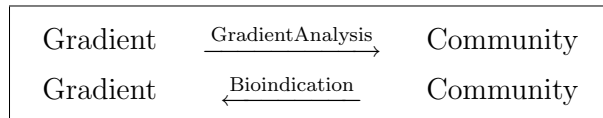


Gradient Analysis

Relation of species and environmental variables or gradients.

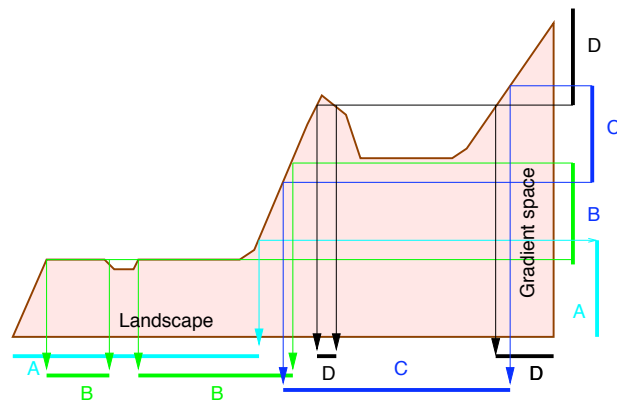
Individualistic species responses.



Gradient types

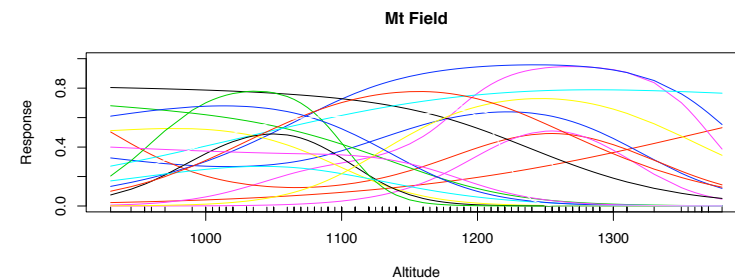
- Direct gradients:** Influence organisms but are not consumed.
 - Correspond to conditions.
- Resource gradients:** Consumed
 - Correspond to resources.
- **Complex gradients.** Covarying direct and/or resource gradients: Impossible to separate effects of single gradients.
 - Most observed gradients.

Gradients and landscape



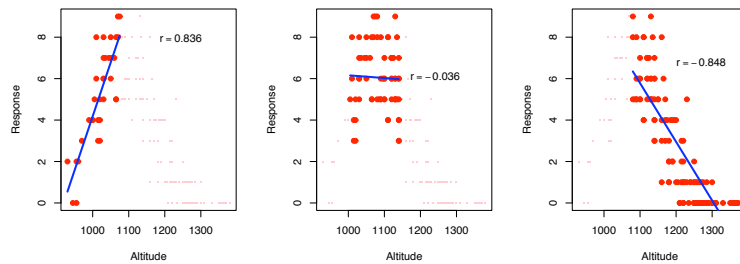
Species responses

Species have non-linear responses along gradients.



Linear models are inadequate

The slope, sign and significance depend on the studied range on the gradient

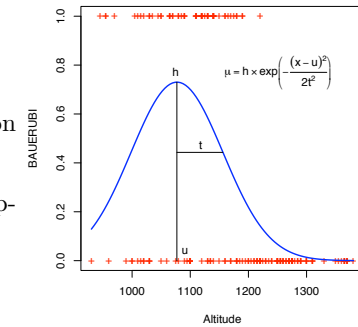


Gaussian response

$$\mu = h \exp\left(-\frac{(x-u)^2}{2t^2}\right)$$

Three interpretable parameters:

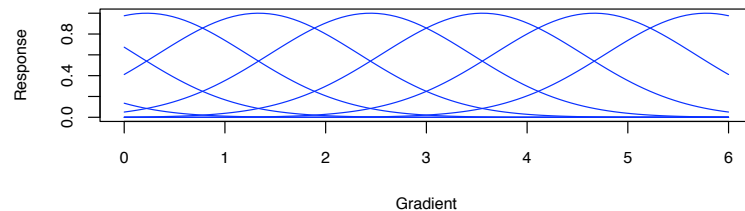
1. Location of optimum u on gradient x
2. Expected height h at the optimum
3. Width t of the response



Dream of species packing

Species have Gaussian responses and divide the gradient optimally:

- Equal heights h .
- Equal widths t .
- Evenly distributed optima u .



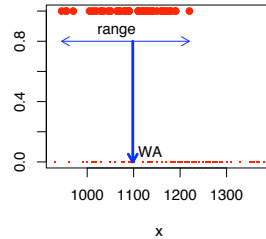
Evidence for Gaussian response

- Whittaker described many response types: multimodal, skewed, flat, plateaux and symmetric.
- Only a small part of responses were regarded as symmetric, still became the standard.
- First canonized in coenocline simulations.
- Species packing is the theoretical basis of (canonical) correspondence analysis.

Weighted averages

- Weights: Species abundances y .
- Gives u as the average on x .
- Presence-absence data: the average of site values where species occurs.
- Quantitative data: more weight to sites where species is more abundant.
- Symmetric: Species optima u to estimate gradient values x

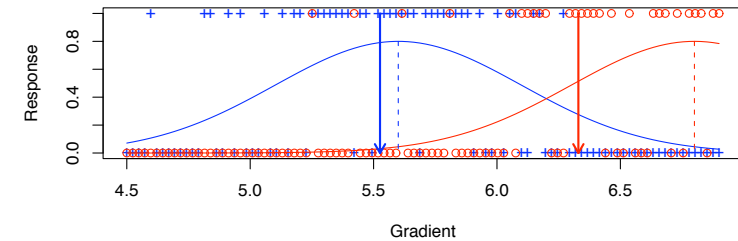
$$\tilde{u}_j = \frac{\sum_{i=1}^N y_{ij} x_i}{\sum_{i=1}^N y_{ij}}$$



Bias and truncation

Weighted averages are good estimates of Gaussian optima, *unless* the response is truncated.

Bias towards the gradient centre: shrinking.

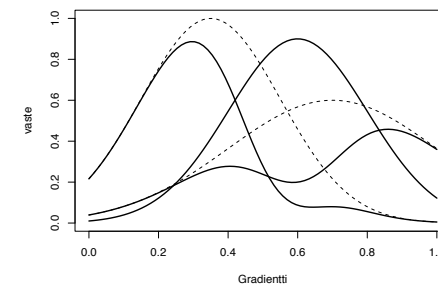


Popular response models

- **Gaussian response model:** The most popular model that gives symmetric responses, and is the basis of much of theory of ordination and gradient analysis.
- **Beta response:** Able to produce responses of varying skewness and kurtosis, and challenges the Gaussian dominance.
- **HOF response:** A family of hierarchic models which can be produce skewed, symmetric or different monotone responses, and can be used to analyse the response shape.
- **GAM models:** Can find any smooth shape and fit any kind of smooth response.

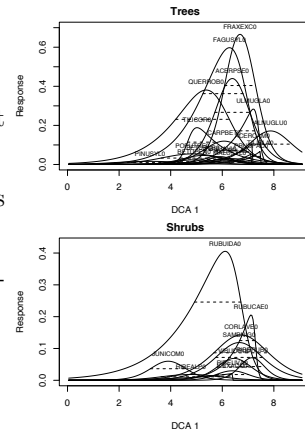
Shape matters

Fundamental response can be symmetric, but **realized** response skewed or multimodal due to species interactions.



Real World (almost)

- In Danish beech forests, dominant species skew other species away
- Austin predicted skewed responses at gradient ends
- A decent gradient would be nice instead of a DCA axis...



Gaussian response: a case of GLM

Can be reparametrized as a generalized linear model:

$$\mu = h \exp\left(\frac{(x-u)^2}{-2t^2}\right)$$

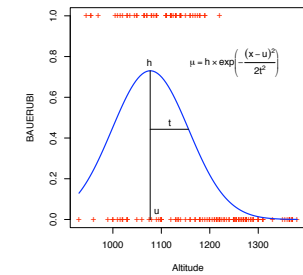
$$\log(\mu) = b_0 + b_1x + b_2x^2$$

- Gradient as a 2nd degree polynomial.
- Logarithmic link function.

$$u = -\frac{b_1}{2b_2}$$

$$t = \sqrt{-\frac{1}{2b_2}}$$

$$h = \exp\left(b_0 - \frac{b_1^2}{4b_2}\right)$$



Generalized linear models: a refresher

1. **Linear predictor** η : a linear function of explanatory variables, which can be continuous or classes, and can be transformed variables, or powers or polynomials

$$\eta = b_0 + b_1x_1 + b_2x_2 + \dots + b_px_p$$

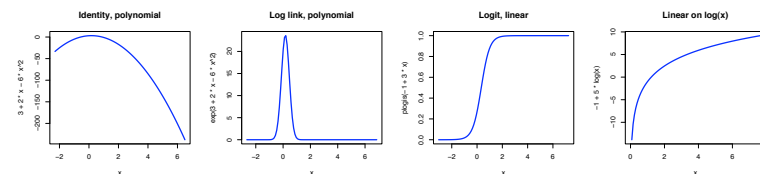
2. **Link function** $g(\cdot)$ that transforms the fitted values μ to the linear predictor η

$$g(\mu) = \eta$$

3. **Error distribution** from the exponential family to describe the distribution of residuals about fitted values.

Special cases of GLM

Model	Link	Error	Variance
Linear model	Identity $\mu = \eta$	Normal	Constant
Log-linear	Logarithmic	Poisson	μ
Logistic	Logistic	Binomial	$\mu(1 - \pi)$



Ecologically meaningful error distributions

Normal error rarely adequate in ecology, but GLM offer ecologically meaningful alternatives.

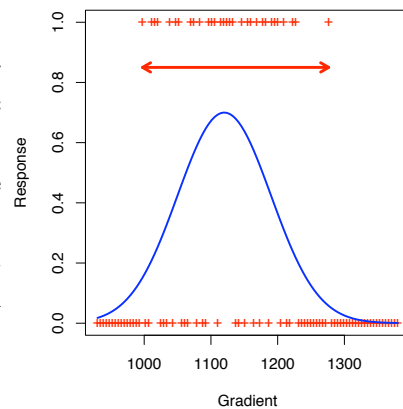
- **Poisson.** Counts: integers, non-negative, variance increases with mean.
- **Binomial.** Observed proportions from a total: integers, non-negative, have a maximum value, variance largest at $\pi = 0.5$
- **Gamma.** Concentrations: non-negative real values, standard deviation increases with mean, many near-zero values and some high peaks.

Goodness of fit and inference

- Deviance: Measure of goodness of fit
 - Derived from the error function: Residual sum of squares in Normal error
 - Distributed approximately like χ^2
- Residual degrees of freedom: Each fitted parameter consumes one degree of freedom and (probably) reduces the deviance.
- Inference: Compare change in deviance against change in degrees of freedom
- Overdispersion: Deviance larger than expected under strict likelihood model
- Use F -statistic in place of χ^2 .

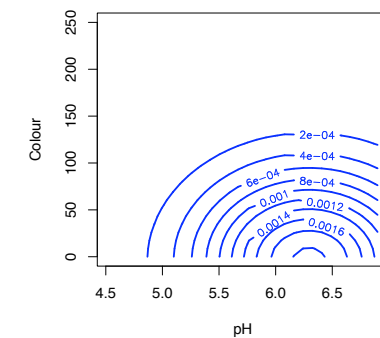
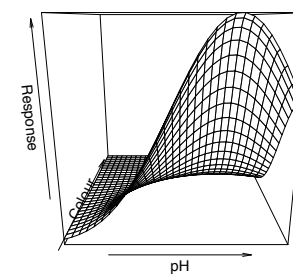
Gaussian model and response range

- Gaussian response is never exactly zero: Asymptotic model
- Observed abundances have a discrete component
- The observed range depends on parameters t and h



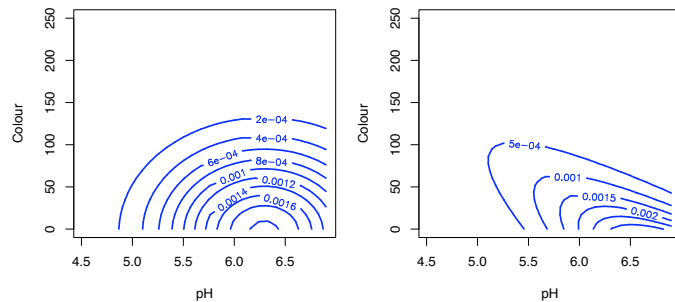
Several gradients

- Gaussian response can be fitted to several gradients: Bell



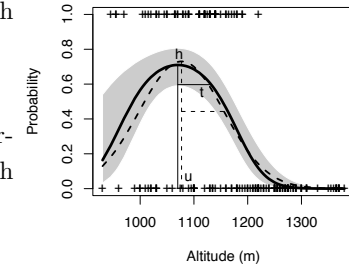
Interactions in Gaussian responses

- No interactions: Responses parallel to the gradients
- Interactions: The optimum on one gradient depends on the other



Logistic Gaussian response

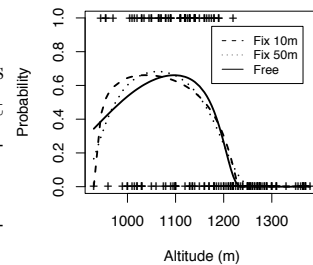
- Polynomial often used with other link functions than log.
- Binomial error: logistic link.
- The Gaussian parameters correct only with log link: Width t has different interpretation.



Beta response

- Responses with varying skewness and kurtosis.
- Simulated coenoclines to test robustness of ordination.
- Commonly fitted fixing endpoints p_1 and p_2 and using GLM: Not flexible any longer, but greatly influenced by endpoints.
- **Must** be fitted with non-linear regression.

$$\mu = k(x - p_1)^\alpha (p_2 - x)^\gamma$$



Parameters of Beta response

- No clearly interpreted parameters.
- α and γ define:
 1. The location of the mode.
 2. The skewness of the response.
 3. The kurtosis of the response.
- Response is zero at p_1 and p_2 : absolute endpoints of the range.
- k is a scaling parameter: height depends on other parameters as well.

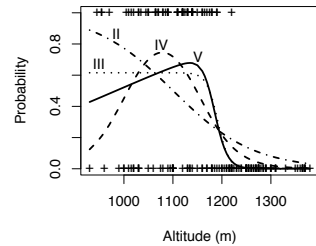
$$\mu = k(x - p_1)^\alpha (p_2 - x)^\gamma$$

HOF models

Huisman–Olf–Fresco: A set of five hierarchic models with different shapes.

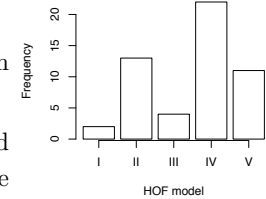
$$\mu = \frac{M}{[1+\exp(a+bx)] \cdot [1+\exp(c-dx)]}$$

Model	Parameters
V Skewed	\underline{a} \underline{b} \underline{c} \underline{d}
IV Symmetric	\underline{a} \underline{b} \underline{c} b
III Plateau	\underline{a} \underline{b} \underline{c} ∞
II Monotone	\underline{a} \underline{b} 0 0
I Flat	\underline{a} 0 0 0



HOF: Inference on response shape

- Alternative models differ only in response shape.
- Selection of parsimonious model with statistical criteria.
- ‘Shape’ is a parametric concept, and parametric HOF models may be the best way of analysing differences in response shapes.

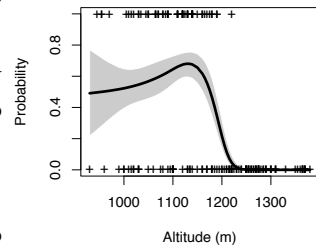


Most parsimonious HOF models on Altitude gradient in Mt. Field, Tasmania.

Generalized Additive Models (GAM)

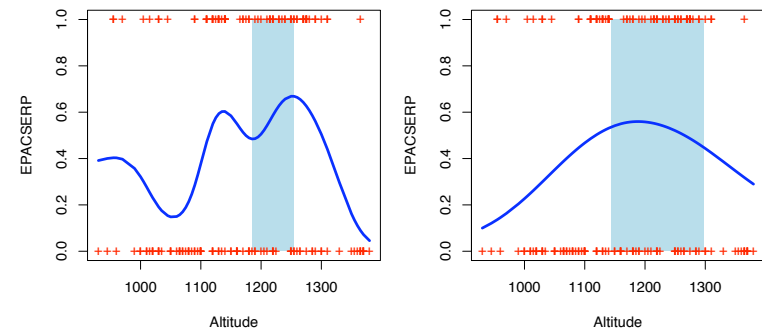
- Generalized from GLM: linear predictor replaced with smooth predictor.
- Smoothing by regression splines or other smoothers.
- Degree of smoothing controlled by degrees of freedom: analogous to number of parameters in GLM.
- Everything else like in GLM.
- Enormous use in ecology — also outside gradient modelling.

$$g(\mu) = \text{smooth}(x)$$



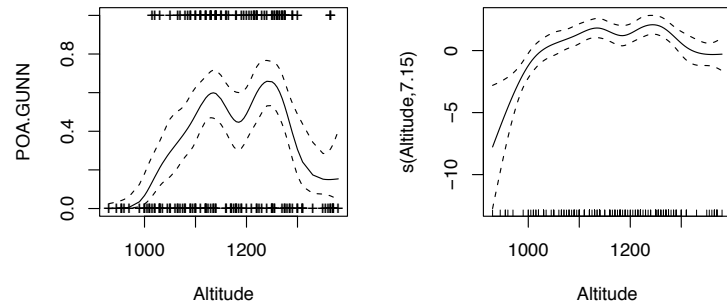
Degrees of Freedom

The width of a smoothing window = Degrees of Freedom



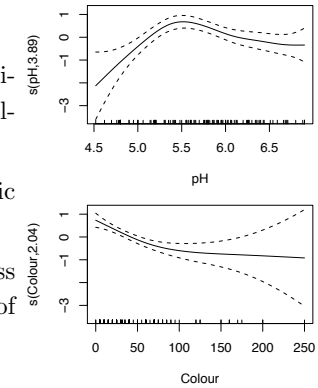
Linear scale and response scale

GAM is smooth in the link scale, but the user prefers the response



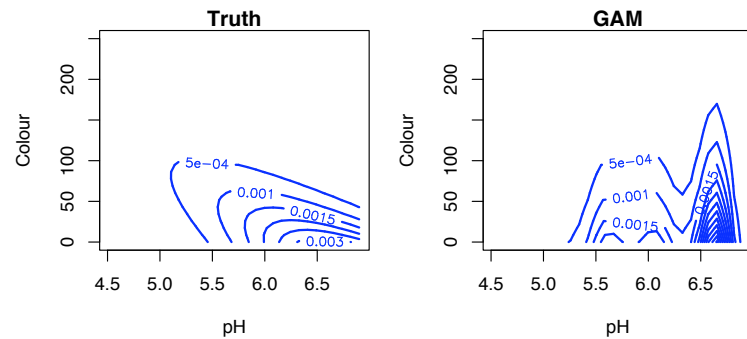
Multiple gradients

- Each gradient is fitted separately
- Interpretation easy: Only the individual main effects shown and analysed
- Possible to select good parametric shapes
- Thin-plate splines: Same smoothness in all directions and no attempt of making responses parallel to axes



Interactions

GAM are designed to show the main effects beautifully in panel plots
'Equivalent kernel' is parallel to the axes



Diversity and spatial scale

- Whittaker suggested several concepts of diversity
 - α : Diversity on a sample plot, or 'point' diversity.
 - β : Diversity along ecological gradients.
 - γ : Diversity among parallel gradients or classes of environmental variables.
 - δ : The total diversity of a landscape: sum of all previous.

Many faces of beta diversity

What are we talking about when we are talking about beta diversity?

1. General heterogeneity of a community.
2. Decay of similarity with gradient separation.
3. Widths of species responses along gradients.
4. Rate of change in community composition along gradients.

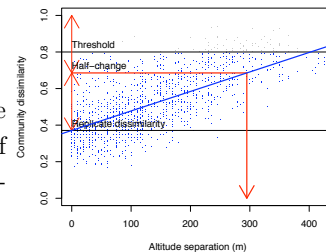
General heterogeneity

- “Whittaker’s index”: Proportion of average species richness on a single plot \bar{S} and the total species richness in all plots S_{TOT} .
- Total richness increases with increasing sample size.
- Average richness stabilizes with increasing sampling effort.
- \bar{S}/S_{TOT} decreases with sample size.
- No reference to gradients: even with a single location, replicate sampling decreases the index.
- Pattern diversity: Within site diversity.

Similarity decay with gradient separation

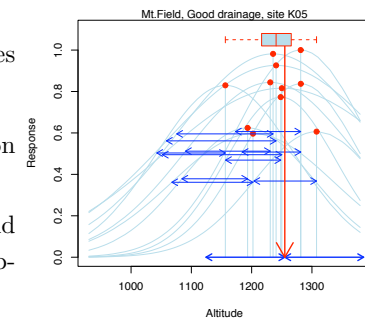
- Intercept: (Dis)similarity at zero-distance – ‘noise’, replicate (dis)similarity, general heterogeneity or pattern diversity.
- Slope: Beta diversity.
- Half-change: Gradient distance where expected similarity is half of the replicate similarity (intercept).

Plot community (dis)similarity against gradient separation and fit a linear regression.



Hill indices of beta diversity

1. Average width of species responses.
 2. Variance of optima of species occurring in one site.
- Used with scaling of ordination axes.
 - The first index discussed and described, but the second applied.
 - Equal only to degenerated species packing gradients.

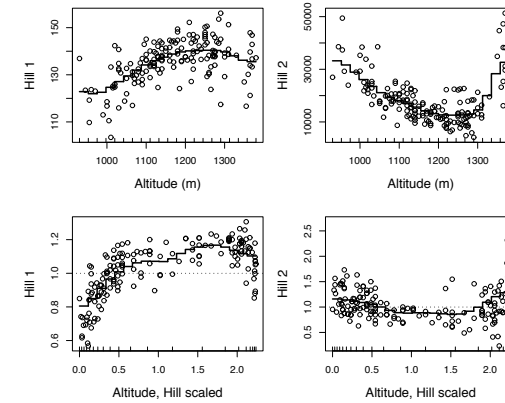


Gaussian responses fitted to species occurring in one site.

Hill scaling in practice

- Hill index spaced on species occurrences in sites: random variation.
- Smoothed by segments.
- Each segment made equally long in terms of the Hill index — almost. . .
- Four cycles commonly performed, but not enough to stabilize the Hill index (with half steps taken).

Hill rescaling of gradients

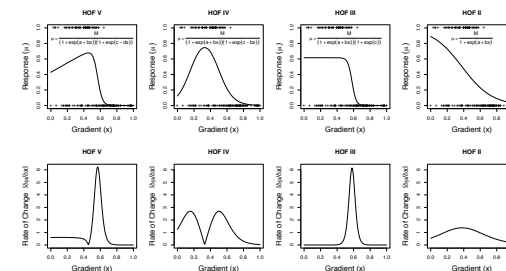


Are there species in common at '4sd' distance?

- Confounds Normal probability density and Gaussian response: Density had 95 % of its survice at $\mu \pm 2\sigma$, but the height of the response is $0.135h$
- The range of species depends on h , but in many cases a more realistic limit is $u \pm 3t$, where $\mu = 0.01h$
- If widths t vary, some species occur at longer distances.
- Look at your data before saying that there are no species in common at 4 'sd'.

Rate of change along gradients

Instantaneous rate of change δ at any gradient point x estimated from fitted species response functions μ :



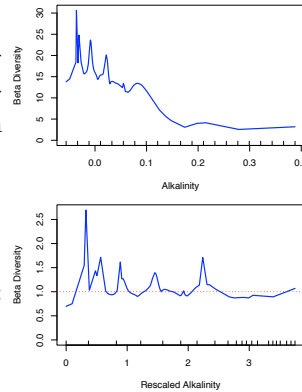
$$\delta(x) = \sum_{j=1}^S \left| \frac{\partial}{\partial x} \mu_j(x) \right|$$

Rescaling to constant rate of change

- Make interval between any two gradient points a and b equal to the total accumulated change Δ_{ab} between points:

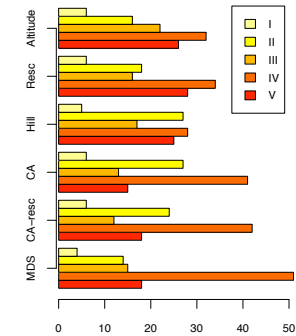
$$\Delta_{ab} = \int_a^b \delta(x) \cdot dx$$

- Can be based on any response model: The example uses HOF.



Alternative rescaling and response shapes

- Direct rescaling and Hill rescaling are inconsistent.
- Two Hill indices of beta diversity are inconsistent.
- None of the rescaling methods produce symmetric response shapes.
- Ordination axes tend to produce symmetric responses.



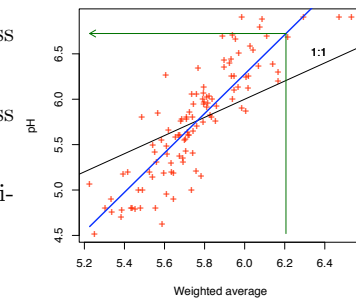
Weighted averages in bioindication

$$\tilde{x}_i = \frac{\sum_{j=1}^S y_{ij} u_j}{\sum_{j=1}^S y_{ij}}$$

- Weighted average of indicator values of species occurring in a site.
- Can use species weighted averages \tilde{u}_j or other indicator values u_j .
- Repeated cycling $x \rightarrow \tilde{u}, \tilde{u} \rightarrow \tilde{x}, \dots, \tilde{x} \rightarrow \tilde{u}$ gives a solution of first axis in correspondence analysis.
- The range and variance of weighted averages is smaller than the range of values they are based on: “deshrinking” to restore the original variance.

Deshrinking: stretch weighted averages

- Inverse regression: regress gradient values on WAs.
- Classical regression: regress WAs on gradient values.
- Simple stretching: make variances equal.

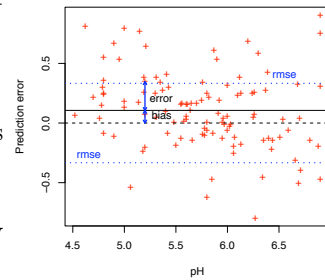


Goodness of prediction: Bias and error

Goodness: prediction error.

Correlation bad: depends on the range of observations.

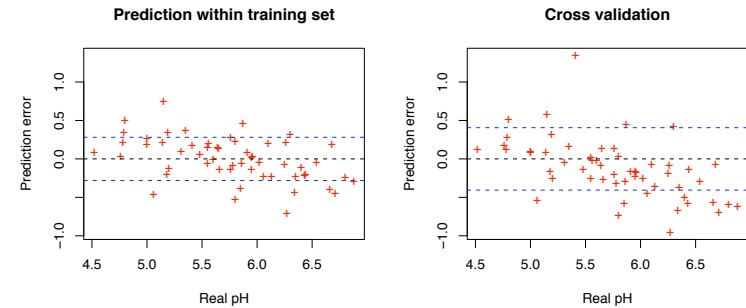
- Root mean squared error $\epsilon = \sqrt{\sum_{i=1}^N (\tilde{x}_i - x_i)^2 / N}$.
- Bias b : systematic difference.
- Error ε : random error about bias
- $\epsilon^2 = b^2 + \varepsilon^2$



Must be cross-validated or badly biased

Cross validation

Leave-one-out ('jackknife'), each in turn, or divide data into training and test data sets.



Bioindication: Likelihood approach

- **Likelihood** is the probability of a given *observed* value with a certain *expected* value
- Maximum likelihood estimation: Expected values that give the best likelihood for observations.
 - ML estimates are close to observed values, and the proximity is measured with the likelihood function
 - Commonly we use the negative logarithm of the likelihood, since combined probabilities may be very small

Regression and Bioindication

- **Regression:**
 - We know the gradient values x and observed species abundances y
 - We find the most likely expected values $\hat{\mu}$ for species
- **Bioindication:**
 - We know the observed species abundances y
 - We have a gradient model that gives the expected abundances μ for any gradient value x
 - We find the most likely gradient values \hat{x} that maximize the likelihood of observing y when expecting μ
- ML Bioindication can be used with many response models and with many gradients

Finding elevation from species composition

