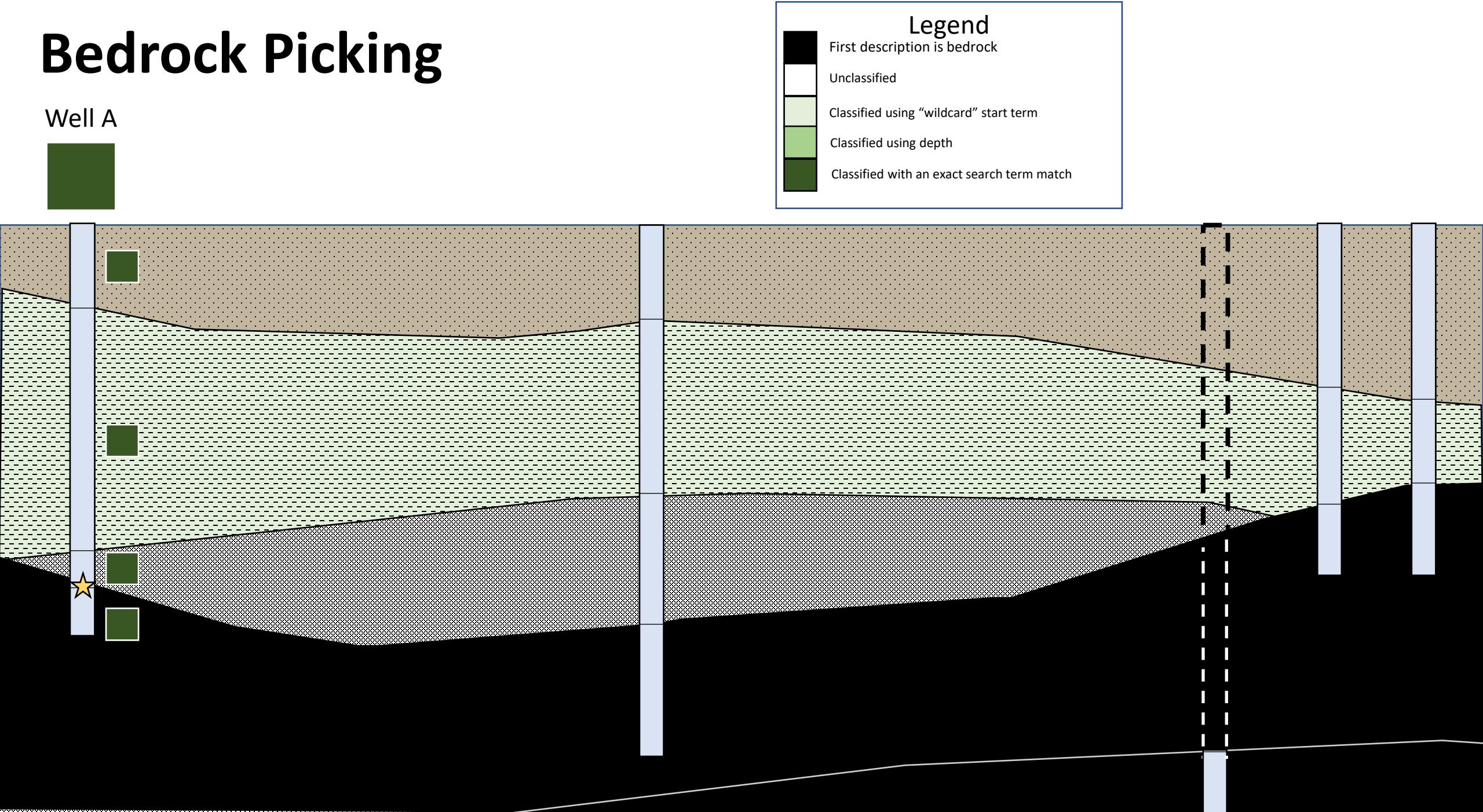
Well D

Well E



# Bedrock Picking

Well A



# Bedrock Picking

Well A

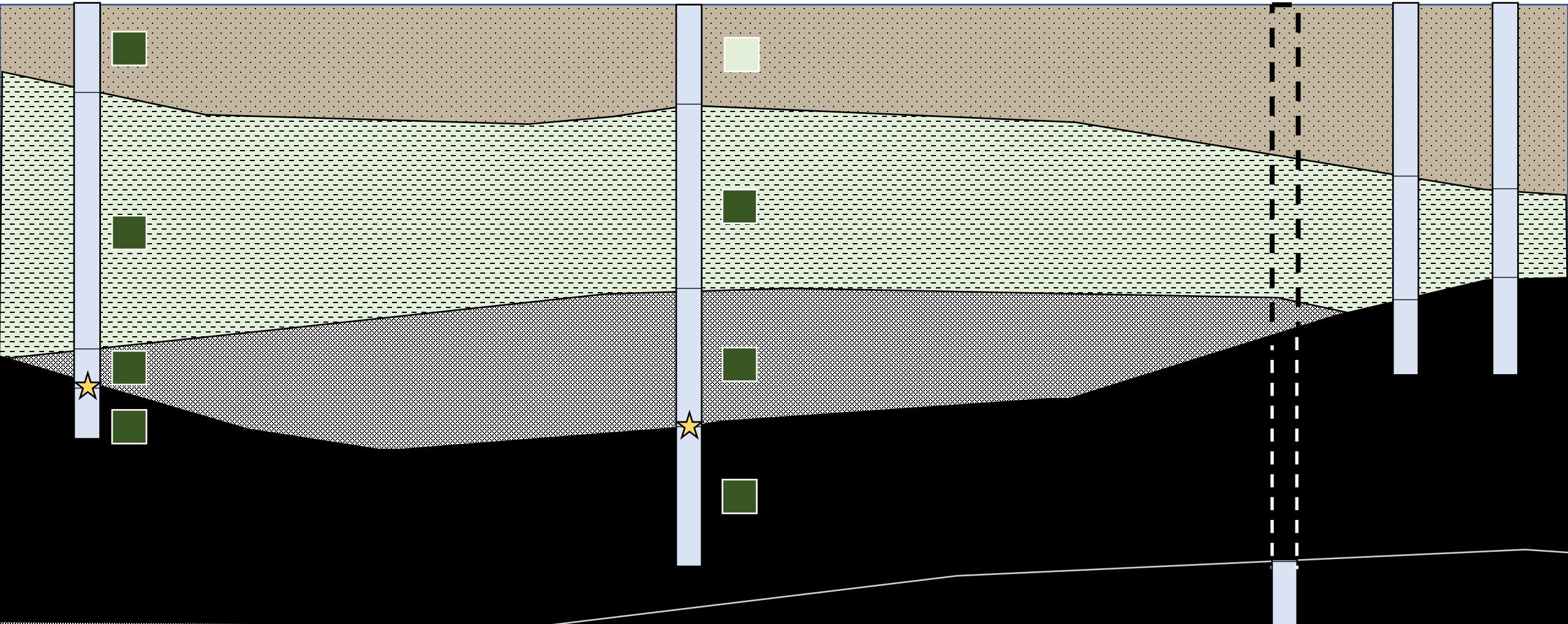


Well B



## Legend

- First description is bedrock
- Unclassified
- Classified using "wildcard" start term
- Classified using depth
- Classified with an exact search term match

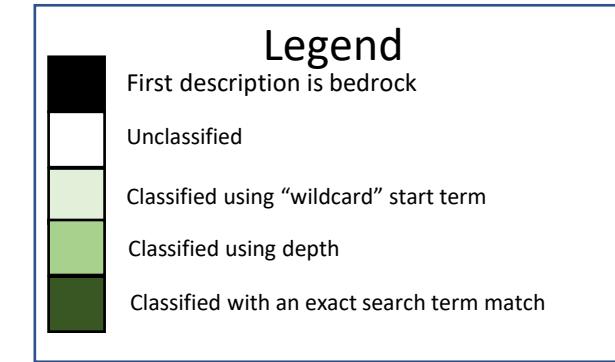


# Bedrock Picking

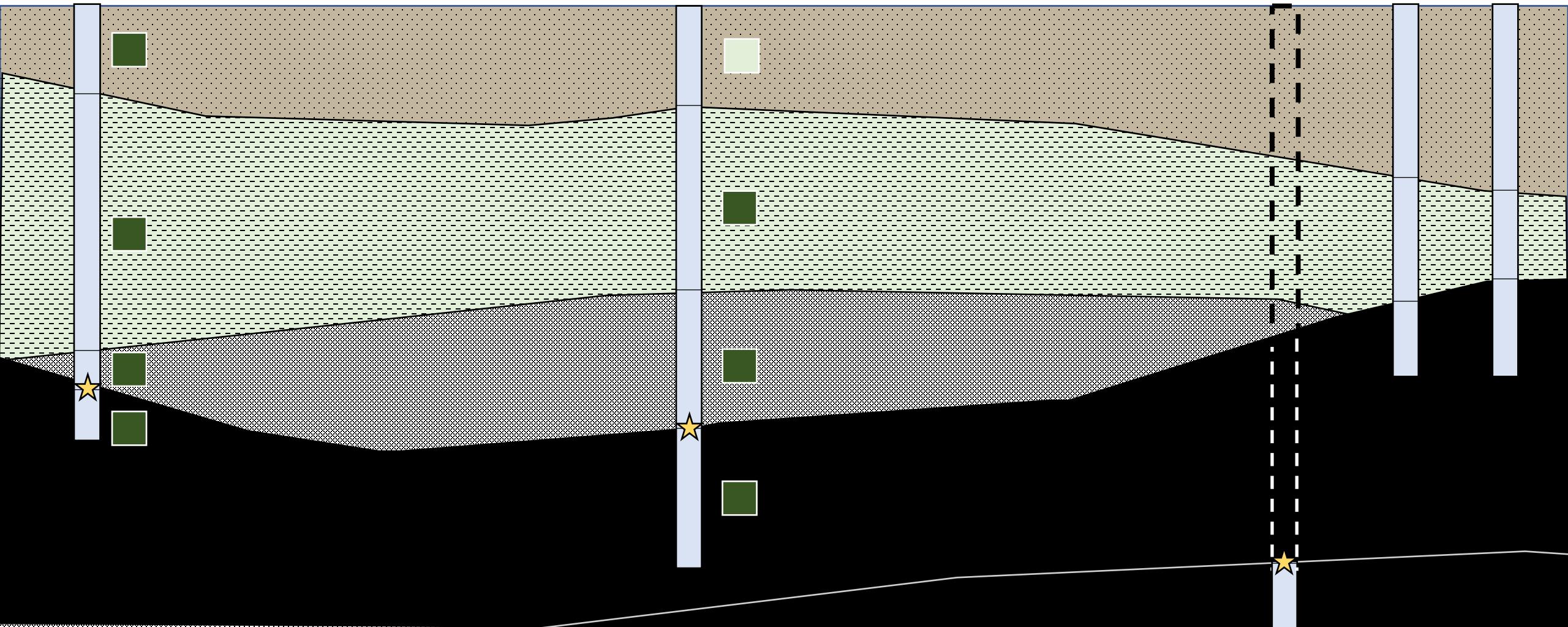
Well A



Well B



Well C



# Bedrock Picking

Well A



Well B



## Legend

First description is bedrock

Unclassified

Classified using "wildcard" start term

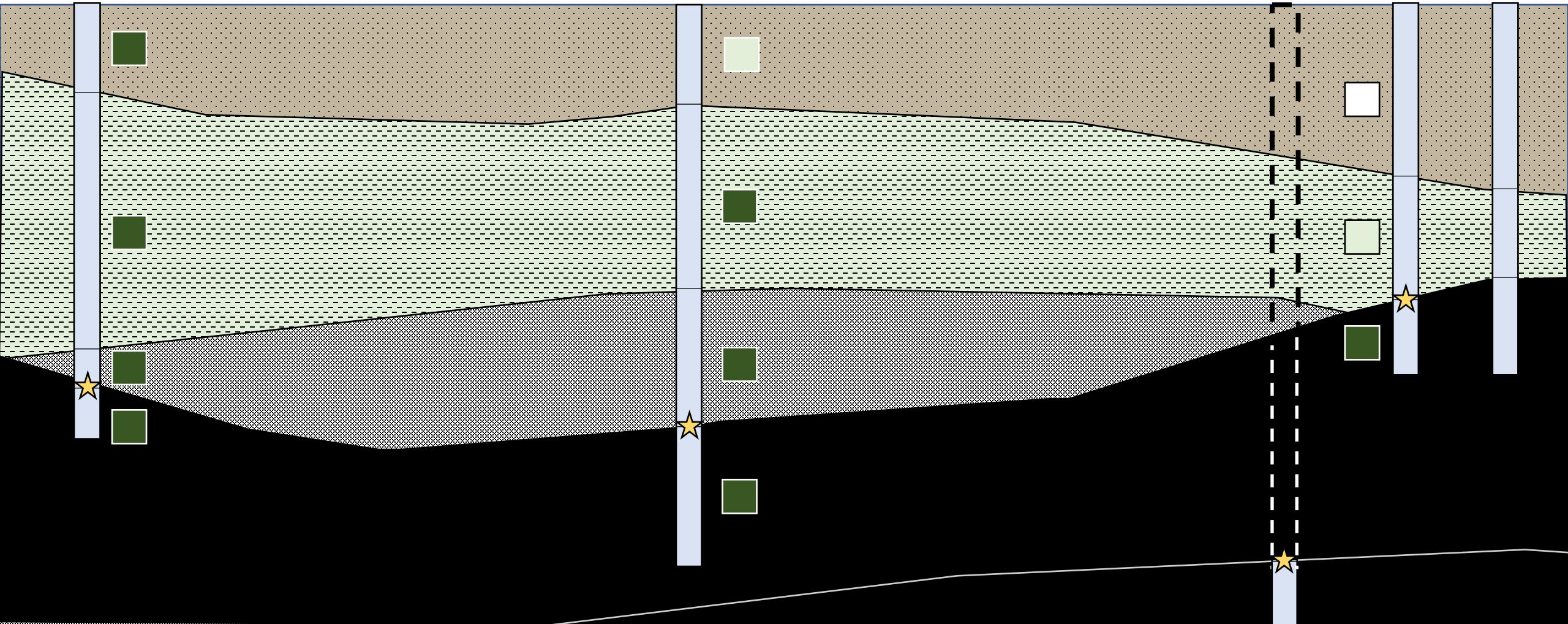
Classified using depth

Classified with an exact search term match

Well C



Well D



# Bedrock Picking

Well A



Well B



## Legend

First description is bedrock

Unclassified

Classified using "wildcard" start term

Classified using depth

Classified with an exact search term match

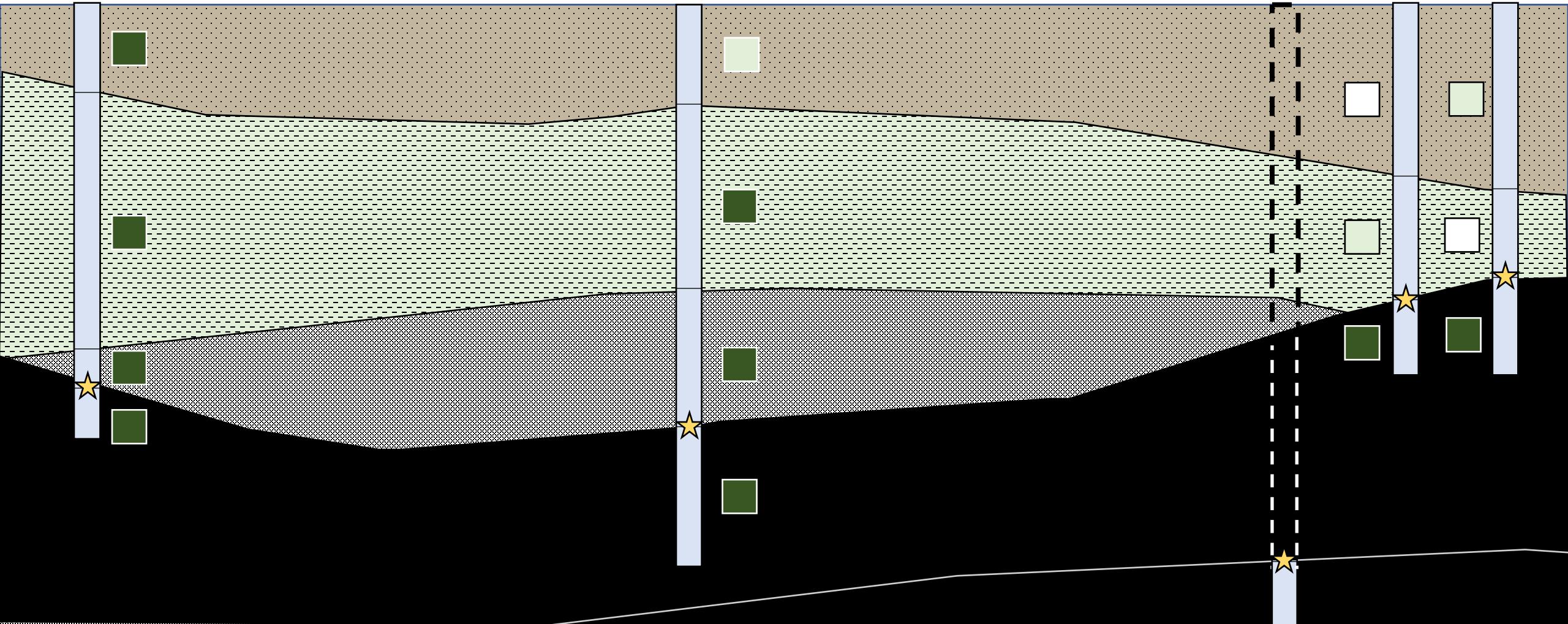
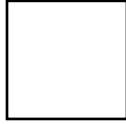
Well C



Well D



Well E



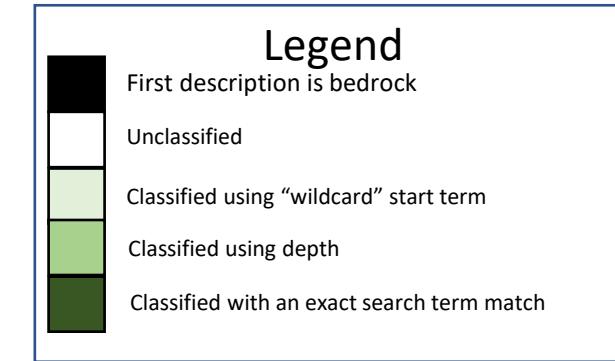
# Bedrock Picking

Well A

2

Well B

1



Well C

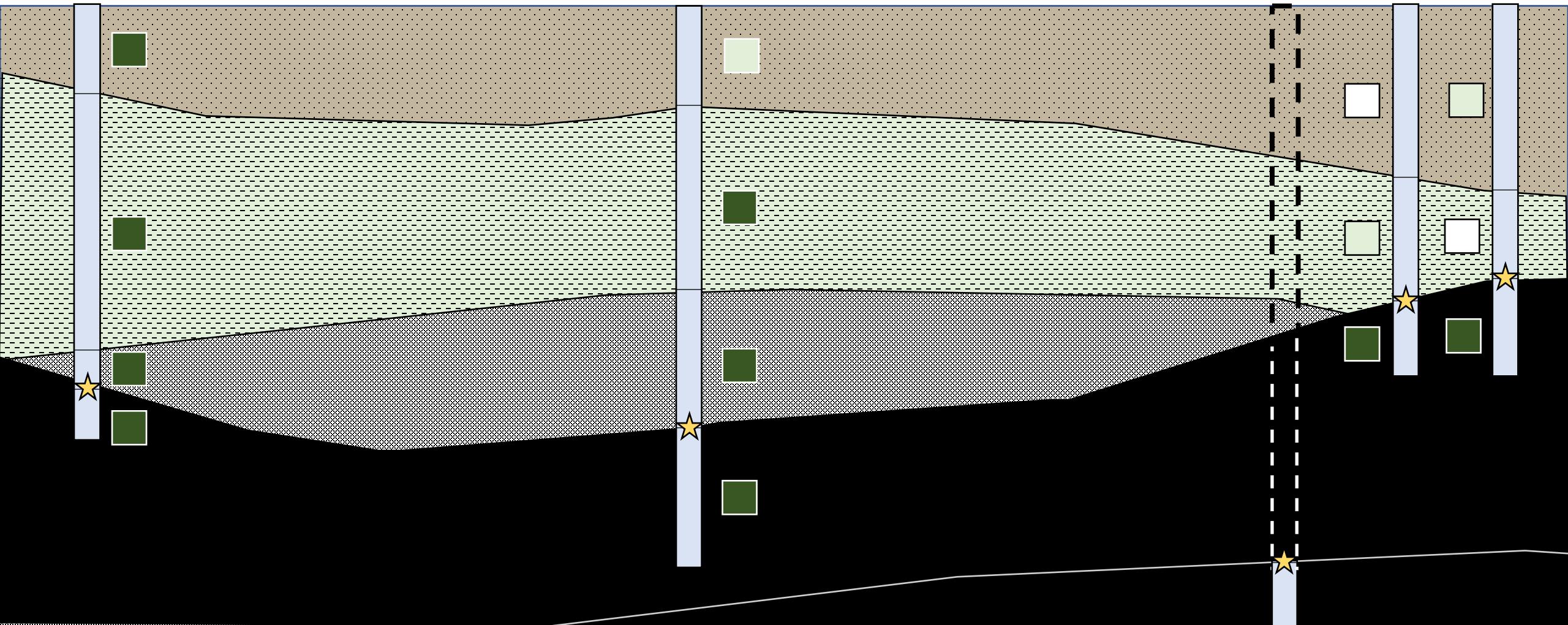
0

Well D

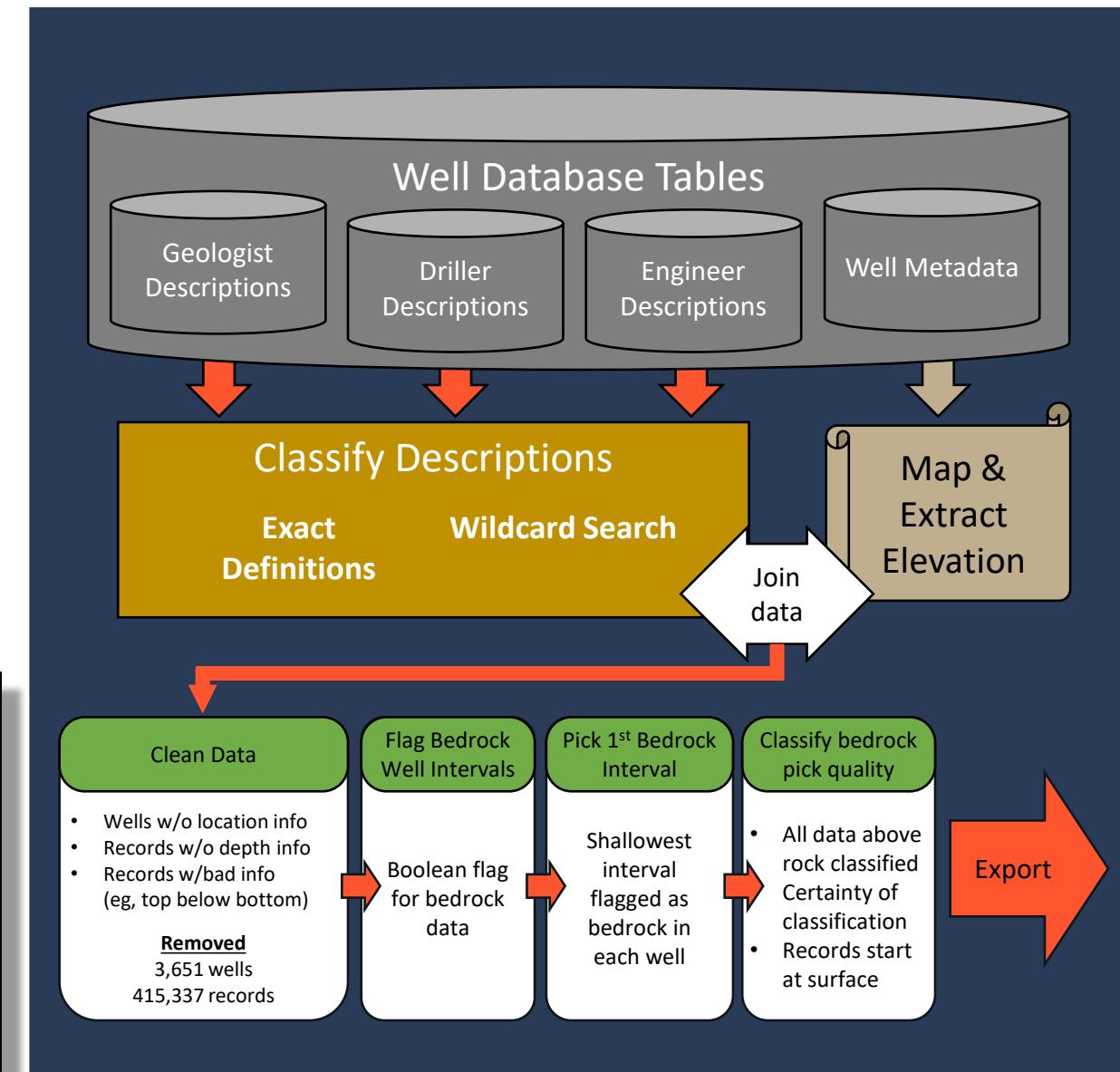
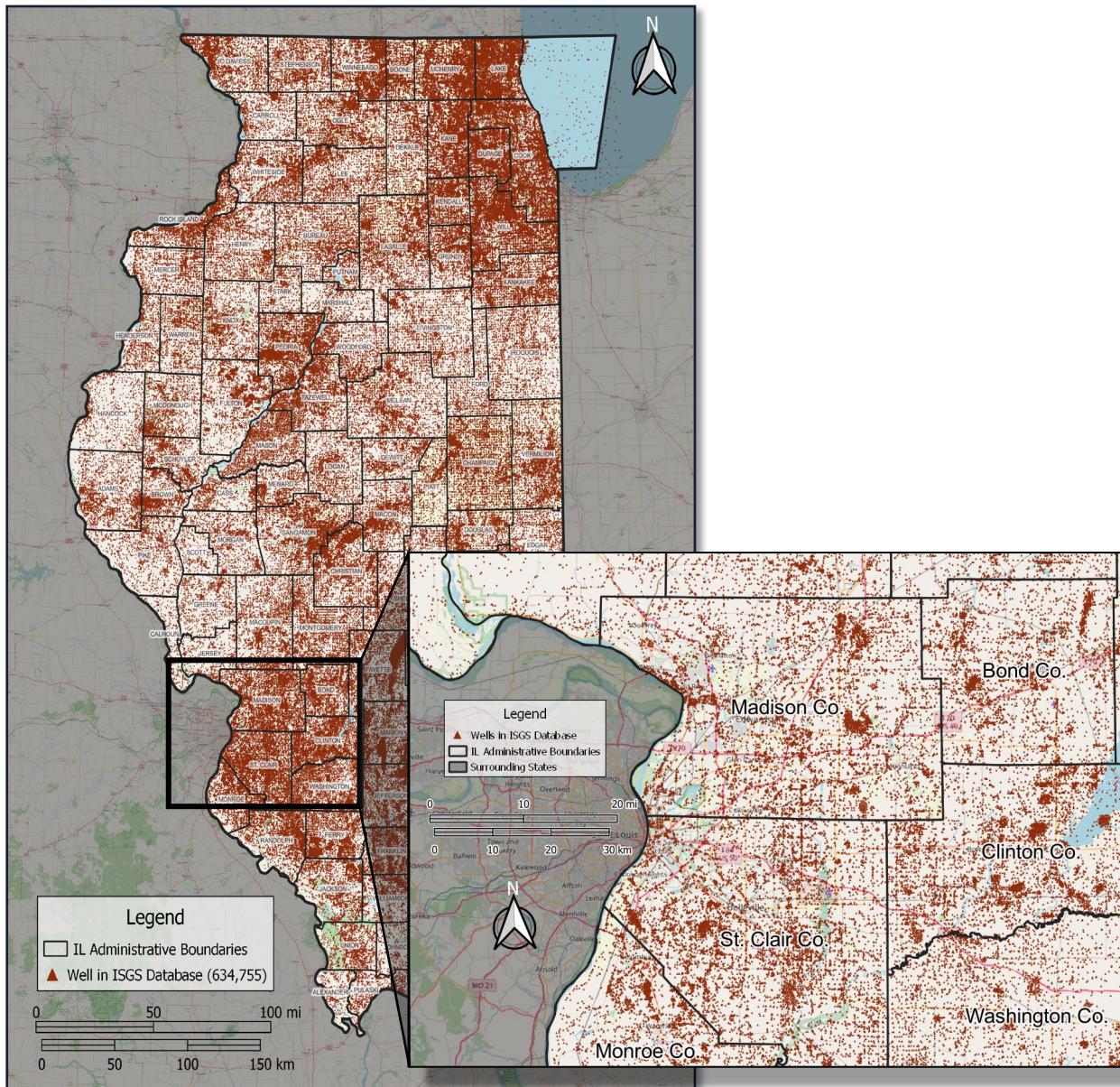
-1

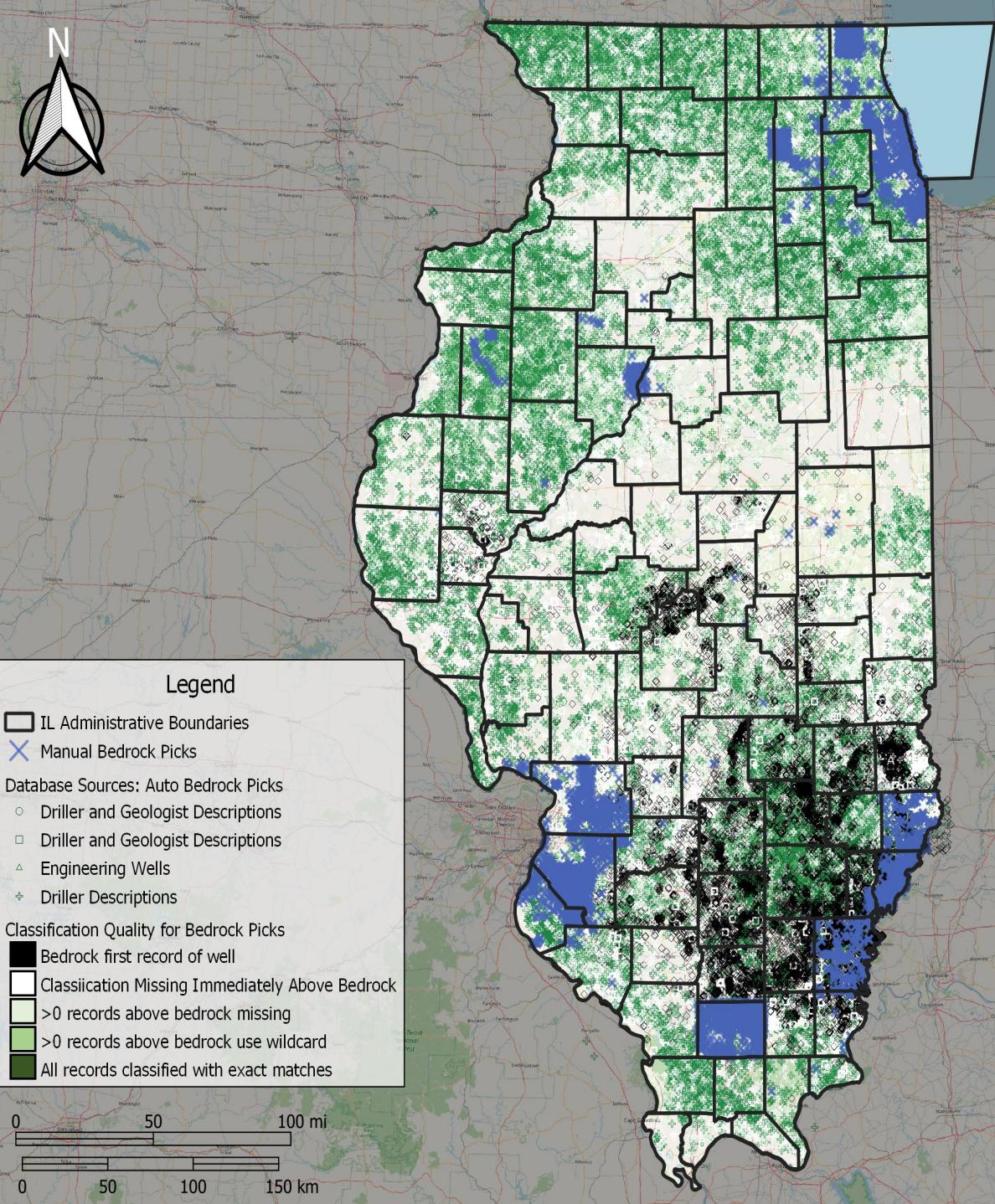
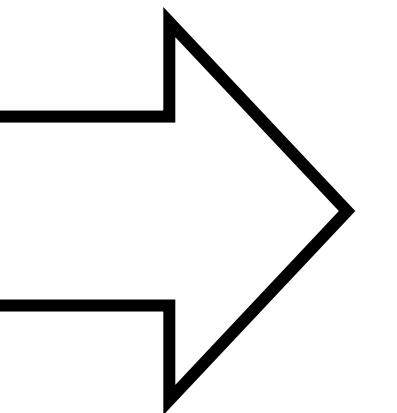
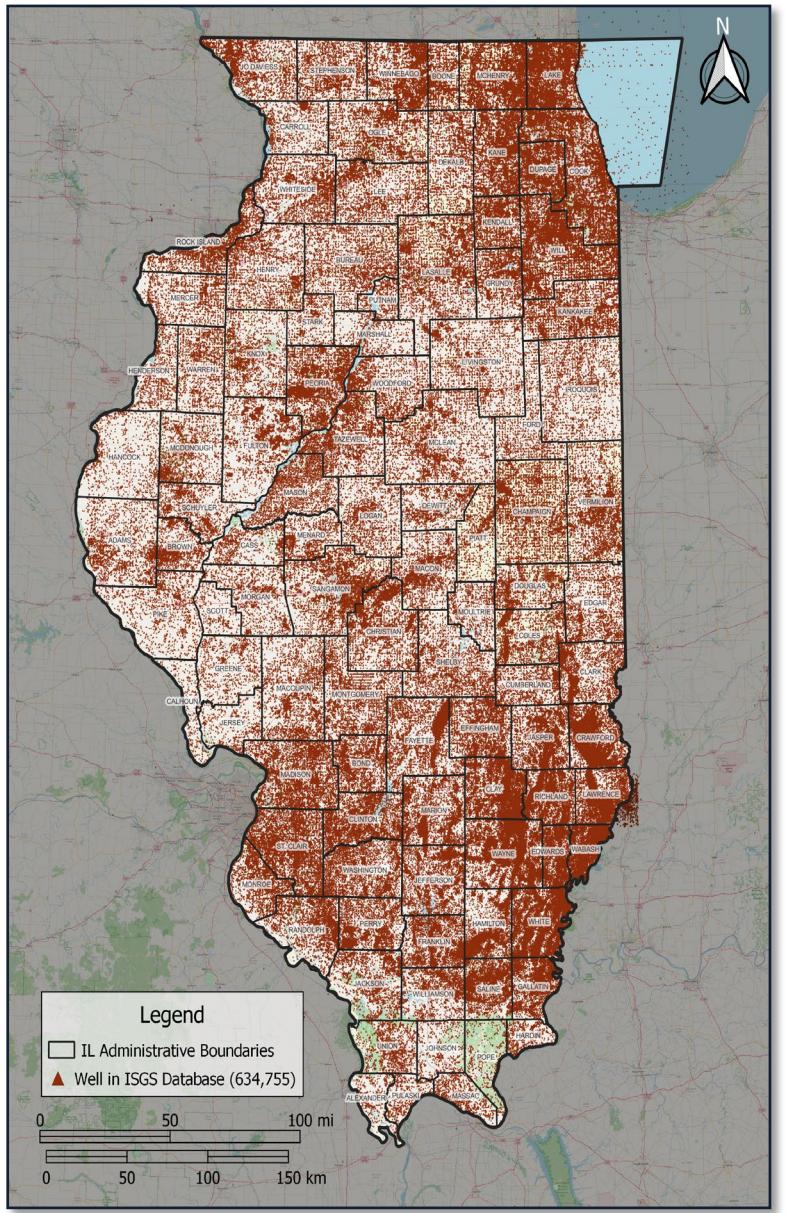
Well E

-2

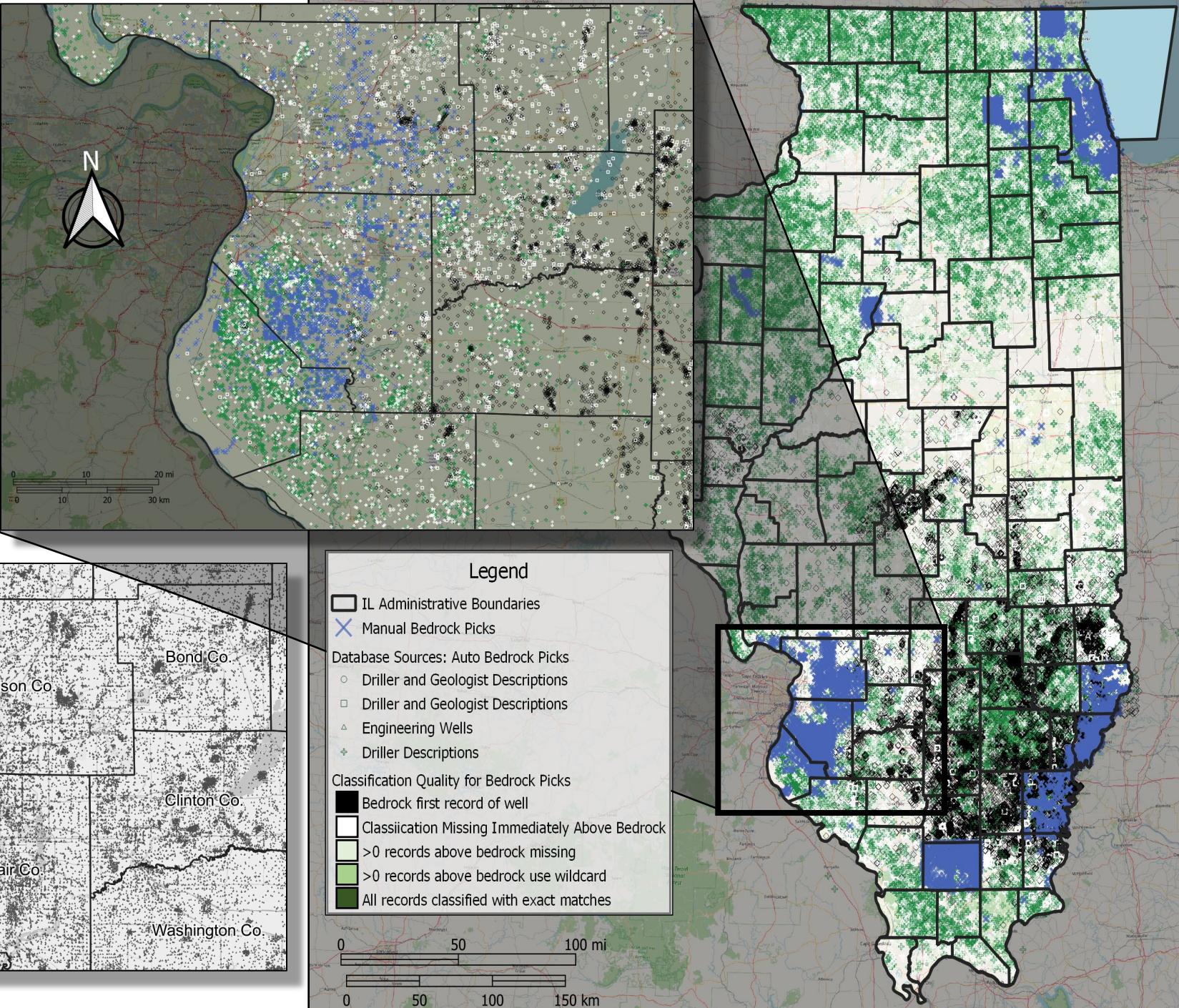
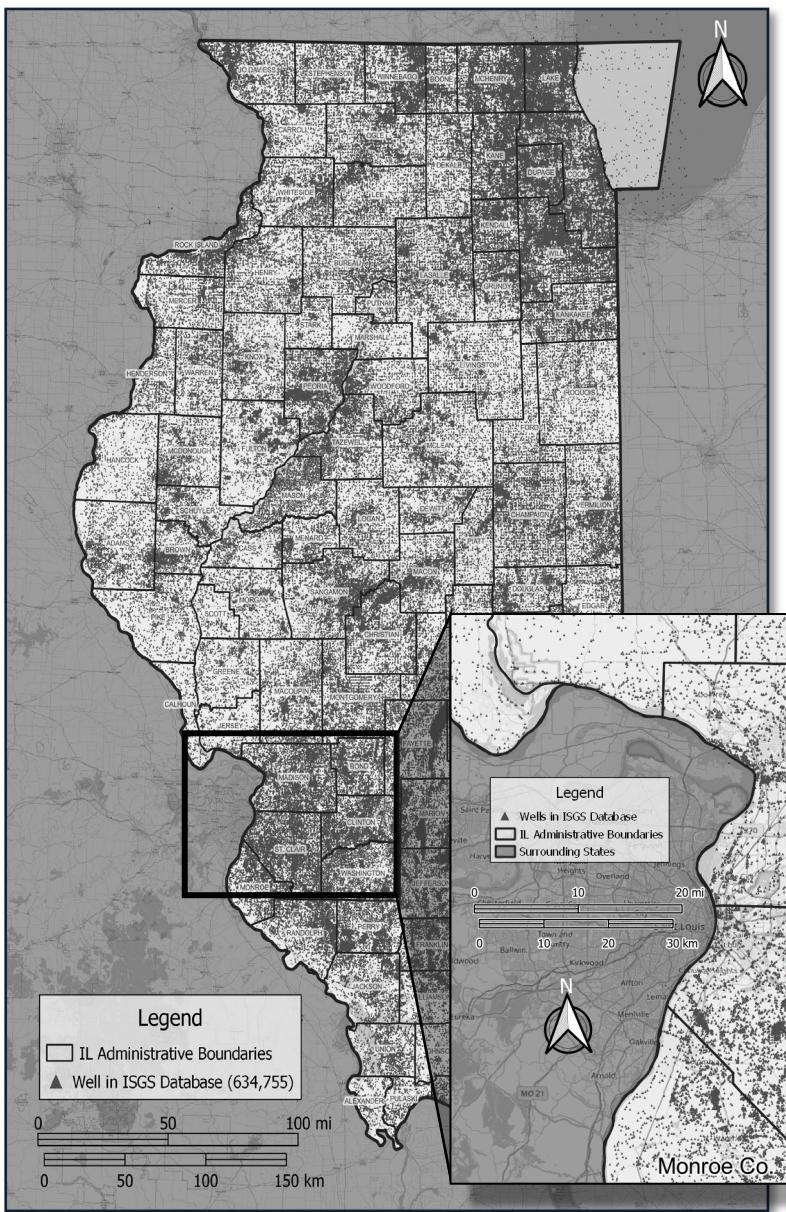


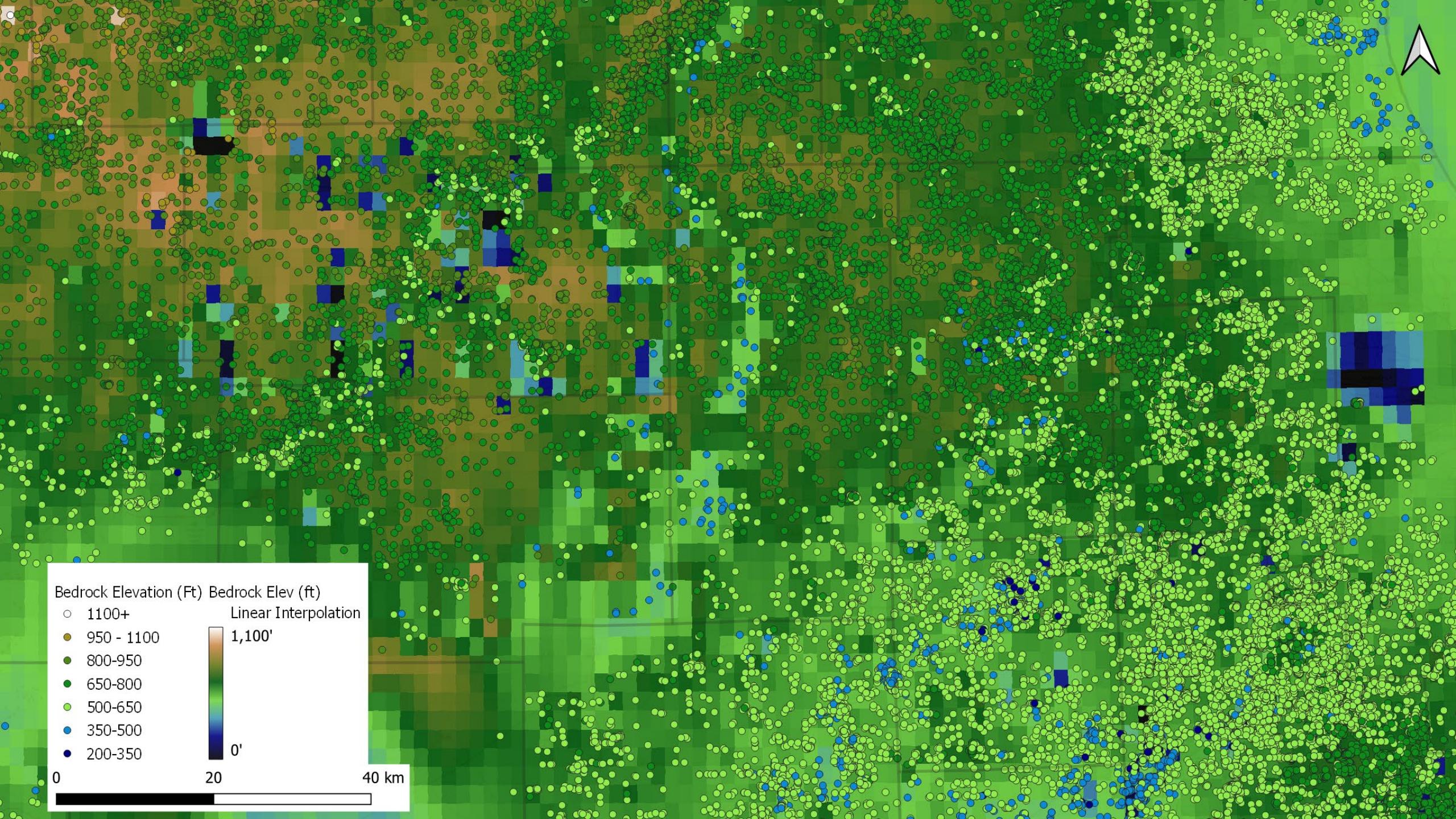
# Unsupervised classification of well logs





# Unsupervised classification of well logs





# Statistics Consulting Needs:

## From the output

- How to identify “outliers”?
  - Variable surface texture
  - Can the weightings/pick classifications be used to help with this?
  - The source tables may be of help with weighting too



## In the workflow:

- Other methods for identifying poor descriptions?
- Other methods for classification or weighting of well records?
- Other methods for classification or weighting of bedrock?

Other suggestions?