

1 **METADATA**

2 **CLASS I. DATA SET DESCRIPTORS**

3 **A. Data set identity:** Morphometric and biomass data for 11 emergent marsh plant species

4 **B. Data set identification code:** GCRew_Allom_SCAM_Data.csv,
5 GCRew_Allom_Other_Data.csv, GCRew_Allom_General_Data.csv,
6 GCRew_Variable_Description.csv, GCRew_Allom_Code.R,
7 GCRew_Allom_Model_Summary.pdf.

8 **C. Data set description:** Data sets contain measurements of tidal marsh plants that can be used
9 to build allometric models of biomass as a function of height and width. Measurements come
10 from plants collected at the Smithsonian Institution's Global Change Research Wetland
11 (GCRew), which is a brackish marsh on the Chesapeake Bay. Eleven graminoid, forb, and shrub
12 species that are common in the mid-Atlantic region of North America are included. Allometric
13 models, evaluations of the models, and the code used to generate them are also provided.

14 **Originators:**

15 Meng Lu¹, Joshua S. Caplan^{1,2}, Jonathan D. Bakker³, J. Adam Langley^{1,4}, Thomas J. Mozdzer^{1,2},
16 Bert G. Drake¹, and J. Patrick Megonigal¹

17 ¹ Smithsonian Environmental Research Center, Edgewater, MD, USA

18 ² Department of Biology, Bryn Mawr College, Bryn Mawr, PA, USA

19 ³ School of Environmental & Forest Sciences, University of Washington, Seattle, WA, USA

20 ⁴ Department of Biology, Villanova University, Villanova, PA, USA

22 **Abstract:** Coastal marshes are highly valued for ecosystem services such as protecting inland
23 habitats from storms, sequestering carbon, removing nutrients and other pollutants from surface
24 water, and providing habitat for fish, shellfish, and birds. Because plants largely determine the
25 structure and function of coastal marshes, quantifying plant biomass is essential for quantifying
26 these ecosystem services, understanding the biogeochemical processes that regulate ecosystem
27 function, and forecasting tidal wetland responses to accelerated sea level rise. Allometry is a
28 convenient and efficient technique for non-destructive estimation of plant biomass, and it is
29 commonly applied to studies of carbon and nitrogen cycles, energy flows, and marsh surface
30 elevation. We present plant allometry data and models developed for three long-term
31 experiments at the Smithsonian Global Change Research Wetland, a brackish marsh in the
32 Rhode River subestuary of the Chesapeake Bay. The dataset contains 9771 measurements for
33 stem height, dry mass, and (in 9667 cases) width across 11 plant species. The vast majority of
34 observations are for *Schoenoplectus americanus* (8430) and *Phragmites australis* (311), with
35 fewer observations for other common species: *Amaranthus cannabinus*, *Atriplex patula*, *Iva*
36 *frutescens*, *Kosteletzkya virginica*, *Polygonum hydropiper*, *Solidago sempervirens*, *Spartina*
37 *alterniflora*, *Spartina cynosuroides*, and *Typha angustifolia*. Allometric relationships take the
38 form of linear regressions of biomass (transformed using the Box-Cox procedure) on either stem
39 height and width, or stem height alone. Allometric relationships for *Schoenoplectus americanus*
40 were measurably, but not meaningfully, altered by elevated CO₂, N enrichment, the community
41 context, interannual variation in climate, or year, showing that a single equation can be used
42 across a broad range of conditions for this species. The archived files include: (1) raw data used
43 to derive the allometric equations for each species, (2) reports and evaluations of the allometric
44 equations we derived using the data, (3) R code with which our derivations can be replicated.

45 Methodological details of the experiments, the allometry data collection effort, and our statistical
46 analysis are described in the metadata. The allometric equations can be used for biomass
47 estimation in empirical and modeling studies of North American coastal wetlands, and the data
48 can be used in ecological studies of terrestrial plant allometry.

49 **D. Key words:** allometric equations, biomass estimation, coastal wetlands, elevated CO₂,
50 nitrogen pollution, plant allometry

51 **CLASS II. RESEARCH ORIGIN DESCRIPTORS**

52 **A. Overall project description**

53 **Identity:** Long term global change research in a tidal marsh of the Chesapeake Bay

54 **Originators:** Bert G. Drake, J. Patrick Megonigal, J. Adam Langley, Thomas J. Mozdzer

55 **Period of study:** Ongoing since 1987

56 **Objectives:** Three primary experiments forecast the effects of interacting global change factors
57 (elevated atmospheric CO₂ concentration, nitrogen eutrophication and plant invasion) on native
58 tidal marsh ecosystems. Experiment 1 exposed three plant communities to elevated CO₂ alone,
59 Experiment 2 exposed a single plant community to elevated CO₂ and nitrogen eutrophication,
60 and Experiment 3 exposed *Phragmites australis* and the native plant community it is invading to
61 elevated CO₂ and nitrogen eutrophication. A goal of all of these experiments is to determine how
62 net primary productivity (NPP) shifts as a function of the global change factors individually and
63 in combination. Additional goals include determining how global change factors alters
64 belowground carbon storage and marsh resilience to sea level rise.

65 **Sources of Funding:** GCRew has been supported by the DOE since 1986, the DOE-TES
66 program (Award DE-SC0008339), the NSF-LTREB program (grants DEB-0950080 and DEB-
67 1457100), and the Maryland Sea Grant program (award SA7528114-WW).

68 **B. Specific subproject description**

69 **System description:** The allometry data published here come from three long-term experiments
70 at Kirkpatrick Marsh (38°52.5' N, 76°32.9' W), a 23 ha brackish saltmarsh on the Rhode River
71 subestuary of the Chesapeake Bay in the eastern United States. The site is operated by the
72 Smithsonian Environmental Research Center, and is also known as the Global Change Research
73 Wetland (GCRew). This high-marsh ecosystem is typical of many Atlantic coast tidal wetlands
74 in North America. The native plant community is dominated by the perennial C₃ sedge
75 *Schoenoplectus americanus* and two co-dominant C₄ grasses, *Spartina patens* and *Distichlis*
76 *spicata*. The lineage of *Phragmites australis* that is invasive in North America (haplotype M) has
77 also become widespread at GCRew. Several other plant species are common in the marsh, but
78 occur almost exclusively in the third of the experiments described below. Salinity at the site
79 varies from 4-15 ppt (mean = 10 ppt) and the mean tidal range is approximately 40 cm. The
80 high-marsh platform is 40-60 cm above the mean low water level, and soils contain
81 approximately 80% organic matter (peat) to a depth of 5 m. The region's mean annual
82 temperature is 14.1°C and its total annual precipitation averages 948 mm.

83 **Experimental design:**

84 Experiment 1 (CO₂ × Community): Ten circular open-top chambers (OTCs; 0.8 m diameter, 1.0 m
85 height) were placed in each of three plant communities in 1986 (Drake 2014). One was a C₃

86 community dominated by *S. americanus*, another was a C₄ community dominated by *S. patens*
87 and *D. spicata*, and the third was a mixed C₃-C₄ community. Beginning in 1987, the CO₂
88 concentration was elevated to 700-720 ppm ($\mu\text{L CO}_2 \text{ L}^{-1}$) in five chambers per community (eCO₂
89 treatment), while air in the remaining OTCs was unamended (aCO₂ treatment). To determine if
90 OTCs affected plant growth, five non-chambered control plots in each community were also
91 included in the study. CO₂ exposure began each year on May 1st and ended on October 31st,
92 capturing the period starting when plants emerged in spring and continuing 24 hours per day
93 through senescence in autumn. A survey of all plots conducted at the outset of the experiment
94 showed no significant differences in plant biomass assigned to the three treatments (aCO₂, eCO₂,
95 and Control) within each community.

96 Experiment 2 (CO₂ × Nitrogen): Twenty circular OTCs (2.0 m diameter, 1.5 m height) were
97 established adjacent to the site of Experiment 1 in 2006 (Langley et al. 2009, Langley et al.
98 2010). The plant community was similar to the C₃ community in Experiment 1, as it was a near-
99 complete monoculture of *S. americanus*. In addition to elevating CO₂ in half of the chambers (to
100 the same eCO₂ level as used in Experiment 1), interactions between CO₂ and nitrogen (N) were
101 evaluated by adding N to half of the chambers in each CO₂ treatment group. N was added by
102 spraying NH₄Cl monthly from May to September each year, totaling 25 g N m⁻² yr⁻¹. This
103 yielded four treatments applied to each of five chambers. CO₂ exposure occurred during daylight
104 hours between May 1st and October 31st.

105 Experiment 3 (Phragmites × CO₂ × Nitrogen): Twelve rectangular OTCs (1.25 × 2.5 × 4.4 m
106 height) were installed at the leading edge of an expanding *Phragmites australis* stand in 2011
107 (Caplan et al. 2015). The native plant community was primarily comprised of *S. americanus*, *S.*

108 *patens*, and *D. spicata*, though additional C₃ species were present in this experiment that were
109 absent from or rare in the others. These included *Amaranthus cannabinus*, *Atriplex patula*, *Iva*
110 *frutescens*, *Kosteletzkya virginica*, *Polygonum hydropiper*, *Solidago sempervirens*, *Spartina*
111 *alterniflora*, *Spartina cynosuroides*, and *Typha angustifolia*. As in Experiment 2, half of the
112 OTCs were elevated to 700-720 ppm CO₂, and three OTCs of each CO₂ treatment type were
113 fertilized with 25 g N m⁻² yr⁻¹. CO₂ exposure occurred during daylight hours between May 1st
114 and October 31st.

115 **Research methods**

116 **a. Biomass estimation:** We estimate peak annual aboveground biomass for plants in all
117 experimental chambers. For species with appropriate forms, we developed allometric equations
118 to predict biomass from measurements of stem height and width (provided in the Model
119 Summary). An allometric approach was used rather than direct measurements of biomass (i.e.,
120 via harvests) in order to minimize disturbance to the long-term experimental plots. To generate
121 allometric equations, we used destructively harvested shoots collected outside of chambers
122 and/or from a small subset of plants within chambers. Aboveground biomass in each plot could
123 then be determined from morphometric measurements collected during an annual census of
124 individual shoots in the experimental plots. Because the short-stature grasses (*S. patens* and *D.*
125 *spicata*) occur at high density and have a form that does not lend itself well to allometry (BG
126 Drake, personal observation), we instead measured the biomass of these species by destructively
127 harvesting replicate subplots (5×5 cm in Experiments 1 and 2; 7×7 cm in Experiment 3); these
128 data are not reported here.

129 Allometry data were collected using consistent methods across experiments through time.
130 Methods were consistent across species except where noted. Plants were harvested in late July to
131 early August each year, at the time of maximum standing biomass. *Schoenoplectus americanus*
132 stems were harvested from within or near the OTCs, whereas stems of all other species were
133 harvested from outside the OTCs. All stems were clipped within 1 cm of the ground surface.
134 Stem heights were measured to the nearest centimeter, usually within 1-3 hours of harvesting and
135 always within 24 hours. Measurements were made by placing the central axis of a plant along a
136 measuring tape and determining the maximum distance between stem base and the stem tip.
137 More precisely, we measured from a line perpendicular to the stem base to a line parallel to it but
138 tangent to the most apical part of the plant, akin to the maximum Feret diameter. Stem widths
139 were measured at the same time to the nearest millimeter with digital calipers; for *S. americanus*
140 calipers were aligned perpendicular to the widest edge of the three-edge stem, and in the middle
141 third of the stem's height. For all other species, widths were measured 40 cm above the base of
142 the plant. We found a measurement taken in the middle third of *S. americanus* stems to be
143 sufficient because width is nearly constant in that section of the plant. Plants were weighed after
144 drying for at least 96 hours at 60°C. Data for *S. americanus* extend back to 1987, while data for
145 *P. australis* and most other species were collected after Experiment 3 began in 2011. Data were
146 collected annually for four species (including *S. americanus* and *P. australis*), over 2-4 years for
147 five species, and only in 2015 for two species (*S. alterniflora* and *T. angustifolia*).

148 The large dataset for *S. americanus* was initially collected under the assumption that allometry
149 would shift with experimental factors. A new set of equations for *S. americanus* was determined
150 annually in each of the treatments in Experiment 1 (Control, aCO₂, and eCO₂) and the two
151 communities in which *S. americanus* was abundant (C₃ and mixed), but not in the C₄-dominated

152 community where *S. americanus* stem density was low. Three to five shoots from each OTC
153 were harvested annually to build the allometric equations. However, treatment and community
154 had negligible effects on allometry (Curtis et al. 1989). We therefore sought to establish a single
155 set of equations that can be used to estimate shoot dry mass for all of the common plant species
156 in this tidal marsh ecosystem, excluding short-stature grasses. In addition to using these
157 equations for quantifications of NPP at GCRew (which we assume is equivalent to peak season
158 biomass), we also intend them to be used by others for estimating aboveground plant biomass in
159 tidal marshes. Note that our equations are based on plants harvested at the time of maximum
160 standing biomass and may be less accurate if applied to plants measured earlier or later in the
161 growing season.

162 **b. Allometric models:** We defined allometric models that could be used to accurately estimate
163 aboveground biomass (Biomass) for 11 of the most common plant species in Rhode River,
164 Chesapeake Bay high-marsh communities (see Model Summary). We also defined models
165 combining data from graminoid or forb species that could be used for estimating the biomass of
166 species not included in the dataset. For these “general” models, data from 30 individuals from
167 each of the 5 graminoid or 5 forb species were combined into new datasets. To do this, we
168 randomly selected three individuals from each decile of Biomass for inclusion (n = 30 per
169 species, or all data if <30 were available); we only included complete cases that were not outliers
170 in the allometric analysis of the contributing species (see below). For each of the 11 species and
171 the two general plant forms, we developed models that can be used to estimate Biomass from
172 Height and Width, or from Height alone. The latter was done to accommodate situations where
173 measurements of Width (i.e., diameter) are not available, which occurs in our dataset and will
174 likely occur in other datasets as well. Note that the dataset for *S. americanus* contains a number

175 of placeholder records for quadrats where no plants were found, i.e., rows with values of zero for
176 Height, Width, and Biomass; these were removed for modeling purposes. Records of *S.*
177 *americanus* indicating that stem tips had been cut off (typically due to breakage or herbivory)
178 were also excluded because the allometry of such stems is atypical for the species.

179 Our allometric equations take the form of linear models with transformed Biomass but
180 untransformed predictor variables. We transformed Biomass to maximize residual normality,
181 which we did by selecting the value of λ (the exponent of a power transformation on Biomass, or
182 $\ln(\text{Biomass})$ if $\lambda=0$) that maximized a model's log-likelihood (Box and Cox 1964). Candidate
183 values of λ were restricted to the set $\left\{-2, -\frac{3}{2}, -1, -\frac{2}{3}, -\frac{1}{2}, -\frac{1}{3}, 0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1, \frac{3}{2}, 2\right\}$. We also
184 evaluated model formulations with untransformed Biomass, but with transformed and
185 multiplicative combinations of predictor variables (e.g., Height^2 , Width^2 , and $\text{Height}^2 \times \text{Width}$).
186 Because models with optimized transformations of Biomass performed as well or better than the
187 other formulations (i.e., greater R^2 , residuals more normal and homoscedastic), we restricted
188 subsequent analysis to linear regressions of transformed Biomass with predictors including
189 combinations of Height, Width, and their interaction ($\text{Height} \times \text{Width}$).

190 After selecting a value of λ for the full model, we compared models with all subsets of terms in
191 that model, including the intercept only (i.e., the null) model. We primarily used the coefficient
192 of determination (R^2) to evaluate model performance, as our priority was maximizing predictive
193 ability. However, we also computed the Bayesian information criterion (BIC) for all models,
194 which accounts for both predictive ability and model size (it does the latter by penalizing the
195 addition of terms). Reductions in BIC (i.e., improvements in performance) could thus be used to
196 quantify the value of adding Width (or $\text{Height} \times \text{Width}$) to the Height model, or, by comparison to

197 the null, determine if a model had strong predictive ability to start with; reductions in BIC > 4
198 are typically considered substantial improvements. Because no additional data were needed to
199 compute the Height×Width interaction but always improved R², full models were selected for all
200 Height & Width models. In two cases, Height models had a different optimal value of λ than the
201 full model for the same species; for these species a separate model evaluation was performed and
202 the appropriate transformation was used.

203 Most species had at least one data point that was an outlier with respect to Height or Width. Our
204 quality control procedures determined that some of these (~10 observations) resulted from
205 typographical errors; the dataset was updated with correct values in these cases. While some
206 remaining outliers may have been due to mismeasurement, others may have resulted from the
207 inclusion of plants with atypical morphologies. Because our objective was to generate equations
208 for estimating Biomass for plants with the typical morphology of their species, we excluded
209 outliers in the final linear modeling phase. The criteria used to define outliers were consistent
210 across species; we considered outliers to be points with Studentized residuals ≥ 3 (i.e., those
211 extremely different from other values in the dataset) or Studentized residuals ≥ 2 together with
212 Cook's distance ≥ 0.05 (i.e., those moderately different from other values but with high influence
213 on model coefficients). Points that met these criteria are shown in orange on the *Effect of*
214 *Transformation* plots in the Model Summary, though they have been excluded from the analysis
215 and therefore do not appear in *Diagnostic Plots*. Across species, outliers comprised an average of
216 2.6% of the observations used in Height & Width models and 2.3% of the observations used in
217 Height models. Outliers comprised no more than 6.7% of observations.

218 We evaluated whether CO₂, N and Community treatment groups, or variation in climate (i.e.,
219 temperature and precipitation) affected the allometry of *S. americanus*. We carried out a multi-
220 model comparison like that described above, but with terms for CO₂ status, N status, the CO₂×N
221 interaction, and Community included in the full model. We focused on *S. americanus* for this
222 evaluation because very few plants from other species were collected within chambers. In
223 addition to the analysis with the full dataset, we fit models with random subsamples of the data
224 (n = 100 to 1000) to determine how power affected the outcome. In no case was model fit
225 appreciably improved by the including any or all of the four additional terms (R² increased by no
226 more than 0.003). However, BIC was highly sensitive to sample size, and was minimized with
227 larger models for sample sizes >~500. Separately, we conducted a multi-model comparison with
228 terms for total annual precipitation, mean monthly temperature, minimum monthly temperature,
229 and year (as a surrogate for other variables like mean annual sea level). Again, effects were
230 statistically discoverable given the size of the dataset, but negligible for the purposes of
231 predicting allometry (R² improved by no more than 0.005). We conclude that treatments,
232 community, and climate did not have meaningful effects on allometry, and that Height and
233 Width are sufficient to accurately estimate biomass in the context of quantifying ecosystem level
234 NPP.

235 **Project personnel:**

236 Data compilation: Meng Lu

237 Allometric modeling: Joshua S. Caplan, Jonathan D. Bakker

238 Principle Investigators: J. Patrick Megonigal, Bert G. Drake, J. Adam Langley, Thomas J.
239 Mozdzer

240 Technicians: James Duls, Gary Peresta, Andrew Peresta

241 **CLASS III. DATA SET STATUS AND ACCESSIBILITY**

242 **A. Status**

243 **Latest data update:** August 2016

244 **Latest metadata update:** August 2016

245 **Data verification:** Potential typographical errors were identified and corrected when found.

246 **B. Accessibility**

247 **Storage location and medium:** Additional copies of the data are posted on the Global Change

248 Research Wetland website: <https://serc.si.edu/gcrew/>

249 **Contact person(s):** Patrick Megonigal, Smithsonian Environmental Research Center,
250 Edgewater, Maryland, USA. Email: megonigalp@si.edu, Tel: 443-482-2346

251 **Copyright or proprietary restrictions:** Data are freely available provided that the source of the
252 data is properly attributed to this publication and to the Smithsonian Environmental Research
253 Center. We request that you alert us when you download the datasets so we can be aware of how
254 they are being used.

255 **Citation:**

256 **Disclaimers:** The above descriptions of the experiments at GCRew are not comprehensive.
257 Please contact the authors to discuss your intended use of the data and how nuances of the
258 experimental methods may affect your analysis.

259 **Costs:** None

260 **CLASS IV. DATA STRUCTURAL DESCRIPTORS**

261 **A. Data Set Files**

262 **Identity and Size:** 3 data files, 1 variable description file; 1 R code file; 1 model summary file.

263 1. GCRew_Allom_SCAM_Data.csv: Allometry data for *Schoenoplectus americanus*. (Size:
264 670KB)

265 2. GCRew_Allom_Other_Data.csv: Allometry data for species other than *S. americanus*. (Size:
266 40KB)

267 3. GCRew_Allom_General_Data.csv: Data used generate general allometric equations for forbs
268 and graminoids. (Size: 11KB)

269 4. GCRew_Variable_Description.csv: Descriptions of the variables in the datasets. (Size: 4KB)

270 5. GCRew_Allom_Code.R: R code for generating allometric equations. (Size: 10KB)

271 6. GCRew_Allom_Model_Summary.pdf: Summary of allometric models. (Size: 3380KB)

272 **Format and storage mode:** Data files and the variable description file contain text stored as
273 comma separated values (CSV), R code is stored as text; and the Model Summary is a portable
274 document format (PDF) file.

275

276 **ACKNOWLEDGMENTS**

277 We thank Gary Peresta, James Duls, Andrew Peresta, Liza McFarland, Rachel Hager, and the
278 hundreds of volunteers who helped collect data during annual censuses. We are grateful to
279 Alison Cawood of the SERC Citizen Science program for coordinating the volunteers. This
280 research was largely supported by the NSF-LTREB program (awards DEB-0950080 and DEB-
281 1457100), the DOE-TES program (award DE-SC0008339), two decades of funding from DOE to
282 B.G. Drake, and the Maryland Sea Grant program (award SA7528114-WW).

283 **LITERATURE CITED**

284 Box, G. E. P., and D. R. Cox. 1964. An analysis of transformations. *Journal of the Royal*
285 *Statistical Society Series B.* 26:211-252.

286 Caplan, J. S., R. N. Hager, J. P. Megonigal, and T. J. Mozdzer. 2015. Global change accelerates
287 carbon assimilation by a wetland ecosystem engineer. *Environmental Research Letters*
288 10:115006.

289 Curtis, P. S., B. G. Drake, P. W. Leadley, W. J. Arp, and D. F. Whigham. 1989. Growth and
290 senescence in plant communities exposed to elevated CO₂ concentrations on an estuarine marsh.
291 *Oecologia.* 78:20-26.

292 Drake, B. G. 2014. Rising sea level, temperature, and precipitation impact plant and ecosystem
293 responses to elevated CO₂ on a Chesapeake Bay wetland: Review of a 28-year study. *Global*
294 *Change Biology*. 20:3329-3343

295 Langley, J. A., and J. P. Megonigal. 2010. Ecosystem response to elevated CO₂ levels limited by
296 nitrogen-induced plant species shift. *Nature*. 466:96-99

297 Langley, J. A., M. V. Sigrist, J. Duls, D. R. Cahoon, J. C. Lynch, and J. P. Megonigal. 2009.
298 Global change and marsh elevation dynamics: Experimenting where land meets sea and biology
299 meets geology. Pages 391-400 *in* M. A. Lang, I. G. Macintyre, and K. Rutzler editors.
300 Proceedings of the Smithsonian Marine Sciences Symposium. Contributions to the Marine
301 Sciences 38. Smithsonian Institution Scholarly Press, Washington DC.