Predicting community composition using site and species characteristics: a data management perspective

Steve C. Walker, Guillaume Guénard, and Pierre Legendre

Université de Montréal
Département de Sciences Biologiques

August 12, 2011
Ecological Society of America
Austin, Texas

Introduction

Traits and the ecology of communities

Illustrative example

Complex example

Data management

Methods

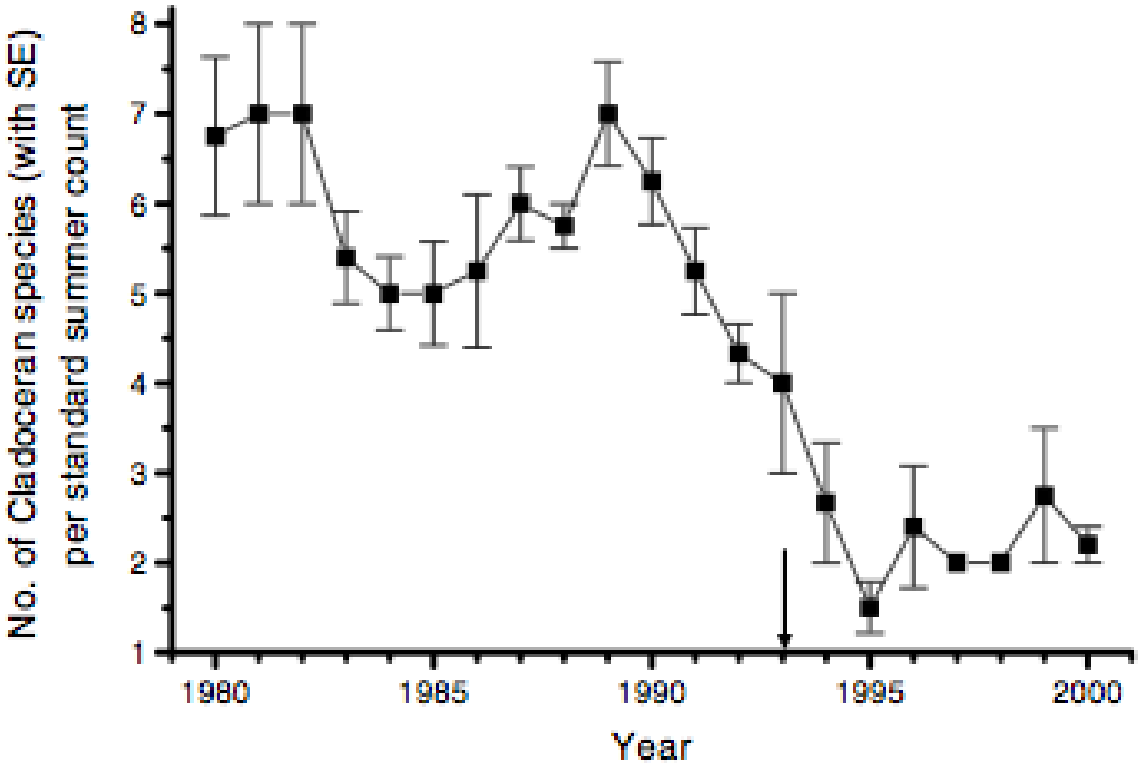
Results

Bythotrephes longimanus

Wisconsin Department of Natural Resources

Steve C. Walker
Guillaume Guénard
Pierre Legendre

Bythotrephes longimanus



Yan et al. (2002)

Steve C. Walker
Guillaume Guénard
Pierre Legendre

Fourth corner

	sp 1	sp 2	sp 3	sp 4
site 1	0.1	2.1	0.1	1.5
site 2	0.7	-0.9	1.8	3.7
site 3	1.1	0.5	1.5	2.8
site 4	1.3	-2.0	3.0	-0.2
site 5	1.7	2.0	1.3	1.2
site 6	0.8	-0.1	2.0	1.1
site 7	-2.6	-1.4	1.8	4.1
site 8	-0.0	1.5	2.3	2.3

Steve C. Walker
Guillaume Guénard
Pierre Legendre

Fourth corner

	sp 1	sp 2	sp 3	sp 4	environment
site 1	0.1	2.1	0.1	1.5	-0.3
site 2	0.7	-0.9	1.8	3.7	1.4
site 3	1.1	0.5	1.5	2.8	-0.1
site 4	1.3	-2.0	3.0	-0.2	0.4
site 5	1.7	2.0	1.3	1.2	-0.3
site 6	0.8	-0.1	2.0	1.1	-0.6
site 7	-2.6	-1.4	1.8	4.1	2.0
site 8	-0.0	1.5	2.3	2.3	0.7

Steve C. Walker
Guillaume Guénard
Pierre Legendre

Fourth corner

	sp 1	sp 2	sp 3	sp 4	environment
site 1	0.1	2.1	0.1	1.5	-0.3
site 2	0.7	-0.9	1.8	3.7	1.4
site 3	1.1	0.5	1.5	2.8	-0.1
site 4	1.3	-2.0	3.0	-0.2	0.4
site 5	1.7	2.0	1.3	1.2	-0.3
site 6	0.8	-0.1	2.0	1.1	-0.6
site 7	-2.6	-1.4	1.8	4.1	2.0
site 8	-0.0	1.5	2.3	2.3	0.7
trait	-1.0	-1.0	1.0	1.0	

Fourth corner

	sp 1	sp 2	sp 3	sp 4	environment
site 1	0.1	2.1	0.1	1.5	-0.3
site 2	0.7	-0.9	1.8	3.7	1.4
site 3	1.1	0.5	1.5	2.8	-0.1
site 4	1.3	-2.0	3.0	-0.2	0.4
site 5	1.7	2.0	1.3	1.2	-0.3
site 6	0.8	-0.1	2.0	1.1	-0.6
site 7	-2.6	-1.4	1.8	4.1	2.0
site 8	-0.0	1.5	2.3	2.3	0.7
trait	-1.0	-1.0	1.0	1.0	→ ?? ←

Steve C. Walker
Guillaume Guénard
Pierre Legendre

Introduction

- Traits and the ecology of communities
- Illustrative example
- Complex example
- Data management

Methods

Results

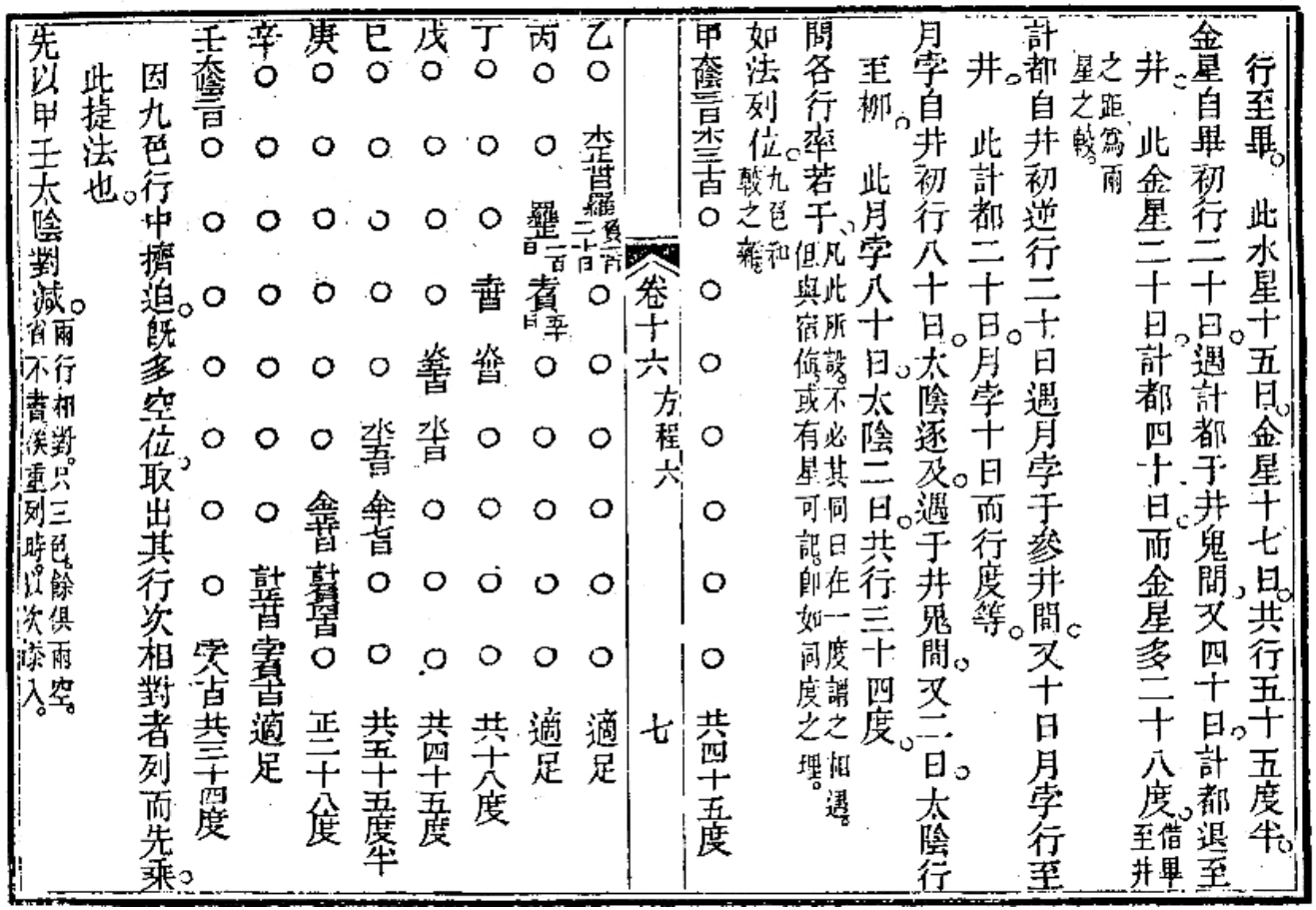
Introduction

- Traits and the ecology of communities
- Illustrative example
- Complex example
- Data management

Methods

Results

Linear algebra as data management



行至畢。此水星十五日。金星十七日。共行五十五度半。
金星自畢初行二十日。遇計都于井鬼間。又四十日。計都退至
井。此金星二十日。計都四十日。而金星多二十八度。（借畢至井之距爲兩星之較。）
計都自井初逆行二十日遇月孛于參井間。又十日月孛行至
井。此計都二十日。月孛十日而行度等。
月孛自井初行八十日。太陰逐及。遇于井鬼間。又二日。太陰行
至柳。此月孛八十日。太陰二日。共行三十四度。
問各行率若干。（凡此所設。不必其同日在一度謂之相遇。但與宿值。或有星可記。卽如同度之理。）
如法列位。（九色和較之雜。）

甲叅言曰三百〇　〇　〇　〇　〇　共四十五度

卷十六　方程六　七

乙〇　曰百羅負二百　〇　〇　〇　〇　適足
丙〇　〇　羅一百　曰貴辛　〇　〇　〇　適足
丁〇　〇　〇　曾　曾　〇　〇　〇　共十八度
戊〇　〇　〇　〇　叅　曾　〇　〇　〇　共四十五度
己〇　〇　〇　〇　〇　曾　叅　曾　〇　共五十五度半
庚〇　〇　〇　〇　〇　〇　金　曾　甘　曾　〇　正二十八度
辛〇　〇　〇　〇　〇　〇　〇　甘　曾　實　曾　適足
壬叅言曰〇　〇　〇　〇　〇　〇　〇　〇　〇　共三十四度

因九色行中擣迫。既多空位。取出其行次相對者列而先乘。
此提法也。
先以甲壬太陰對減。（兩行相對。只三色。餘俱兩空。省不書。後重列時。位次添入。）

Ancient Chinese text (∼150 BCE)

Linear algebra as data management

$$\begin{bmatrix} 3 & 30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 45 \\ 0 & 30 & -120 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 100 & -50 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 30 & 30 & 0 & 0 & 0 & 0 & 18 \\ 0 & 0 & 0 & 0 & 50 & 10 & 0 & 0 & 0 & 45 \\ 0 & 0 & 0 & 0 & 0 & 15 & 17 & 0 & 0 & \frac{111}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 20 & -40 & 0 & 28 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 20 & -10 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 80 & 34 \end{bmatrix}$$

Hart (2009)

Steve C. Walker
Guillaume Guénard
Pierre Legendre

Linear algebra as data management

$$\begin{bmatrix} k_1 & l_1 & 0 & \cdots & 0 & b_1 \\ 0 & k_2 & l_2 & \ddots & \vdots & b_2 \\ \vdots & \ddots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & k_{N-1} & l_{N-1} & b_{N-1} \\ l_N & 0 & \cdots & 0 & k_N & b_N \end{bmatrix}$$

Hart (2009)

Steve C. Walker
Guillaume Guénard
Pierre Legendre

Linear algebra as data management

$$\mathbf{Y} = \mathbf{XB} \tag{1}$$

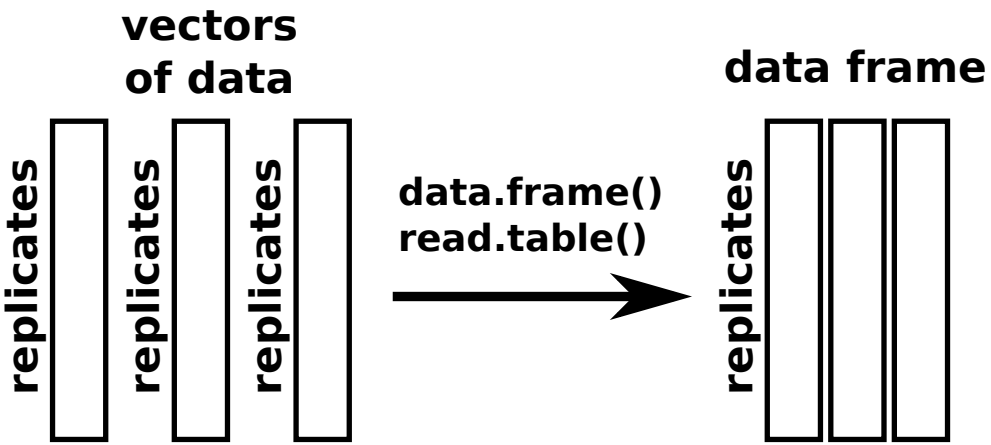
Linear algebra as data management

$$\begin{aligned} \mathbf{Y} &= \mathbf{XB} \\ \mathbf{X}^{\mathsf{T}}\mathbf{Y} &= \mathbf{X}^{\mathsf{T}}\mathbf{XB} \end{aligned} \tag{1}$$

Linear algebra as data management

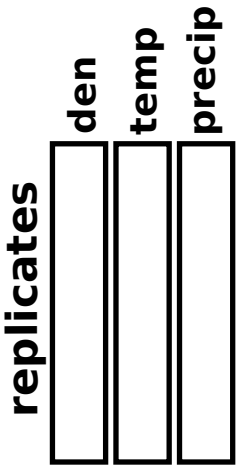
$$\begin{aligned} \mathbf{Y} &= \mathbf{XB} \\ \mathbf{X}^{\mathsf{T}}\mathbf{Y} &= \mathbf{X}^{\mathsf{T}}\mathbf{XB} \\ (\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}\mathbf{Y} &= \mathbf{B} \end{aligned} \tag{1}$$

The R paradigm of data management



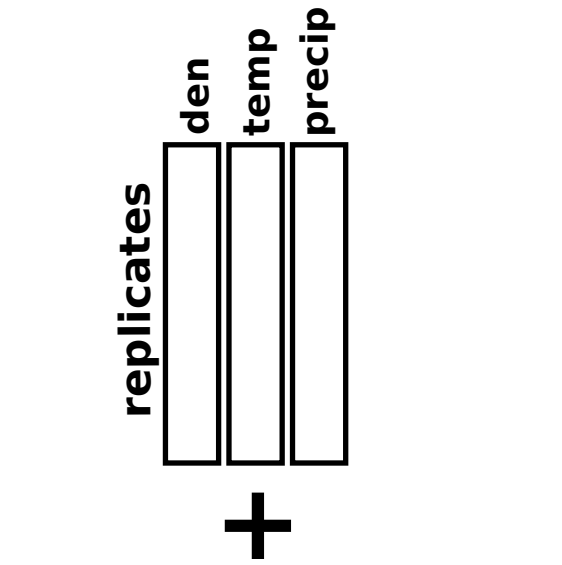
Steve C. Walker
Guillaume Guénard
Pierre Legendre

The R paradigm of data management



Steve C. Walker
Guillaume Guénard
Pierre Legendre

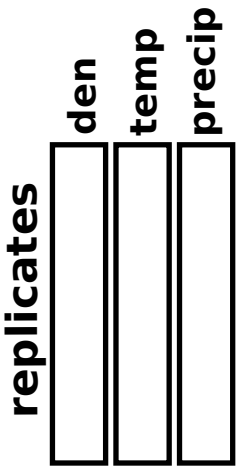
The R paradigm of data management



den ~ temp + precip

Steve C. Walker
Guillaume Guénard
Pierre Legendre

The R paradigm of data management



+

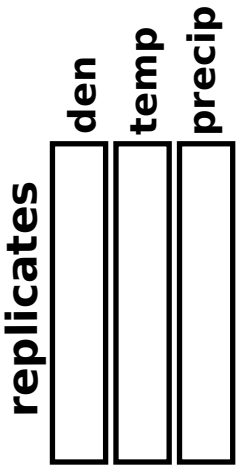
den ~ temp + precip

+

lm / glmer / plot / xyplot

Steve C. Walker
Guillaume Guénard
Pierre Legendre

The R paradigm of data management



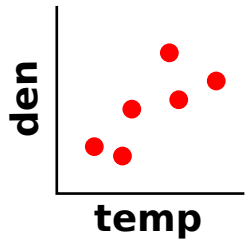
+

den ~ temp + precip

+

lm / glmer / plot / xyplot

=



p < 0.0001

	coef	s.e.
(intcpt)	-1.2	0.4
temp	2.1	0.1
precip	-0.1	0.1

The R paradigm of data management

$$\text{DATA FRAME} + \text{FORMULA} + \text{FUNCTION} = \text{ANALYSIS}$$

Steve C. Walker
Guillaume Guénard
Pierre Legendre

The R paradigm of data management

DATA FRAME + FORMULA + FUNCTION = ANALYSIS

data management

The R paradigm of data management

Steve C. Walker
Guillaume Guénard
Pierre Legendre

$$\text{DATA FRAME} + \underset{\text{ecological ideas}}{\textbf{FORMULA}} + \text{FUNCTION} = \text{ANALYSIS}$$

The R paradigm of data management

$$\text{DATA FRAME} + \text{FORMULA} + \underset{\text{algorithms}}{\textbf{FUNCTION}} = \text{ANALYSIS}$$

Steve C. Walker
Guillaume Guénard
Pierre Legendre

The R paradigm of data management

DATA FRAME + FORMULA + FUNCTION = **ANALYSIS**

Steve C. Walker
Guillaume Guénard
Pierre Legendre

Acknowledgements

- Natural Sciences and Engineering Research Council of Canada
- Laura Timms (McGill University)
- Ben Bolker (McMaster University)
- The many people who gave their time to develop free software:  and LaTeX