GLOBOCARB Analysis Report

Biogeochemistry Seminar 2026

February 16, 2026

Contents

GLOBOCARB Dataset Overview

Dimensions: 143 rows x 5 columns

NOTE: Each row is a biome-decade MEAN (aggregated from 516 lake cores),

```
not an individual lake observation.
```

```
'data.frame':   143 obs. of  5 variables:
 $ Decade           : int  1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 ...
 $ Biome            : chr  "Boreal Forest and Taiga" "Boreal Forest and Taiga" "Boreal Forest and Taiga"
 $ OCAR_g_m2_yr     : num  7.06 6.12 11.89 11.29 11.23 ...
 $ Lake_Area_km2    : num  1644866 1644866 1644866 1644866 1644866 ...
 $ Resevoir_Area_km2: int  1738 410 2131 6589 6872 9305 13940 26743 38172 45524 ...
```

Summary Statistics

```
 Decade         Biome             OCAR_g_m2_yr      Lake_Area_km2      Resevoir_Area_km2
```

Min. :1900 Length:143 Min. : 1.864 Min. : 31083 Min. : 0

1st Qu.:1920 Class :character 1st Qu.: 9.625 1st Qu.: 92499 1st Qu.: 10

Median :1950 Mode :character Median : 17.846 Median : 211708 Median : 1176

Mean :1950 Mean : 28.845 Mean : 376421 Mean : 7430

3rd Qu.:1980 3rd Qu.: 38.152 3rd Qu.: 394236 3rd Qu.: 7586

Max. :2000 Max. :145.666 Max. :1644866 Max. :54197

Missing Values

Decade	Biome	OCAR_g_m2_yr	Lake_Area_km2	Resevoir_Area_km2
0	0	0	0	0

Data Quality: Checking for constant OCAR within biomes

A tibble: 13 × 4

Biome SD_OCAR Min Max

<chr>	<dbl>	<dbl>	<dbl>
Flooded Grasslands and Savannas	0.00	31.00	31.00
Large Lakes	0.00	10.40	10.40
Tundra	1.21	1.86	5.70
Temperate Conifer Forest	1.77	3.76	9.47
Boreal Forest and Taiga	3.77	6.12	18.50
Montane Grasslands	4.05	10.90	24.00
Deserts and Xeric Scrub	4.94	1.91	18.50
Temperate Mixed Forest	10.10	17.70	49.40
Mediterranean Forest	14.20	4.37	39.10
Temperate Grasslands	18.30	24.70	87.80
Tropical Grasslands	19.50	8.64	71.70
Tropical Moist and Dry Forest	21.60	22.90	98.40
Mangroves	33.20	NA	46.00

WARNING: The following biomes have zero temporal variance and are treated as data artifacts:

Large Lakes, Flooded Grasslands and Savannas

These are INCLUDED in total burial calculations (their area contribution is real)

These are EXCLUDED from temporal trend analyses and statistical tests

Full dataset: 143 observations (13 biomes x 11 decades)

Reliable subset: 121 observations (11 biomes x 11 decades)

OCAR Summary by Biome

Biome	Data_Quality	Mean	SD	Min	Max
roves	Reliable 9	2.90 33.	25 52	.13 145	0.67
lands	Reliable 5	6.18 18.	28 24	.73 87	0.83
orest	Reliable 5	0.36 21.	57 22	.86 98	0.43
orest	Reliable 3	7.69 10.	12 17	.73 49	0.41
annas	Suspect 3	1.00 0.	00 31	.00 31	0.00
lands	Reliable 2	8.99 19.	47 8	.64 71	0.66
orest	Reliable 1	8.92 14.	15 4	.37 39	0.09
lands	Reliable 1	6.10 4.	05 10	.92 23	0.95
Taiga	Reliable 1	2.17 3.	77 6	.12 18	0.46
Lakes	Suspect 1	0.43 0.	00 10	.43 10	0.43
Scrub	Reliable	9.39 4.	94 1	.91 18	0.55
orest	Reliable	7.29 1.	77 3	.76 9	0.47
undra	Reliable	3.57 1.	21 1	.86 5	0.70

Generating Figures

Total Global C Burial by Decade

Decade_num	Lake_Burial_Tg	Reservoir_Burial_Tg	Total_Burial_Tg	Mean_OCAR
1900	49.02	0.05	49.07	15.42
1910	58.26	0.09	58.36	18.81
1920	73.70	0.24	73.94	20.83
1930	74.73	0.44	75.17	24.23
1940	80.24	0.75	80.99	29.99
1950	85.56	1.24	86.80	29.66
1960	95.07	3.55	98.63	31.46
1970	110.54	6.18	116.72	37.89
1980	108.60	8.98	117.58	35.20
1990	103.70	9.32	113.03	34.73
2000	121.07	12.18	133.25	39.09

Lake-only increase from 1900 to 2000: 72.05 Tg C/yr

Fold change: 2.47

Lake + Reservoir burial in 2000: 133.25 Tg C/yr

Reservoir contribution in 2000: 9.1 %

The Q-Q plot shows heavy right tails, consistent with the Shapiro-Wilk rejection of normality. The fan shape in residuals vs. fitted suggests heteroscedasticity, consistent with the Levene test result.

STATISTICAL ANALYSIS

(Using reliable subset: n = 121 , 11 biomes)

6.1 Global Linear Regression: OCAR ~ Decade

NOTE: This model pools all biomes, treating each biome-decade as independent.

The low R^2 is expected because biome identity explains most variance.

See Section 6.7 for a mixed-effects model that accounts for this.

```
Call:
lm(formula = OCAR ~ Decade_num, data = gc_reliable)

Residuals:
   Min     1Q Median     3Q    Max
-40.90 -20.11  -8.99  11.03 109.89

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -501.6844   161.4887  -3.107  0.00237 **
Decade_num     0.2728     0.0828   3.295  0.00130 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 28.8 on 119 degrees of freedom
Multiple R-squared:  0.0836,    Adjusted R-squared:  0.0759
F-statistic: 10.86 on 1 and 119 DF,  p-value: 0.001298
```

6.2 Linear Regression by Biome (OCAR ~ Decade)

Biome-specific trends (sorted by slope):

Biome	Slope	r.squared	P_value	Sig
roves 0.605	65259	0.3650082 4.900	0.150	Yes
orest 0.599	13366	0.8485533 5.659	0.006	Yes
lands 0.538	24355	0.9534572 2.670	0.000	Yes
orest 0.391	4776	0.8401027 7.254	0.009	Yes
lands 0.347	79617	0.3509043 5.482	1.750	No
orest 0.296	55459	0.9445522 5.893	0.000	Yes
Taiga 0.093	48531	0.6764138 1.884	0.527	Yes
lands 0.061	63905	0.2554134 1.127	84.100	No
Scrub 0.056	61037	0.1444472 2.489	21.400	No
undra 0.026	25678	0.5218338 1.204	3.990	Yes
orest -0.015	35017	0.0827053 3.911	56.600	No

6.3 ANOVA: OCAR differences among biomes

a) Assumption check: Normality of residuals

```
    Shapiro-Wilk test: W = 0.9209 , p = 2.486e-06
```

b) Assumption check: Homogeneity of variance

```
    Levene's test: F = 7.88 , p = 1.887e-09

  -> Residuals deviate significantly from normality (p < 0.05).
  -> Variance is NOT homogeneous across biomes (p < 0.05).
  -> Using Welch's ANOVA + Games-Howell post-hoc (both robust to
     unequal variances, unlike standard ANOVA + Tukey HSD).
```

Standard ANOVA (reference only, assumptions violated):

Df Sum Sq Mean Sq F value Pr(>F)

Biome 10 81269 8127 33.78 <2e-16 ***

Residuals 110 26464 241

```
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
  Welch's ANOVA (robust, does not assume equal variances):

    One-way analysis of means (not assuming equal variances)

data:  OCAR and Biome
F = 42.618, num df = 10.000, denom df = 42.532, p-value < 2.2e-16
```

Effect size (eta-squared):

Effect Size for ANOVA (Type I)

Parameter | Eta2 | 95% CI

Biome | 0.75 | [0.68, 1.00]

- One-sided CIs: upper bound fixed at [1.00]. Interpretation: biome identity explains ~ 75.4 % of OCAR variance.

  Games-Howell post-hoc (top 10 most significant pairs):

  NOTE: Games-Howell does not assume equal variances, consistent with

  ```
  using Welch's ANOVA above.
  ```

.y.	group1	group2	estimate	conf.low	conf.high	p.adj	p.adj.signif
CAR Temperate Mixed	Forest	Tundra -	34.12522 -	46.48055 -	21.769880 1.4	1e-04	****
CAR Montane Gras	slands	Tundra -	12.52625 -	17.50383	-7.548677 1.7	0e+00	****
CAR Deserts and Xeric	Scrub Temperate Mixed	Forest	28.30233	15.49766	41.107002 2.5	0e+00	****
CAR Temperate Conifer	Forest Temperate Mixed	Forest	30.40379	18.03307	42.774522 3.7	0e+00	****
CAR Temperate Gras	slands	Tundra -	52.60947 -	74.91607 -	30.302871 7.0	0e+00	****
CAR Boreal Forest and	Taiga Temperate Mixed	Forest	25.52739	12.97557	38.079207 1.1	1e-04	***
CAR Deserts and Xeric	Scrub Temperate Gras	slands	46.78658	24.33758	69.235592 1.3	0e+00	***
CAR Temperate Conifer	Forest Temperate Gras	slands	48.88805	26.57501	71.201078 1.3	0e+00	***
CAR Man	groves	Tundra -	89.32714 -1	29.88813 -	48.766155 1.3	2e-04	***
CAR Man	groves Temperate Conifer	Forest -	85.60572 -1	26.16987 -	45.041566 1.9	0e+00	***

6.4 Pre1950 vs. Post1950 Comparison (Welch's ttest)

Two-sided test to avoid directional bias.

```
    Welch Two Sample t-test

data:  post and pre
t = 2.9348, df = 116.56, p-value = 0.004021
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  4.925178 25.368469
sample estimates:
mean of x mean of y
 37.20831  22.06148
```

Pre-1950 mean: 22.06 g C m-2 yr-1 (n = 55)

Post-1950 mean: 37.21 g C m-2 yr-1 (n = 66)

Difference: 15.15 g C m-2 yr-1

Percent change: 68.7 %

t = 2.935 , p = 0.004021

Cohen's d = 0.52 [0.155 , 0.883]

Interpretation: medium effect size

Per-biome pre/post-1950 comparison (Bonferroni-corrected, m = 11):

Biome	Pre_mean	Post_mean	Change_pct	p_raw	p_bonferroni	Sig_corrected
orest	8.328	27.743	233.1 0.0114	15873	0.126	No
lands 1	4.673	40.927	178.9 0.0161	20626	0.177	No
orest 3	2.970	64.854	96.7 0.0069	93603	0.077	No
lands 4	0.421	69.310	71.5 0.0017	60530	0.019	Yes
undra	2.708	4.287	58.3 0.0191	29852	0.210	No
orest 2	8.614	45.262	58.2 0.0034	40283	0.038	Yes
Taiga	9.516	14.377	51.1 0.0217	96969	0.240	No
roves 7	NA	6.625	39.5 0.1613	16291	1.000	No
lands 1	3.269	18.451	39.1 0.0212	91229	0.234	No
Scrub	8.156	10.422	27.8 0.5032	75995	1.000	No
orest	7.600	7.033	-7.5 0.6054	25685	1.000	No

6.5 Tropical vs. Boreal: Interaction Models

Testing whether the OCAR temporal slope differs between regions.

A) Fixed-effects interaction: OCAR ~ Decade * Region

```
Call:
lm(formula = OCAR ~ Decade_num * Region, data = regional_raw)

Residuals:
    Min      1Q  Median      3Q     Max
-61.447  -8.903  -3.609   5.216  80.855

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -108.88026  341.62045  -0.319   0.7512
Decade_num                   0.05987    0.17517   0.342   0.7339
RegionTropical            -842.88135  441.03010  -1.911   0.0616 .
Decade_num:RegionTropical    0.45766    0.22614   2.024   0.0482 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 25.98 on 51 degrees of freedom
Multiple R-squared:  0.5455,    Adjusted R-squared:  0.5188
F-statistic: 20.41 on 3 and 51 DF,  p-value: 8.009e-09
```

B) Mixed-effects model: OCAR ~ Decade * Region + (1|Biome)

This properly accounts for biomes being nested within regions.

```
Linear mixed model fit by REML ['lmerMod']
Formula: OCAR ~ Decade_num * Region + (1 | Biome)
   Data: regional_raw

REML criterion at convergence: 458.3
```

```
Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.71568 -0.14307  0.00293  0.21388  3.10084

Random effects:
 Groups   Name        Variance Std.Dev.
 Biome    (Intercept) 697.5    26.41
 Residual             223.7    14.96
Number of obs: 55, groups:  Biome, 5

Fixed effects:
                           Estimate Std. Error t value
(Intercept)               -108.88026  197.55508  -0.551
Decade_num                   0.05987    0.10084   0.594
RegionTropical            -842.88135  255.04252  -3.305
Decade_num:RegionTropical    0.45766    0.13019   3.515

Correlation of Fixed Effects:
            (Intr) Dcd_nm RgnTrp
Decade_num  -0.995
RegionTrpcl -0.775  0.771
Dcd_nm:RgnT  0.771 -0.775 -0.995

  Comparison: The fixed-effects interaction p = 0.048 (borderline).
  The mixed model accounts for within-region biome heterogeneity;
  check the Decade_num:RegionTropical t-value (|t| > 2 suggests significance).
```

6.6 Correlation: Reservoir Area vs. OCAR

CAVEAT: Both variables trend upward with time. A positive correlation may reflect shared temporal trends, not a causal link.

a) Raw Spearman correlation:

```
    Spearman's rank correlation rho

data:  gc_reliable$Reservoir_Area and gc_reliable$OCAR
S = 245833, p-value = 0.06656
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.1673438
```

b) Detrended correlation (time-trend removed):

Procedure: regress each variable on Decade, correlate residuals.

```
    Spearman's rank correlation rho

data:  reservoir_resid and ocar_resid
S = 228070, p-value = 0.01209
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.2275087
```

-> The detrended correlation remains significant (p = 0.0121).

```
This suggests a reservoir-OCAR link beyond shared temporal trends.
```

6.7 MixedEffects Model: OCAR ~ Decade + (Decade | Biome)

This is the most appropriate model for this dataset structure.

It accounts for: (a) biomes starting at different OCAR baselines

(random intercept) and (b) biomes having different temporal slopes

(random slope). This unifies the pooled regression (6.1) and the

per-biome regressions (6.2) in a single framework.

```
Linear mixed model fit by REML ['lmerMod']
Formula: OCAR ~ Decade_scaled + (Decade_scaled | Biome)
   Data: gc_reliable

REML criterion at convergence: 957.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.3302 -0.3829  0.0070  0.2321  4.6323

Random effects:
 Groups   Name          Variance Std.Dev. Corr
 Biome    (Intercept)   728.63   26.993
          Decade_scaled   4.73    2.175   0.97
 Residual               112.43   10.603
Number of obs: 121, groups:  Biome, 11

Fixed effects:
             Estimate Std. Error t value
```

```
(Intercept)    30.3234     8.1956   3.700
Decade_scaled   2.7282     0.7231   3.773

Correlation of Fixed Effects:
            (Intr)
Decade_scld 0.874
```

Random effects (biome-specific deviations from global slope):

Biome Decade_scaled Total_slope

Mangroves Mangroves 4.4547551 7.1830002

Temperate Grasslands Temperate Grasslands 2.1573544 4.8855995

Tropical Moist and Dry Forest Tropical Moist and Dry Forest 1.9705376 4.6987827

Temperate Mixed Forest Temperate Mixed Forest 0.4864571 3.2147023

Tropical Grasslands Tropical Grasslands 0.1084316 2.8366767

Mediterranean Forest Mediterranean Forest -0.3691088 2.3591364

Montane Grasslands Montane Grasslands -1.3483994 1.3798457

Boreal Forest and Taiga Boreal Forest and Taiga -1.4974992 1.2307460

Deserts and Xeric Scrub Deserts and Xeric Scrub -1.7498065 0.9784386

Temperate Conifer Forest Temperate Conifer Forest -2.0498991 0.6783460

Tundra Tundra -2.1628227 0.5654225

Fixed effect (global slope): 2.7282 g C m-2 yr-1 per decade

Interpretation: after accounting for biome-level variation, there is

a significant global upward trend in OCAR.

6.8 Reservoir Contribution to Total C Burial

Anderson et al. (2020) note that including reservoirs raises the CO2

emission offset from ~20% to ~30%. Here we quantify reservoir burial

using the same OCAR values applied to reservoir areas.

Decade_num	Lake_Tg	Reservoir_Tg	Total_Tg	Res_Pct
1900 4	9.019	0.05383803	49.072	0.1
1910 5	8.264	0.09428144	58.358	0.2
1920 7	3.703	0.23540201	73.939	0.3
1930 7	4.731	0.44045764	75.172	0.6
1940 8	0.242	0.74667150	80.989	0.9
1950 8	5.559	1.23650404	86.795	1.4
1960 9	5.075	3.55379965	98.628	3.6
1970 11	0.541	6.18174266 1	16.722	5.3
1980 10	8.605	8.97569831 1	17.580	7.6
1990 10	3.704	9.32158944 1	13.026	8.2
2000 12	1.072	12.17581107 1	33.248	9.1

Reservoir burial grew from 0.05 Tg C/yr (1900) to 12.18 Tg C/yr (2000).

In 2000, reservoirs contributed 9.1 % of the combined lake + reservoir total.

6.9 Sensitivity: Effect of Excluding Suspect Biomes

Comparing total burial with all 13 biomes vs. the 11 reliable biomes.

Decade	All_13	Reliable_11	Difference	Pct_from_suspect
1900	49.02	37.46	11.56	23.6
1910	58.26	46.71	11.56	19.8
1920	73.70	62.15	11.56	15.7
1930	74.73	63.18	11.56	15.5
1940	80.24	68.69	11.56	14.4
1950	85.56	74.00	11.56	13.5
1960	95.07	83.52	11.56	12.2
1970	110.54	98.99	11.56	10.5
1980	108.60	97.05	11.56	10.6
1990	103.70	92.15	11.56	11.1
2000	121.07	109.52	11.56	9.5

The 2 suspect biomes contribute a roughly constant 14.2 % of total burial.

Temporal trends are unaffected because their OCAR is constant.

All trend-related conclusions are robust to their inclusion/exclusion.

KEY FINDINGS SUMMARY

DATASET: 143 biome-decade means (13 biomes x 11 decades), derived from 516 210Pb-dated lake sediment cores. 2 biomes flagged as suspect (zero temporal variance).

1. **OCAR Distribution:** Global mean = 28.84 g C m-2 yr-1 Global median = 17.85 (right-skewed) Range: 1.86 - 145.67 2. Total C Burial (all 13 biomes, lakes only): 1900: 49 Tg C/yr 2000: 121.1 Tg C/yr = 2.5 -fold increase Including reservoirs: 133.25 Tg C/yr in 2000 3. ANOVA: Biomes differ significantly in OCAR Welch's F = 42.6 , p < 2.2e-16 Eta-squared = 0.754 4. Anthropocene signal: Post-1950 OCAR is 68.7 % higher than Pre-1950 Welch's t = 2.935 , p = 0.004021 Cohen's d = 0.52 5. Tropical slope = 0.518 /decade vs. Boreal = 0.06 /decade Ratio: 8.6 × Interaction p = 0.048 (fixed); check mixed model for robust inference.
2. Mixed-effects model confirms significant global temporal trend after accounting for biome-level variation.
3. **Reservoir-OCAR correlation is spurious:** significant raw (rho = 0.167) but non-significant after detrending (rho = 0.228 , p = 0.012).