EDA Ashfield Sediment Data 2019-2022

This document describes exploratory and some detailed data analysis



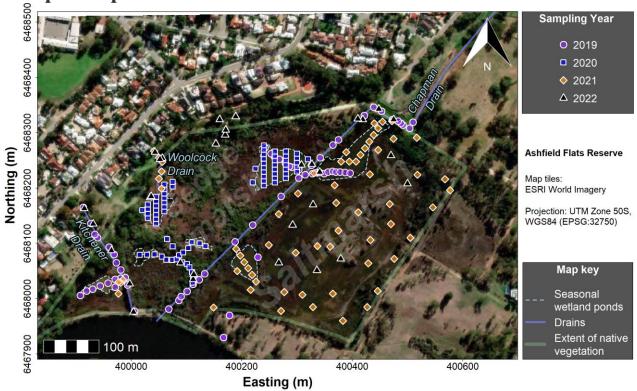


Figure 1: Map of sediment sample locations at Ashfield Flats for all sampling years 2019-2022.

Samples were taken each year from 2019 to 2022 (Figure 1). Different sampling designs were implemented each year depending on the perceived gaps in understanding of the site and the learning objectives desired for the students involved. Apart from three soil/sediment depth profiles in 2019, all samples were subaqueous or subaerial surface sediments (0-10cm) from stormwater drains, wetland ponds (with or without water), saltmarsh, or seasonally flooded woodland.

Data summaries

pH, EC, Al-Cu

Table 1: F	First block o	f elements									
Stat	рН	EC	Al	As	Ba	Ca	Cd	Ce	Co	Cr	Cu
mean	5.97e+00	9.81e+00	3.02e+04	8.44e+00	5.95e+01	3.68e+03	2.03e-01	9.17e+01	1.46e+01	4.90e+01	1.25e+02
sd	8.74e-01	2.11e+01	1.34e+04	8.48e+00	3.02e+01	4.64e+03	2.36e-01	6.17e+01	1.27e+01	2.09e+01	1.30e+02
rsd	14.6%	215.1%	44.4%	100.5%	50.8%	126.1%	116.3%	67.3%	87%	42.7%	104%
0%	3.160	0.014	888.000	0.400	4.600	470.000	0.015	2.200	1.800	1.800	2.000
50%	6.00	6.59	31800.00	6.60	57.00	2320.00	0.12	82.40	12.00	54.30	79.00
100%	8.36	280.00	64400.00	62.00	191.00	48500.00	1.60	271.00	110.00	84.00	1010.00
NAs	32	33	1	13	1	1	98	1	1	1	8

Fe-Ni

Tahla	2.	Second	hlock	$\cap f$	elements
Iable	۷.	SECULIA	DIOCK	UΙ	elellello

Stat	Fe	Ga	Gd	K	La	Li	Mg	Mn	Мо	Na	Nd	Ni
mean	41000.00	17.40	6.53	2490.00	45.20	27.90	3850.00	153.00	2.77	11600.00	35.10	19.40

Tahla	2.	Secon	Ы	hlock	of a	elements

Stat	Fe	Ga	Gd	K	La	Li	Mg	Mn	Мо	Na	Nd	Ni
sd	20500.00	7.53	3.63	1340.00	28.60	15.50	2330.00	346.00	2.10	11000.00	21.70	8.31
rsd	50%	43.3%	55.6%	53.8%	63.3%	55.6%	60.5%	226.1%	75.8%	94.8%	61.8%	42.8%
0%	1560.0	-2.0	0.2	59.0	1.0	0.5	106.0	6.0	0.1	61.0	8.0	2.0
50%	38600.00	17.60	6.28	2720.00	41.40	30.00	3720.00	81.00	2.00	10500.00	34.00	19.70
100%	135000.0	40.7	18.0	5260.0	126.0	80.0	20700.0	3500.0	13.0	95800.0	98.0	50.0
NAs	1	170	7	1	1	1	1	1	4	1	1	8

P-Zn

Table 3: Third block of element

Stat	Р	Pb	Rb	S	Sc	Sr	Th	Ti	V	Υ	Zn
mean	650.00	60.70	37.60	5550.00	6.96	60.20	13.40	211.00	63.30	21.20	550.00
sd	724.00	94.60	24.80	9150.00	3.49	50.20	6.51	81.70	24.50	15.00	1140.00
rsd	111.4%	155.8%	66%	164.9%	50.1%	83.4%	48.6%	38.7%	38.7%	70.8%	207.3%
0%	26.0	3.3	0.6	225.0	0.1	2.3	0.4	47.2	2.3	0.4	11.2
50%	454.0	40.0	41.2	2970.0	8.0	52.8	14.7	219.0	71.0	19.4	246.0
100%	7020.0	831.0	88.0	76600.0	12.4	512.0	26.9	395.0	118.0	67.0	7560.0
NAs	1	1	174	1	170	1	3	167	1	1	5

The elements of primary interest in this study are the rare-earth (*REE* or lanthanide) elements La, Ce, Nd, and Gd; Y (often considered together with the REE); major elements considered to be useful proxies for sediment parameters expected to affect geochemical reactions of REE: Al, Ca, Fe, P and S. Aluminium (Al) is a proxy for clay minerals (although pXRD data show that other aluminosilicates are present, such as feldspars and micas, these are resistant and less likely than the phyllosilicate clays to report Al to an *aqua regia* digest). Calcium (Ca) is a proxy for carbonate minerals (since silicate-bound Ca would also be resistant to dissolution in *aqua regia*). Iron (Fe) concentrations are a proxy for iron oxide (and possibly iron sulfide) minerals. Phosphorus (P) is included since secondary REE phosphates such as rhabdophane may be a REE sink. Sulfur (S) most likely represents sulfate and sulfide minerals, would be expected to accumulate in wet, reducing environments, and is strongly linked to acid sulfate soils.

Sediment pH and EC have numerous but not excessive missing observations and are included due to their substantial effects on sediment geochemical processes.

The trace elements As, Cu, Pb, and Zn are included as common inorganic contaminants. In addition, Pb may be immobilised in environments receiving acid sulfate soil drainage due to the insolubility of PbSO₄. Thorium (Th) may be depleted in oxidised acid sulfate soils.

Note that several elements have too many missing observations to be useful: Cd, Ga, Rb, Sc, and Ti.

Drain sediment

Table 4: Basic statistics for Drain Sediment

Stat	рΗ	EC	Al	Ca	Fe	Р	S	La	Ce	Nd	Gd	Υ
mean	6.35	8.55	18000	3700	33300	589	6390	20.3	39.6	15.9	3.33	9.22
sd	1.13	41.1	13000	4000	22500	601	7840	17.7	39.3	15	2.63	9.57
0%	3.49	0.014	888	470	1560	26	225	1	2.2	8.0	0.2	0.4

Table 1.	Rasin	etatietice	for	Drain	Sediment	
1 4011111111111111111111111111111111111	Dasil	SIGHSHUS	1()1	плаш	ocumen	

Stat	рΗ	EC	Al	Ca	Fe	Р	S	La	Ce	Nd	Gd	Υ
50%	6.64	0.985	14100	2200	32700	389	3910	13.4	24.8	10.6	2.22	4.8
100%	8.27	280	64400	21600	88000	2800	40900	87	183	70	12	45

	Lake sediment Table 5: Basic statistics for Lake Sediment													
Stat	рН	EC	Al	Ca	Fe	Р	S	La	Ce	Nd	Gd	Υ		
mean	6.090	8.940	36,000	4,550	41,400	600	7,530	60.6	124.0	45.0	8.23	29.1		
sd	0.808	6.200	11,500	5,430	16,700	428	12,200	30.7	65.9	22.3	3.76	16.4		
0%	3.160	0.174	2,860	961	8,220	81	380	3.0	3.0	2.0	1.00	2.0		
50%	6.170	8.200	36,900	3,770	38,300	502	3,650	65.0	129.0	47.0	8.60	29.6		
100%	8.360	32.900	59,800	48,500	135,000	2,390	76,600	126.0	271.0	98.0	18.00	67.0		

Saltmarsh

Table 6: Basic statistics for Saltmarsh

Stat	рН	EC	Al	Ca	Fe	Р	S	La	Ce	Nd	Gd	Υ
mean	5.670	15.30	30,900	3,120	48,600	961	3,860	47.2	94.7	38.4	6.88	21.7
sd	0.694	18.20	11,500	4,250	23,600	1,120	5,510	20.3	47.1	17.7	2.72	10.8
0%	4.350	0.12	6,070	584	11,900	85	281	6.2	6.0	5.9	2.00	3.2
50%	5.700	11.50	31,900	1,960	41,400	540	2,780	52.0	107.0	42.0	7.00	22.8
100%	7.240	135.00	53,800	31,400	130,000	7,020	42,300	93.4	218.0	95.7	16.80	53.4

Distribution tests

Table 7: Shapiro-Wilk statistics and p-values for untransformed (_orig) and transformed (_log, _pow) variables, power terms, and dip test of multimodality of log-transformed variable, from soil and sediment analysis at Ashfield Flats Reserve 2019-2022.

Variable	W_orig	p_orig	W_log_tr	p_log_tr	W_pow_tr	p_pow_tr	PowEst	PowRnd	D_diptest	p_diptest
рН	0.991	0.195	0.963	9.45e-06	0.994	0.545	1.47	1	0.0161	0.98
EC	0.301 1.	.07e-28	0.934	1.09e-08	0.964	1.52e-05	0.18	0.18	0.023	0.574
Al	0.974 0	0.00011	0.828	2.36e-16	0.975	0.000157	1.03	1	0.0175	0.889
As	0.7017.	.92e-21	0.986	0.0179	0.987	0.0255	0.0468	0	0.032	0.0706
Ва	0.967 9.	.23e-06	0.934	1.96e-09	0.994	0.323	0.563	0.5	0.0234	0.424
Ca	0.516	3.1e-26	0.982	0.00197	0.995	0.632	-0.206	-0.33	0.0115	0.996
Ce	0.956 4.	.24e-07	0.907	1.08e-11	0.982	0.00197	0.501	0.5	0.018	0.858
Co	0.659 1.	.31e-22	0.98	0.000842	0.98	0.000863	-0.00921	0	0.0203	0.687
Cr	0.945 2.	.61e-08	0.766	4.73e-19	0.954	2.49e-07	1.2	1	0.0165	0.937
Cu	0.774 1.	.91e-18	0.978	0.000521	0.982	0.00228	0.103	0	0.0144	0.991
Fe	0.886 3.	.83e-13	0.827	2.07e-16	0.941	9.69e-09	0.579	0.5	0.0132	0.993
Gd	0.972 6.	.15e-05	0.928	7.51e-10	0.988	0.0313	0.565	0.5	0.0252	0.314
K	0.965 5.	.28e-06	0.821	1.04e-16	0.962	1.9e-06	0.815	0.815	0.0232	0.442
La	0.962 1.	.98e-06	0.906	9.75e-12	0.987	0.0202	0.526	0.5	0.0171	0.913
Li	0.968 1.	.55e-05	0.814	5.16e-17	0.961	1.73e-06	0.8	0.8	0.0289	0.133
Mg	0.903 5.	.88e-12	0.874	7.02e-14	0.971	3.66e-05	0.558	0.5	0.0234	0.426
Mn	0.283 1.	.18e-30	0.913	3.23e-11	0.935	2.52e-09	-0.167	-0.167	0.0128	0.994

Table 7 (continued).

Variable	W_orig p	p_orig	W_log_tr	p_log_tr	W_pow_tr	p_pow_tr	PowEst	PowRnd	D_diptest	p_diptest
Мо	0.85 4.0	05e-15	0.966	8.38e-06	0.989	0.039	0.229	0.33	0.0485	0.000151
Na	0.788 3	.8e-18	0.879	1.55e-13	0.971	3.21e-05	0.366	0.33	0.0324	0.0492
Nd	0.968 1.2	26e-05	0.887	4.67e-13	0.983	0.00295	0.58	0.5	0.0263	0.239
Ni	0.981 0.	.00146	0.89	1.18e-12	0.98	0.0012	0.909	1	0.0286	0.157
Р	0.618 9.5	53e-24	0.971	4.19e-05	0.973	6.1e-05	0.0438	0	0.0133	0.993
Pb	0.404 1.6	61e-28	0.952	1.28e-07	0.959	9.91e-07	-0.0991	-0.0991	0.0153	0.976
S	0.49 8.5	54e-27	0.984	0.00506	0.987	0.0172	-0.0664	0	0.0191	0.782
Sr	0.712 5.6	61e-21	0.964	4.15e-06	0.983	0.00336	0.222	0.33	0.0185	0.824
Th	0.963 3.5	52e-06	0.79	5.63e-18	0.964	3.63e-06	1	1	0.0272	0.196
V	0.914 4.1	19e-11	0.723	1.39e-20	0.943	1.39e-08	1.44	1.44	0.0203	0.68
Υ	0.944 1.8	88e-08	0.923	1.97e-10	0.984	0.00602	0.444	0.5	0.0174	0.897
Zn	0.428 7.0	06e-28	0.967	1.19e-05	0.967	1.23e-05	-0.00973	0	0.0243	0.372

Most variables except pH are not normally distributed (Table 7). No variables have normal distributions when log₁₀-transformed. A few variables have normal distributions when power-transformed: pH, Ba, & Ca. The non-normal distributions of variables, even when transformed, suggest non-parametric tests are required *e.g.* for means comparisons).

check Na bimodality

This is worth doing to see if there is more than one population of samples based on salinity (assuming the main source of Na to *aqua regia* digests is halite).

The resulting map (Figure 2) shows a smaller population of low-Na locations, corresponding with where low salinity would be expected (*i.e.* away from tidal influence, and/or where evaporative concentration of Na salts is unlikely).

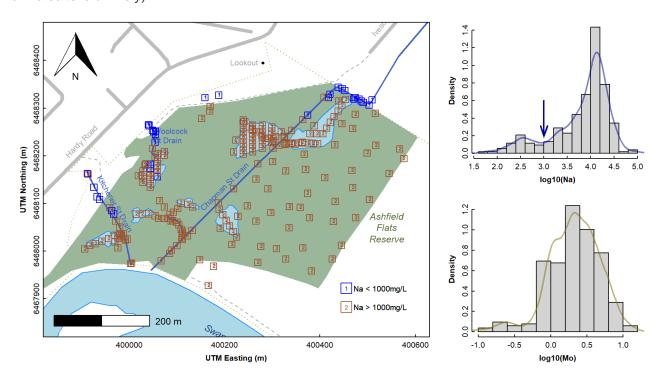


Figure 2: Density histograms for potentially multi-modal variables.

Additional explanatory maps

Surface water wetland naming

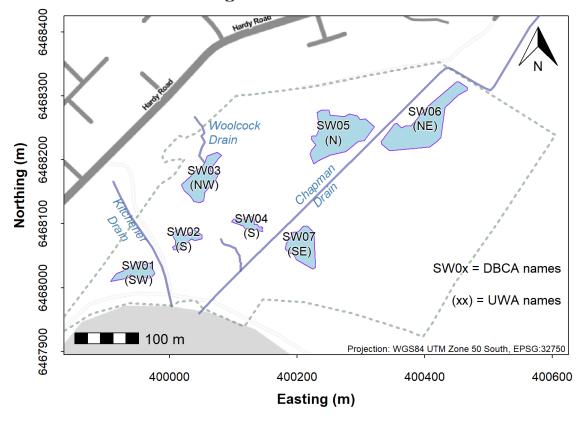


Figure 3: Map of wetland pond locations at Ashfield Flats comparing naming systems.

Sampling zones

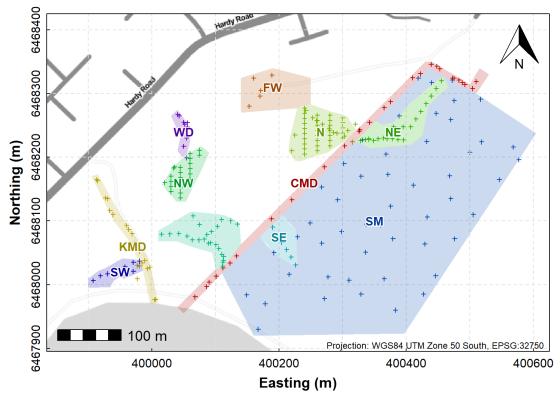


Figure 4: Map of sampling zones at Ashfield Flats 2019-2022

Means by Kruskal-Wallis

Comparing means (strictly mean rank sums) between sampling zones (the factor ZoneSimp) as this seems most appropriate. As shown in Table 8, all overall effects are significant at p<0.001. Some post-hoc pairwise comparisons show significant differences (P\$\$0.05), differing for each variable.

Table 8: Mean comparisons and pairwise letter comparisons from Kruskal-Wallis and pairwise Conover test for selected elements including REEs (REE = Σ REE), by simplified sampling zone, at Ashfield Flats 2019-2022.

CICITION O			Pairwise Compact Letters										
Variable	$KW \chi^2$	KW p	CMD	KMD	LP	N	NE	NW	S	SE	SM	SW	WD
Al	119.6	6.21e-21	abc	abcd	а	е	d	bcd	d	bde	d	bcd	ac
Ca	117.7	1.52e-20	ab	abcde	cde	С	ad	ade	ce	abd	b	cde	cde
Fe	66.56	2.03e-10	ab	ac	С	ab	abd	ac	d	abcd	ab	bd	ac
K	139.6	4.99e-25	а	ab	а	С	bd	а	cd	С	bd	bcd	а
Na	155.1	3.25e-28	а	abc	bcd	е	de	ab	f	cde	cd	f	а
S	107.4	1.81e-18	ab	cd	ae	abc	d	ae	ae	abcd	d	е	bcd
Р	69.1	6.62e-11	ab	а	bcd	cd	cd	ab	С	bcd	ab	bcd	abd
As	83.97	8.34e-14	а	abc	b	а	abc	ac	d	abc	b	abc	bc
Co	70.05	4.34e-11	abc	а	bcde	de	abc	bcde	d	bcde	ab	abce	cde
Cr	128.6	9.13e-23	ab	abc	а	d	ce	ab	ce	de	ce	bcde	а
Cu	127.6	1.41e-22	ab	cd	acd	е	be	С	е	be	ad	abcd	С
Ni	124.6	5.89e-22	ab	а	а	С	d	ab	d	cd	d	bd	а
Pb	71.35	2.43e-11	abc	abcd	d	а	а	bd	а	ac	bc	abcd	bd
Th	103.4	1.12e-17	abc	abde	С	de	de	ac	bd	е	de	abde	С
Zn	150.4	3.15e-27	а	ab	cd	е	а	cd	cde	acde	b	ace	d
La	139.2	6.06e-25	abc	abcd	а	е	ef	abc	bd	bcdf	d	bcd	ac
Ce	143.4	8.27e-26	abc	abcde	а	f	df	abc	bc	bde	е	bce	ac
Nd	141.1	2.47e-25	ab	abc	а	d	cd	ab	b	bcd	С	bc	а
Gd	133.1	1.06e-23	abc	abcd	а	е	ef	abc	bd	bdf	d	bcdf	ac
Υ	140.6	3.19e-25	ab	abc	а	d	de	ab	bc	bce	С	bce	а
REE	132.3	1.58e-23	ab	abcd	а	е	се	ab	bd	bcd	d	abd	ab

REE boxplot by Zone

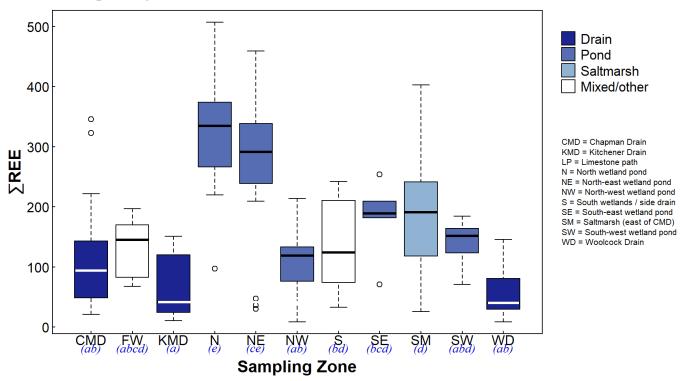


Figure 5: Sum of REE concentrations by sampling zone at Ashfield Flats for all sampling years. Different colours show environment sub-types, and means are different (p<0.05, Kruskal-Wallis) if italic text below x-axis labels has no common letters.

The greatest concentrations of Σ REE are in the north and northeast wetland pond sediments (SW05 (N) and SW06 (NE) in McGrath 2021, see Figure 5). The pattern of Σ REE means across zones is similar to that for Al in Figure 6.

Aluminium boxplot by Zone

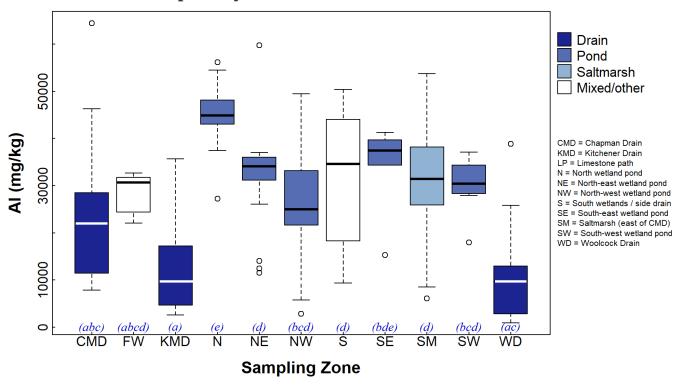


Figure 6: Al concentrations by sampling zone at Ashfield Flats for all sampling years. Different colours show environment sub-types, and means are different (p<0.05, Kruskal-Wallis) if italic text below x-axis labels has no common letters.

Bubble maps

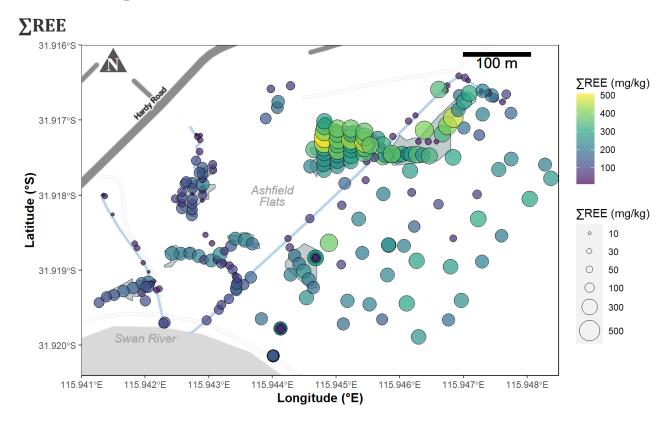


Figure 7: Map of REE concentrations by location at Ashfield Flats for all sampling years.

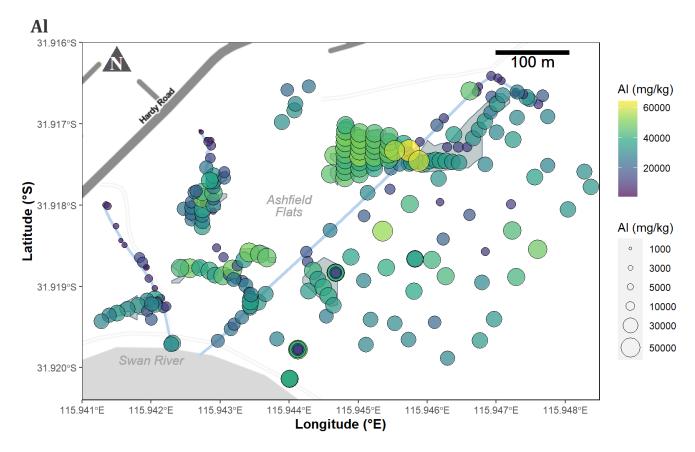


Figure 8: Map of Al concentrations by location at Ashfield Flats for all sampling years.

Boxplots by Type

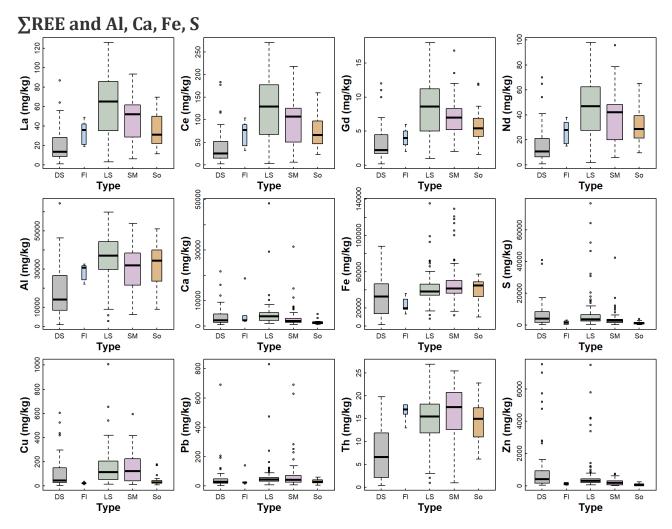


Figure 9: Boxplots of REEs and selected major and trace elements by sample type at Ashfield Flats for all sampling years (DS = Drain sediment; FI = Flooded; LS = Lake Sediment; SM = Saltmarsh; So = Soil).

Boxplots by Zone

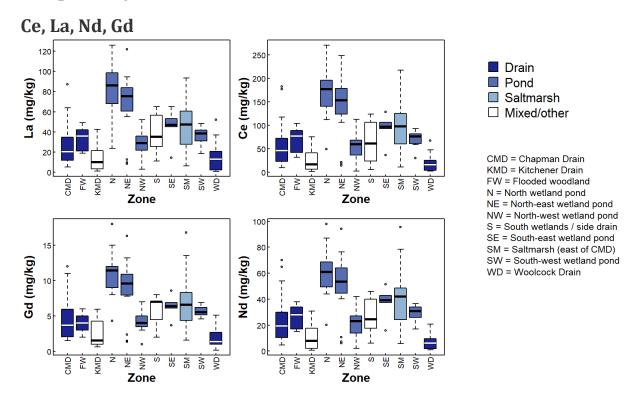


Figure 10: Boxplots of REEs by sampling Zone at Ashfield Flats for all sampling years (DS = Drain sediment; FI = Flooded; LS = Lake Sediment; SM = Saltmarsh; So = Soil).

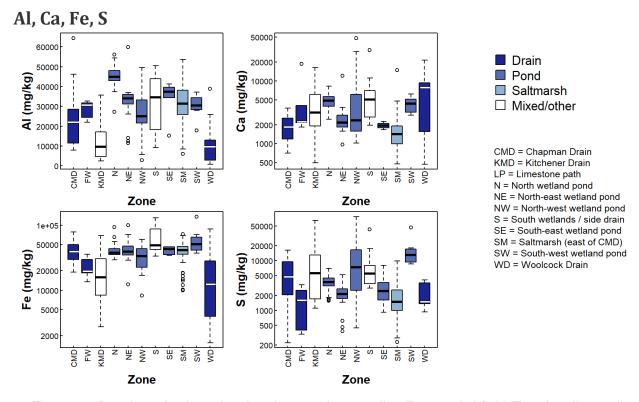


Figure 11: Boxplots of selected major elements by sampling Zone at Ashfield Flats for all sampling years.

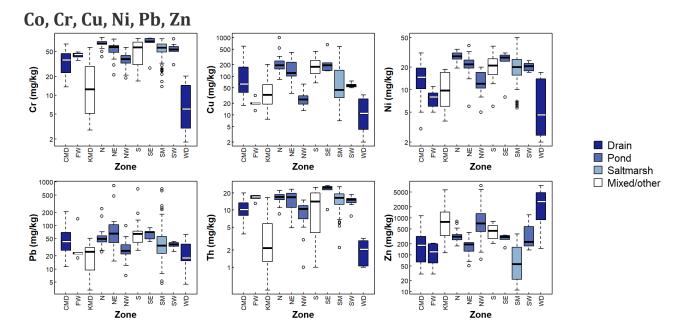


Figure 12: Boxplots of selected major elements by sampling Zone at Ashfield Flats for all sampling years (CMD = Chapman Drain, KMD = Kitchener Drain, LP = Limestone path, N = North wetland pond, NE = North-east wetland pond, NW = North-west wetland pond, S = South wetlands / side drain, SE = South-east wetland pond, SM = Saltmarsh (east of CMD), SW = South-west wetland pond, WD = Woolcock Drain).

Scatter plot matrix majors & REEs raw

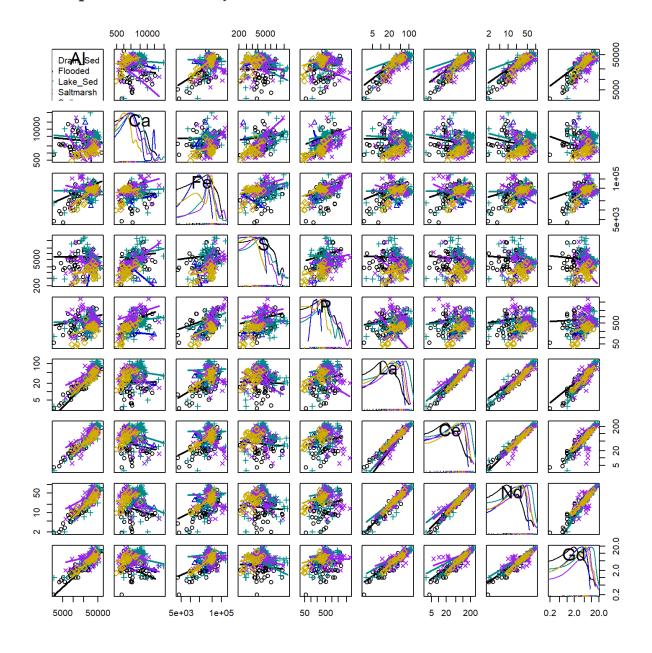


Figure 13: Scatter plot matrix for log₁₀-transformed elements at Ashfield Flats for all sampling years.

Scatter plots

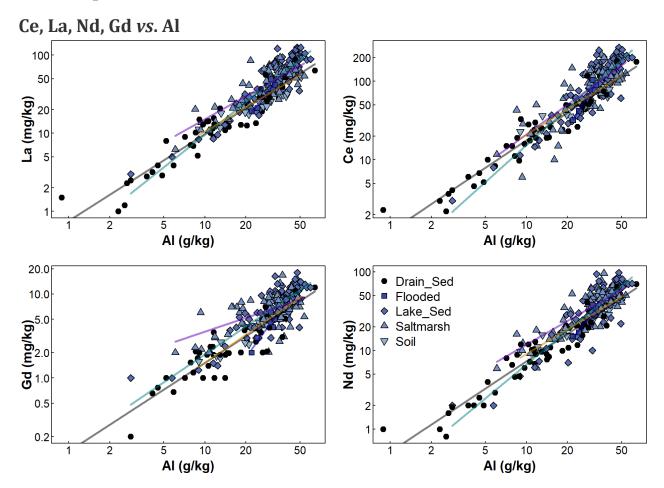


Figure 14: REE-Al relationships at Ashfield Flats for all sampling years.

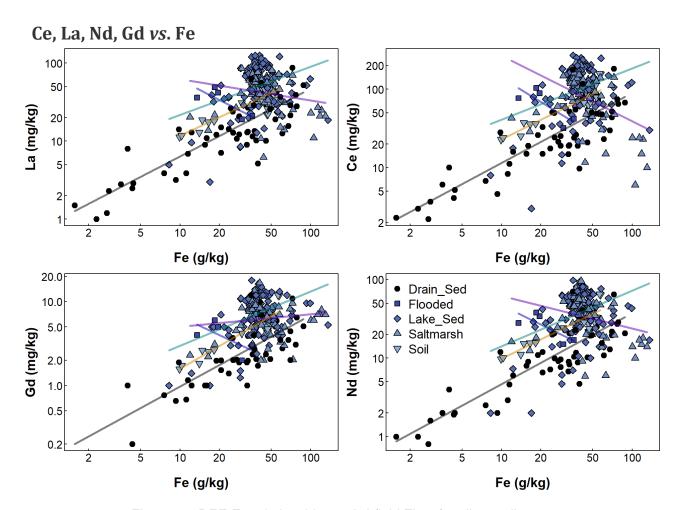


Figure 15: REE-Fe relationships at Ashfield Flats for all sampling years.

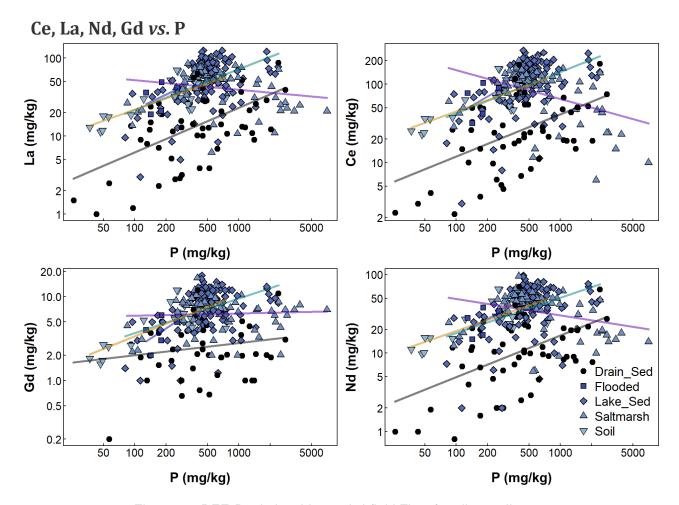


Figure 16: REE-P relationships at Ashfield Flats for all sampling years.

```
Untransformed:
      Al
            Ca
                          Ρ
                                 S
                                      La
                                            Се
                                                  Nd
                                                                         2.1 4180
    9000 5180 28800 630 32400 8.4 11.4
                                                4.7 1.24 19.0 31.1
2 37100 3280 44700 544 17900 48.6 93.0 36.6 6.51 75.8 43.2 15.7 1220
   28700 4430 48600 419 13900 41.5 81.6 32.5 5.73 54.7 32.0 15.3
                             9880 41.9 81.0 33.1 5.90 62.9 41.8 18.6
  33900 2840 53100 533
5 31200 3050 71500 744
                             8540 41.5 82.9 34.9 6.90 50.5 34.7 12.4
                             8910 35.4 72.5 29.0 5.39 58.3 36.9 16.0
6 34800 4220 38400 465
CLR-transformed:
                               S
                                              Ce
                                                      Nd
                                                             Gd
                                                                     Cu
1 4.70 4.15 5.86 2.040 5.98 -2.28 -1.970 -2.86 -4.19 -1.460 -0.969 -3.66 3.930
2 5.28 2.86 5.47 1.060 4.55 -1.36 -0.708 -1.64 -3.37 -0.913 -1.480 -2.49 1.860
3 5.20 3.34 5.73 0.978 4.48 -1.33 -0.658 -1.58 -3.31 -1.060 -1.590 -2.33 0.976
4 5.23 2.75 5.68 1.080 4.00 -1.47 -0.806 -1.70 -3.43 -1.060 -1.470 -2.28 0.396 5 5.31 2.99 6.14 1.580 4.02 -1.31 -0.619 -1.48 -3.10 -1.110 -1.490 -2.52 0.116 6 5.47 3.36 5.57 1.160 4.11 -1.42 -0.702 -1.62 -3.30 -0.920 -1.380 -2.21 0.186
```

Scatter plot matrix majors-REEs CLR-transformed

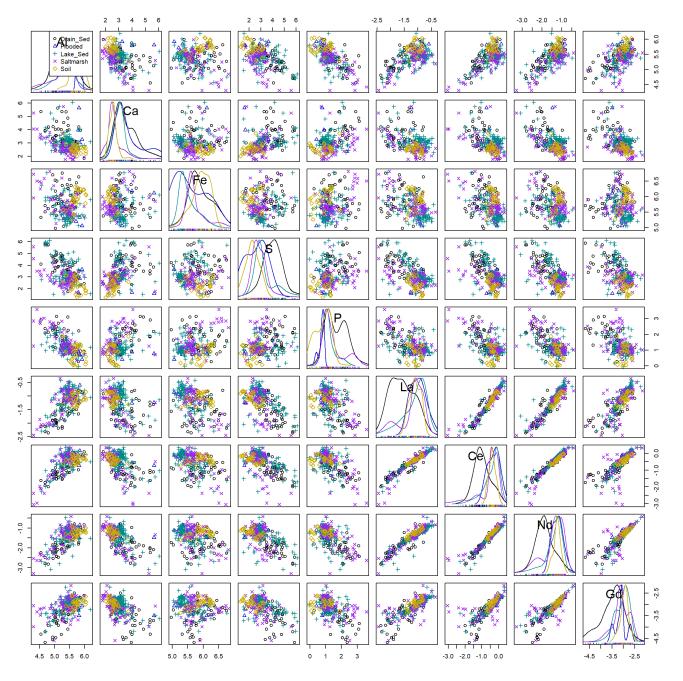


Figure 17: Scatter plot matrix of CLR-transformed elements at Ashfield Flats for all sampling years.

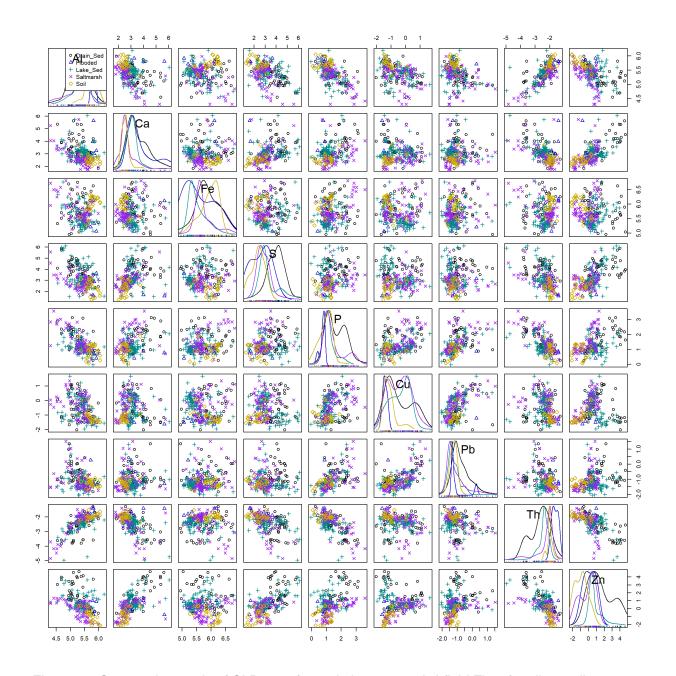


Figure 18: Scatter plot matrix of CLR-transformed elements at Ashfield Flats for all sampling years.

Scatter Plots for CLR-transformed variables

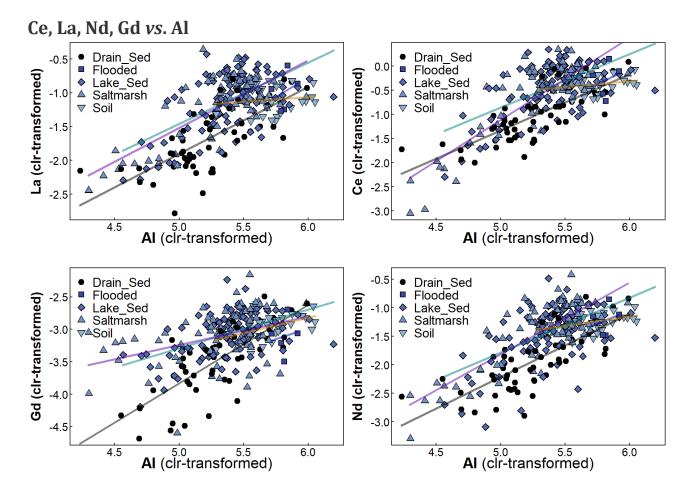


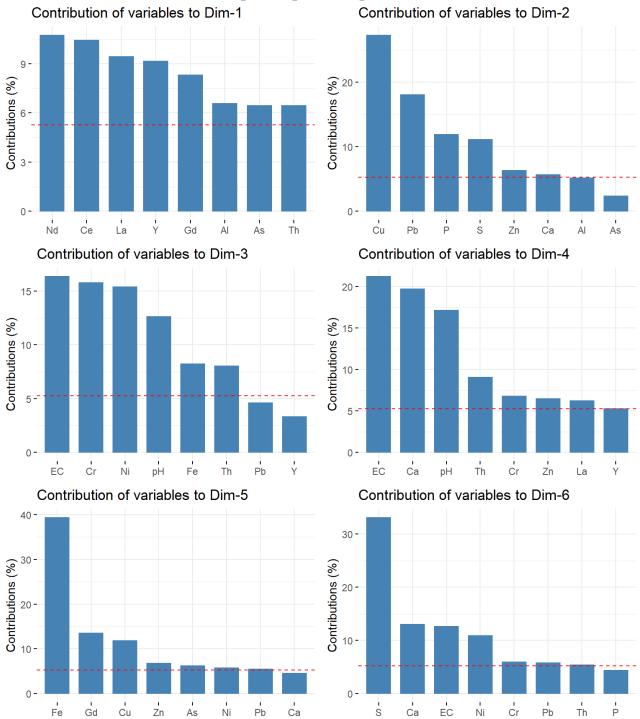
Figure 19: REE-Al relationships (concentrations CLR-transformed) at Ashfield Flats for all sampling years.

Principal components analysis

PCA Summary

```
Importance of components:
                                    PC2
                                           PC3
                                                   PC4
                          2.8922 1.4590 1.4110 1.19798 1.02381 0.9101 0.82625
## Standard deviation
## Proportion of Variance 0.4403 0.1120 0.1048 0.07554 0.05517 0.0436 0.03593
  Cumulative Proportion 0.4403 0.5523 0.6571 0.73263 0.78780 0.8314 0.86733
                                      PC9
                              PC8
                                             PC10
                                                     PC11
                                                             PC12
## Standard deviation
                          0.72077 0.71313 0.67446 0.51658 0.44520 0.41887 0.36647
## Proportion of Variance 0.02734 0.02677 0.02394 0.01405 0.01043 0.00923 0.00707
## Cumulative Proportion 0.89467 0.92144 0.94538 0.95942 0.96986 0.97909 0.98616
                           PC15
                                    PC16
                                            PC17
                                                    PC18
## Standard deviation
                          0.3515 0.24809 0.18308 0.17264 0.12075
## Proportion of Variance 0.0065 0.00324 0.00176 0.00157 0.00077
## Cumulative Proportion 0.9927 0.99590 0.99766 0.99923 1.00000
```

Contribution of variables to principal components



PCA Biplots by Type

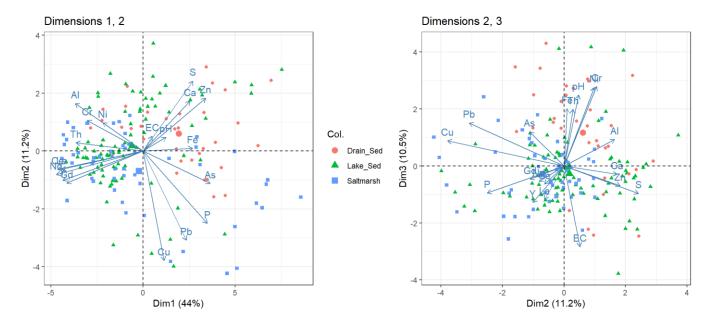


Figure 20: Principal components biplot for 2019-2022 Ashfield data grouped by sample Type.

PCA Biplots by Zone

```
bp12 <- fviz_pca_biplot(af_pca, col.ind = data0$Zone, title="PC1 & PC2 by Zone",</pre>
                                geom="point", lwd = 2, axes = c(1,2),
                                palette=rainbow(nlevels(data0$Zone), v=0.7, end=0.75)) +
     theme bw()
  bp23 <- fviz_pca_biplot(af_pca, col.ind = data0$Zone, title="PC2 & PC3 by Zone",</pre>
                               \overline{\text{geom}}=\text{"point"}, 1\text{wd}=2, axes = c(2,3),
                               palette=rainbow(nlevels(data0$Zone), v=0.7, end=0.75)) +
    theme bw()
  ggarrange (bp12, bp23, nrow = 1)
    PC1 & PC2 by Zone
                                                                     PC2 & PC3 by Zone
                                                   Col.
                                                       CMD
                                                       KMD
                                                       N_transect
                                                       N_wetland
Dim2 (11.2%)
                                                       NE
                                                       NW
                                                       NW wetland
                                                      SW_transect
                                                       Woolcock
                      Dim1 (44%)
                                                                                       Dim2 (11.2%)
```

Figure 21: Principal components biplot for 2019-2022 Ashfield data grouped by sampling Zone.

PCA Biplots by ZoneSimp

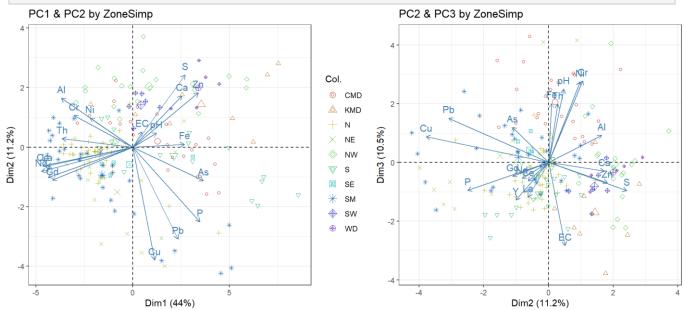


Figure 22: Principal components biplot for 2019-2022 Ashfield data grouped by simplified sampling Zone.

References

McGrath, G. S. (2021). Ashfield Flats Hydrological Study: Summary Report. Kensington, Western Australia, Department of Biodiversity, Conservation, and Attractions (Government of Western Australia). http://www.dbca.wa.gov.au/sites/default/files/2021-12/Ashfield%20Flats%20summary%20report.pdf

end of file