Species distributions: uncertainty and modelling

Uncertainty in species distributions

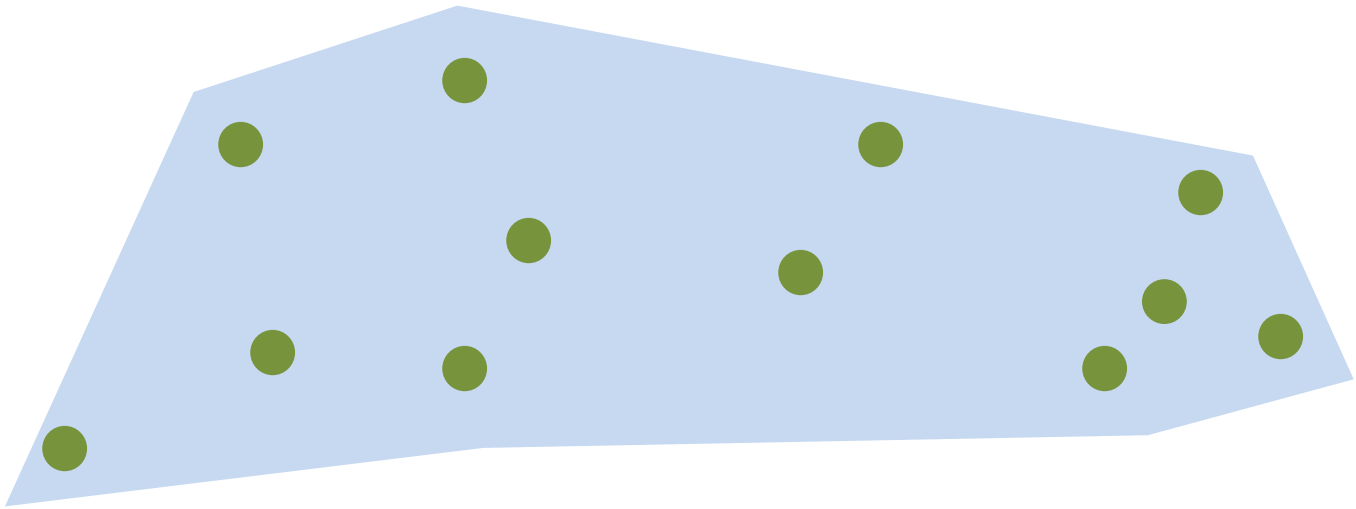
- when is an absence really an absence?
- … and when is a presence really a presence?

Uncertainty in species distributions

- species distributions are dynamic
- occurrence data are just snapshots
- can’t tell where everyone will be at a given moment

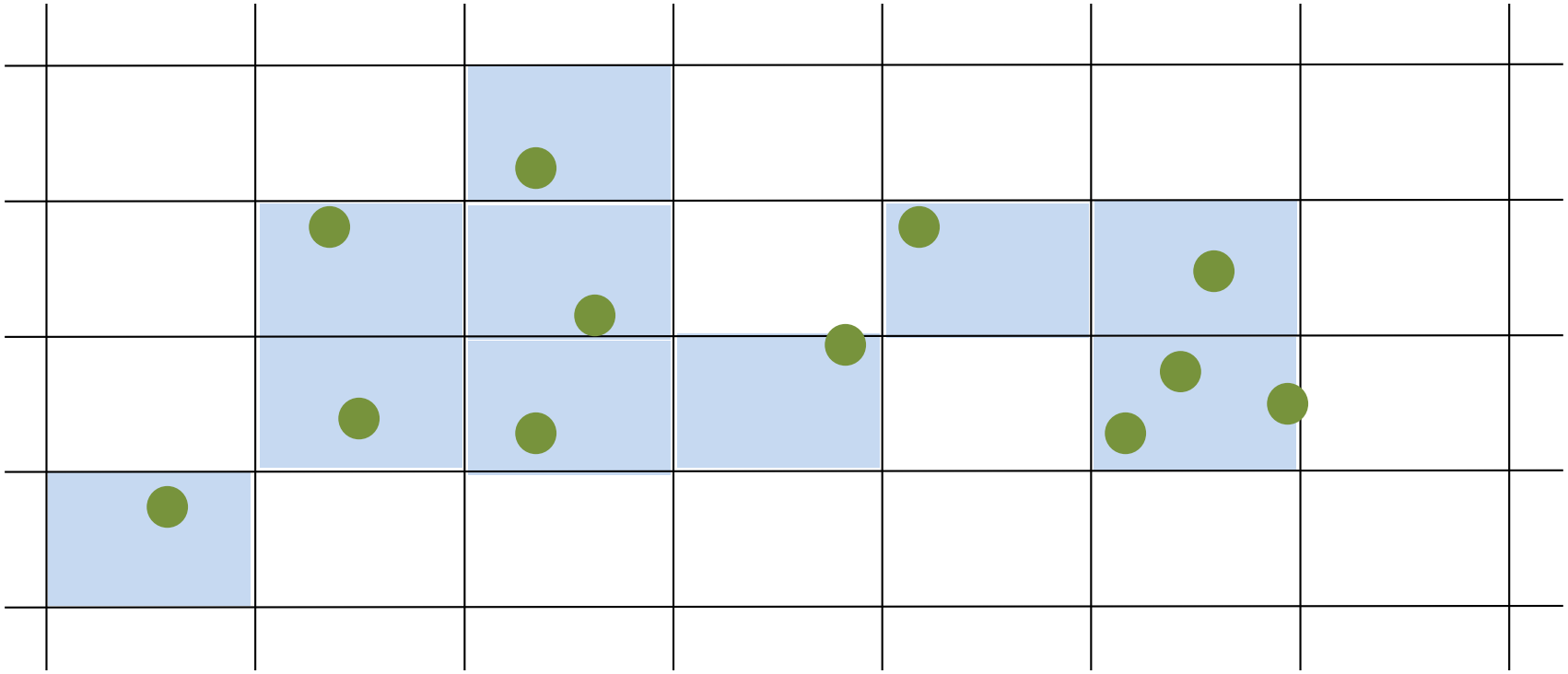
Uncertainty in species distributions

- species distributions are dynamic
- occurrence data are just snapshots
- can’t tell where everyone will be at a given moment



Uncertainty in species distributions

- species distributions are dynamic
- occurrence data are just snapshots
- can’t tell where everyone will be at a given moment

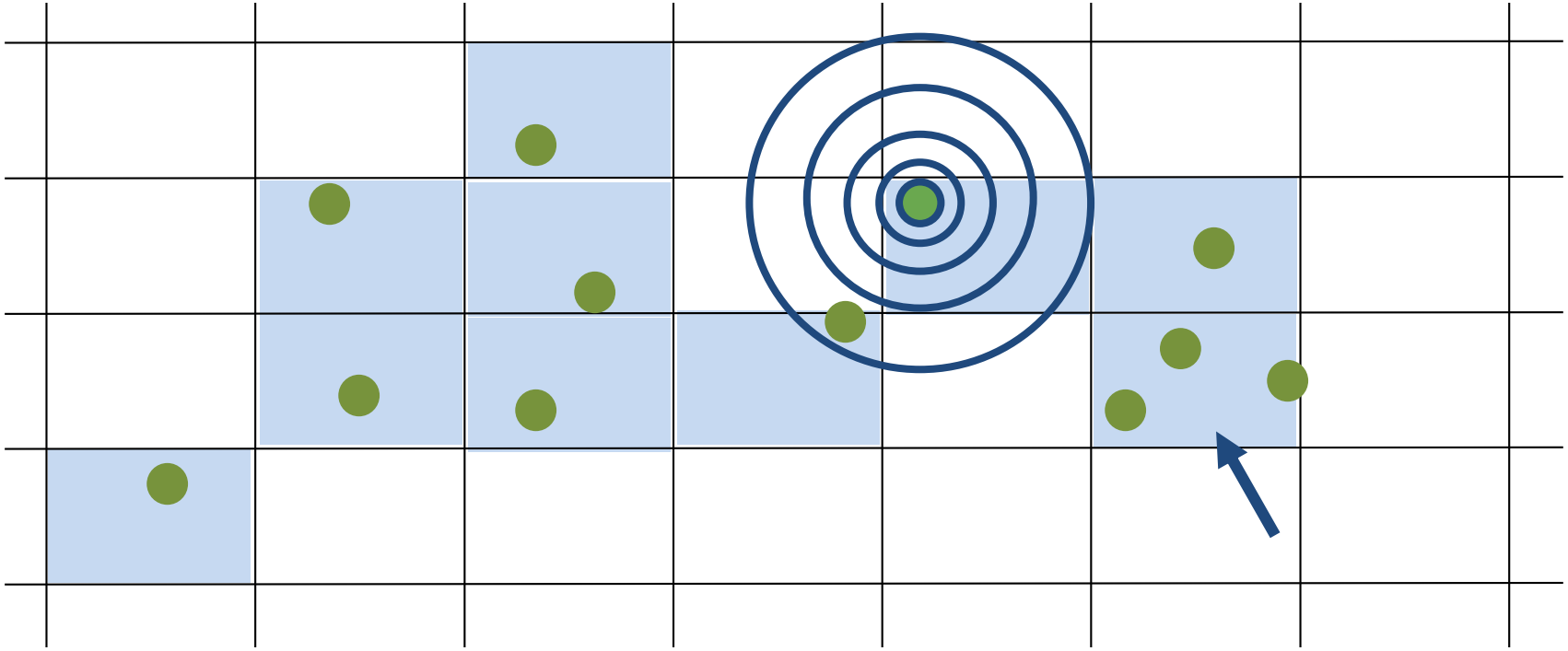


Uncertainty in species distributions

- when is an absence really an absence?
- … and when is a presence really a presence?

  will it still be there when you return?

  and is it equivalent everywhere?

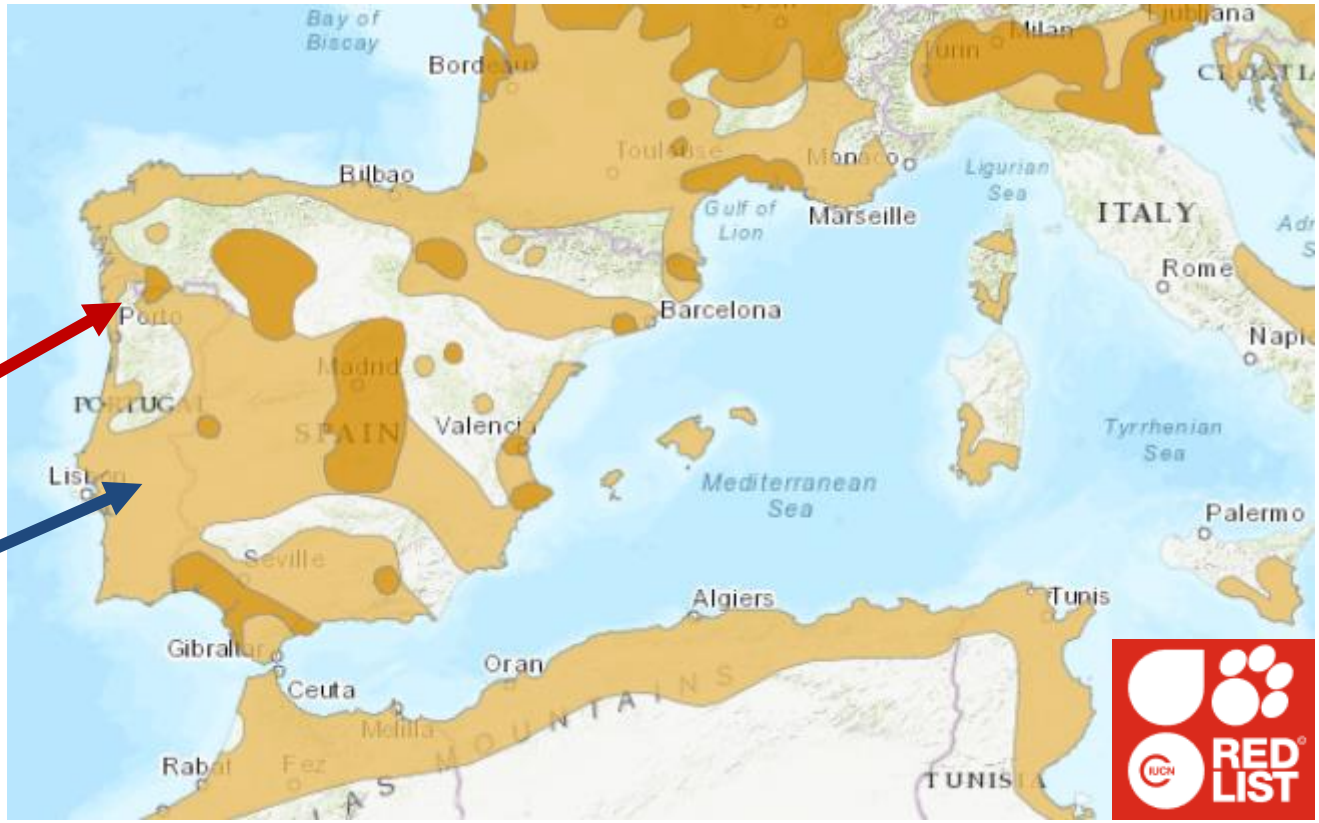


Uncertainty in species distributions

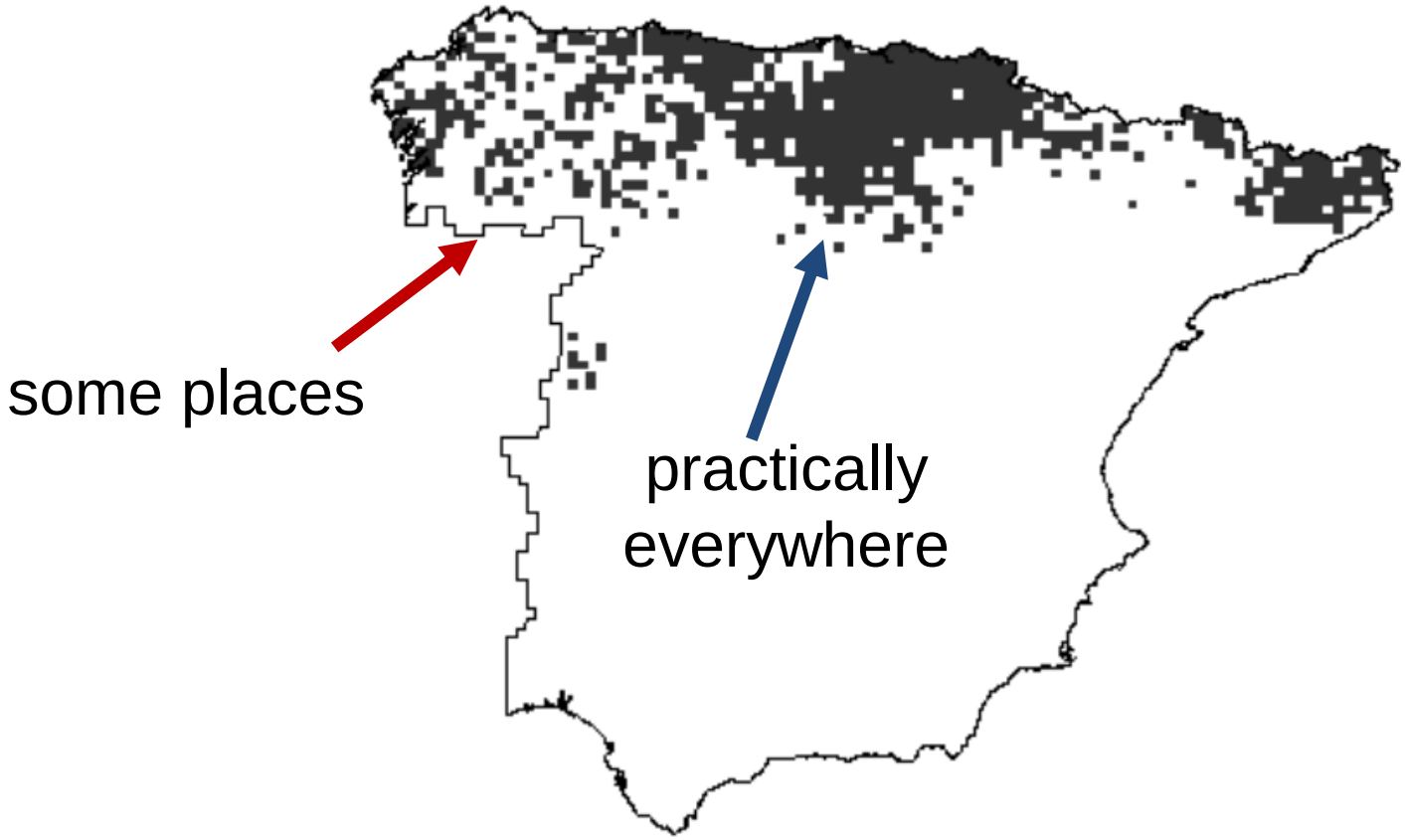
Northern lapwing (*Vanellus vanellus*)

some winters

every winter



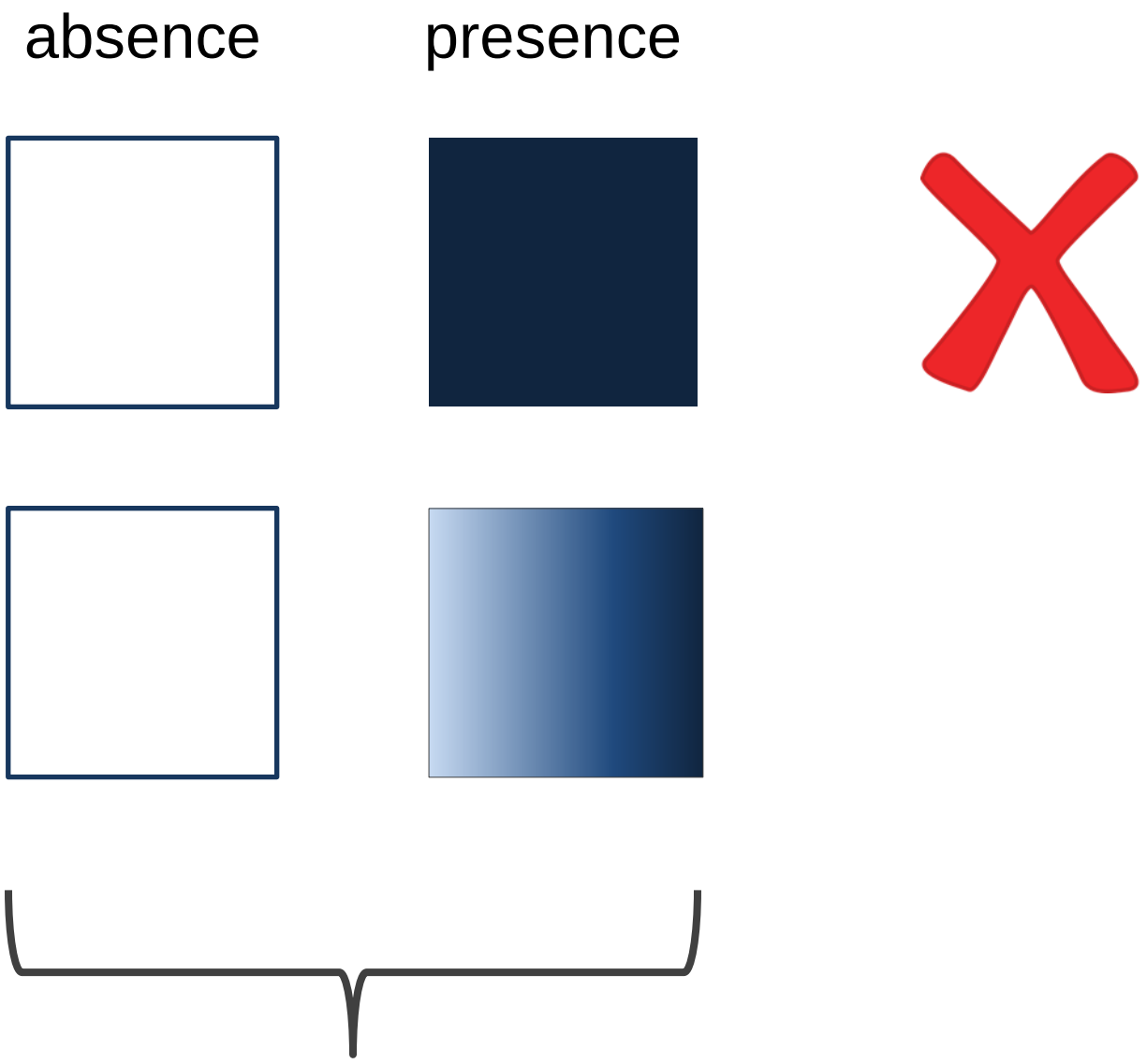
Uncertainty in species distributions



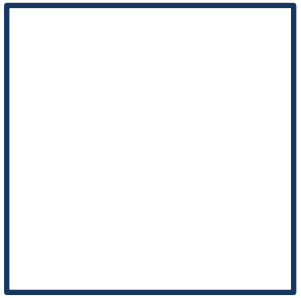
imperfect detection: it's OK!

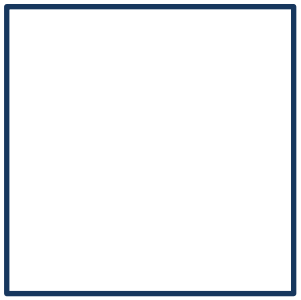
as long as it reflects **lower occurrence frequency**

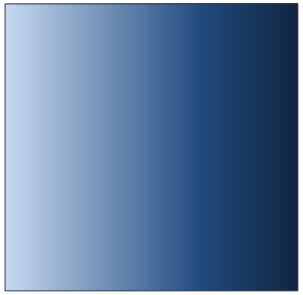
rather than **bias** in survey effort

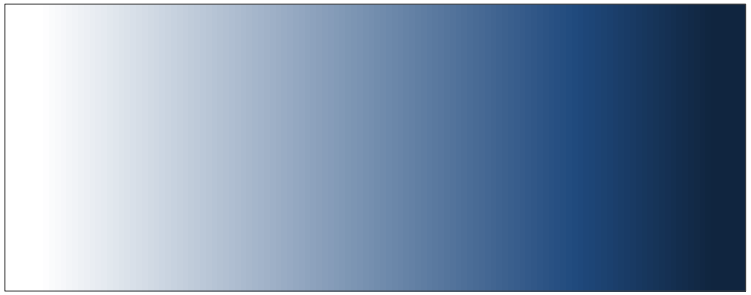

absence
presence
2 classes?!

absence presence





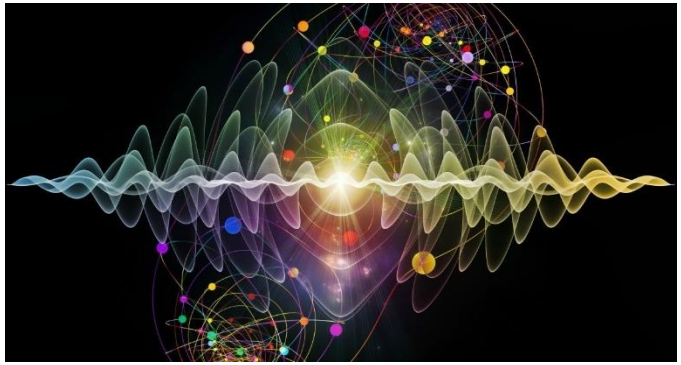




QUANTUM PHYSICS

uncertainty principle: we cannot know where a specific particle is at any given time; we can only know the **probability** of finding that particle at a specific point

QUANTUM PHYSICS

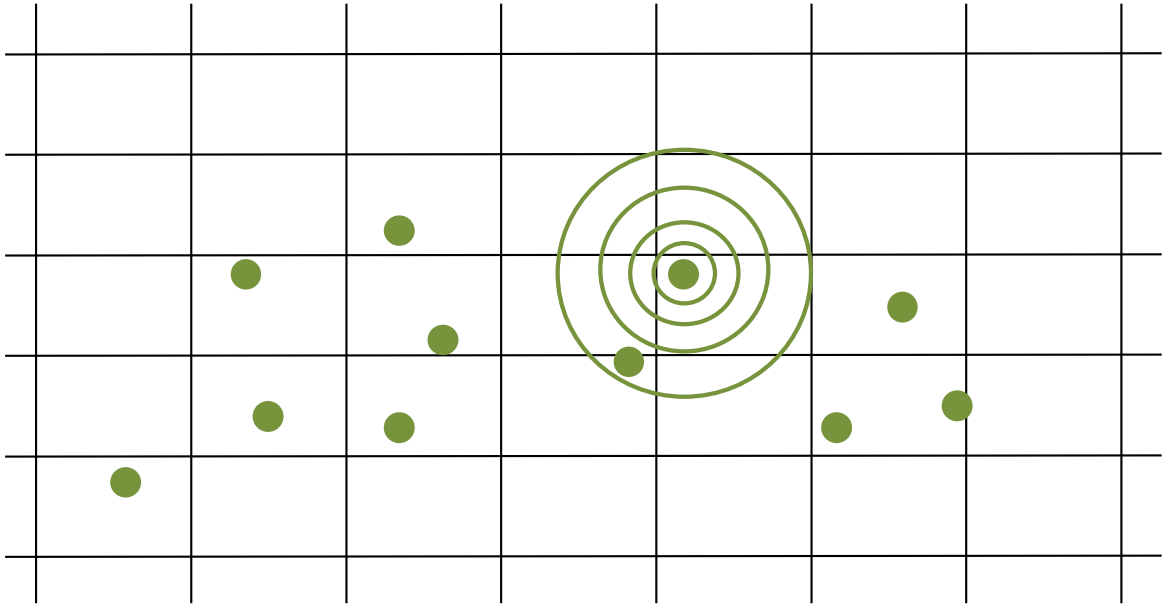


the **complete information** about a particle's location is not where it is, but the wavefunction that describes everywhere it could be and **how likely** it is to be there

just an analogy!

QUANTUM PHYSICS

SPECIES DISTRIBUTIONS



current / momentary location impossible to predict

QUANTUM PHYSICS

A particle can simultaneously be here and over there. This multiple personality is described mathematically by the particle's wavefunction, which gives the probability that it is in each of those states.

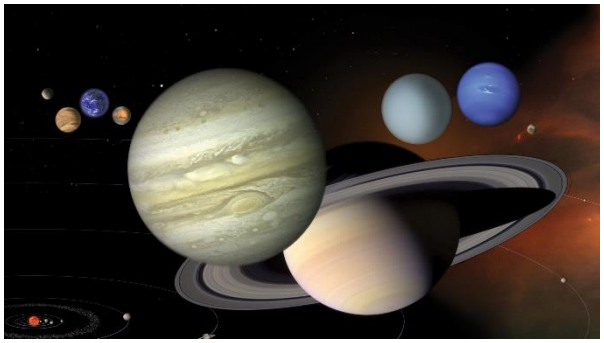
Only when the particle's properties are measured does the wavefunction collapse, choosing a definite state in a single location. Crucially, there is no way, even in principle, to predict the result of a single experiment; the probabilities show up only as a statistical distribution and only when the experiment is repeated many times.

[Merali 2012, *Nature* 492: 22-25]

QUANTUM PHYSICS & SPECIES DISTRIBUTIONS

A **species** *can simultaneously be here and over there. This multiple personality is described mathematically by the* **distribution model**, *which gives the probability that it is in each of those* **sites**.

Only when the **species' position is** *measured does the* **distribution** *collapse, choosing a definite state in a single location. Crucially, there is no way, even in principle, to predict the result of a single* **observation**; *the probabilities show up only as a statistical distribution and only when the* **observation** *is repeated many times.*

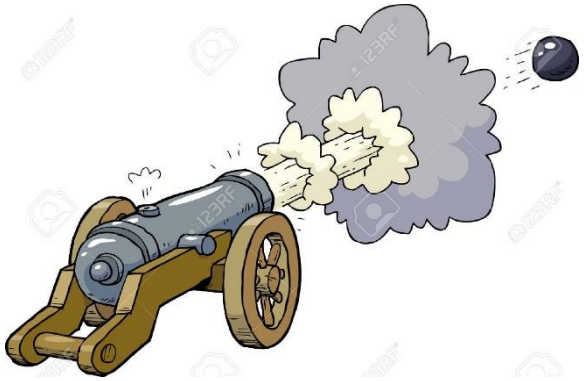


determinism

current / momentary location impossible to predict

live beings have some degree of freedom

response to environment not entirely deterministic



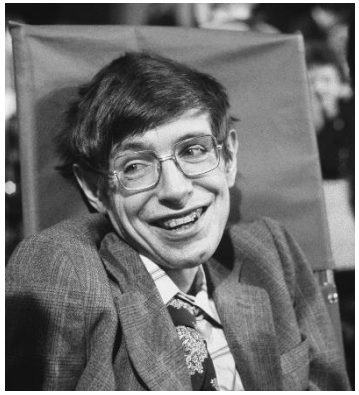
determinism

current / momentary location impossible to predict

live beings have some degree of freedom

response to environment **not entirely deterministic**

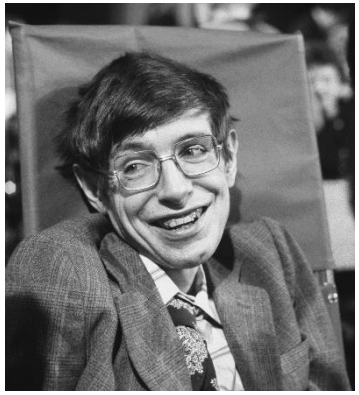
Quantum physics might seem to undermine the idea that nature is governed by laws, but that is not the case. Instead it leads us to accept a ***new form of determinism****: Given the state of a system at some time, the laws of nature determine the* ***probabilities of various futures and pasts*** *rather than determining the future and past with certainty.*



-- Stephen Hawking, *The Grand Design*

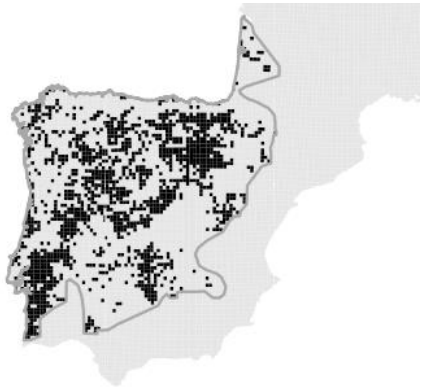
*Despite the **probabilistic nature** of quantum mechanical predictions, its claims **can be rigorously tested** […] Quantum theory is still physics: a rigorous science that gives rise to quantitative predictions that can be verified or falsified by experiment.*

-- Stephen Hawking, *The Grand Design*



Species distribution models

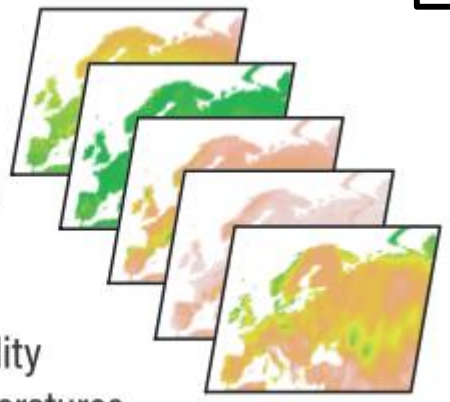
recorded distribution



+

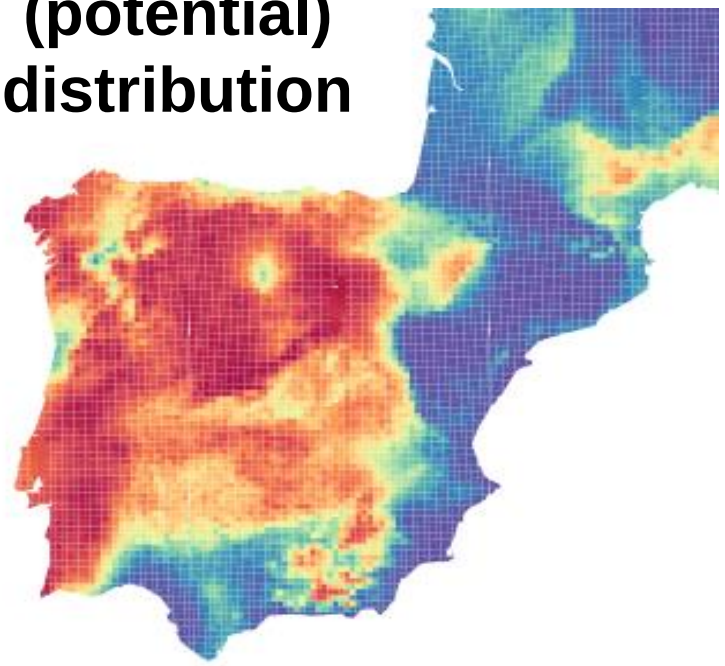
predictor variables

Elevation
Slope
Aspect
Soil properties
Wind speed
Cloud cover
Relative humidity
Max/min temperatures



model formula

predicted (potential) distribution



convert to binary prediction?!

Species distribution models

wavefunction

~

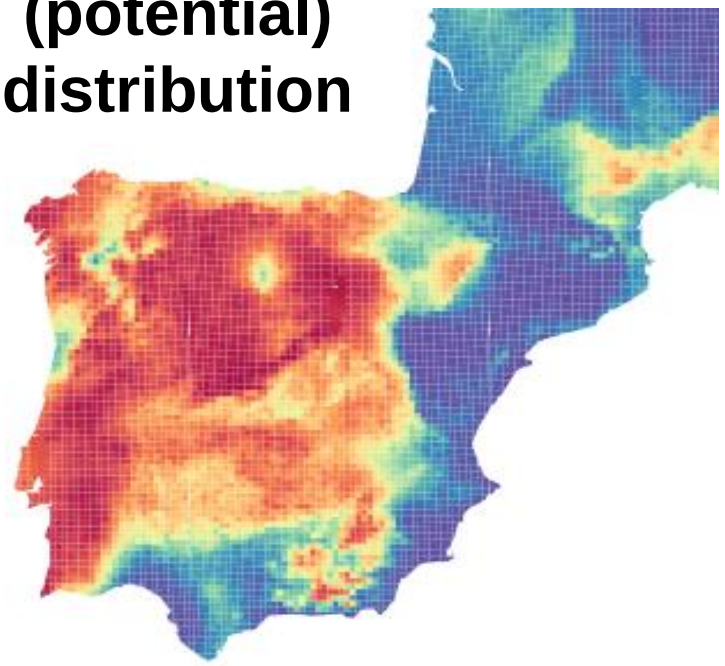
favourability

prevalence

presence

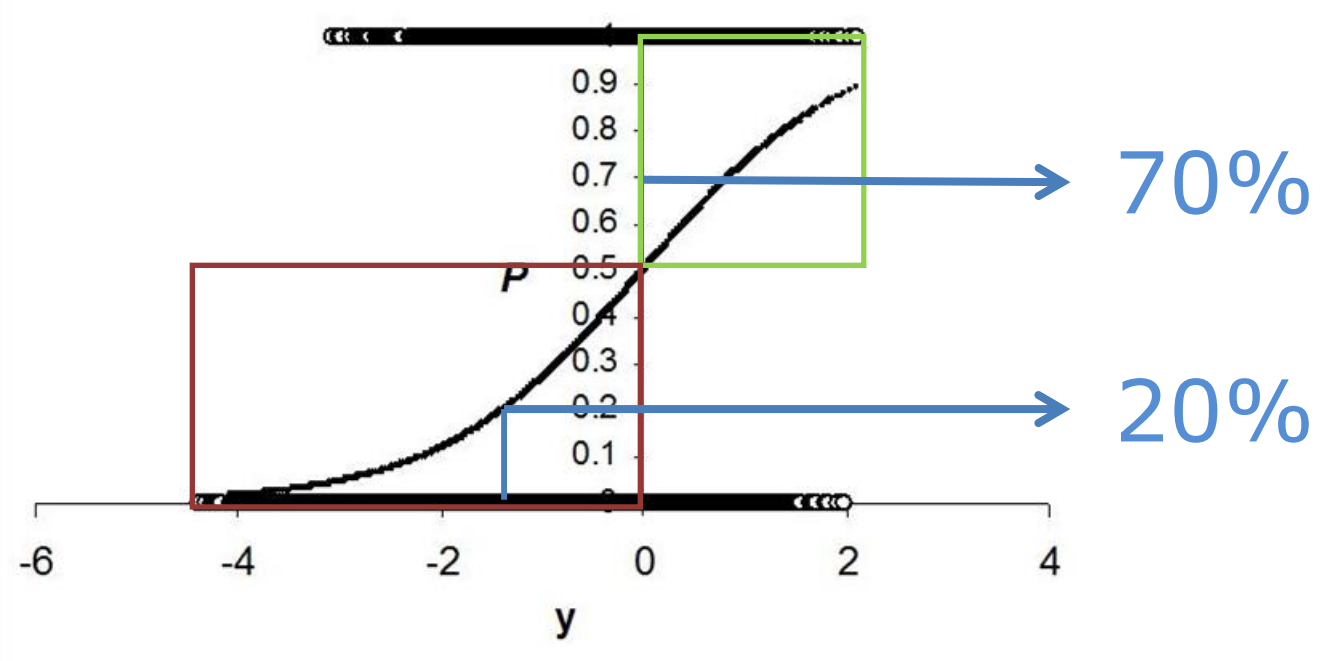
probability

predicted (potential) distribution



convert to binary prediction?!

Species distribution models

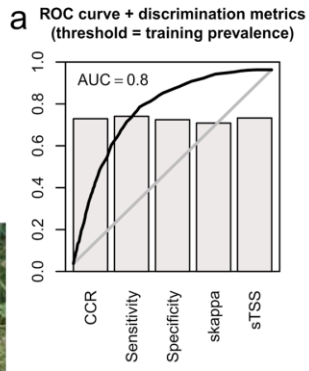


Binarizing predictions introduces an abrupt artificial difference between values near the threshold, and equates those values to the extremes, discarding important quantitative information

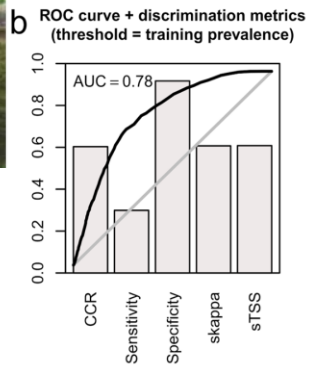
Species distribution models

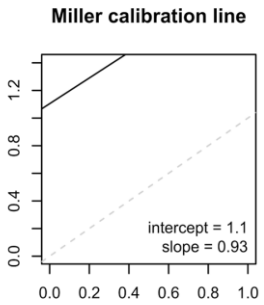
evaluated after 10 years of changes

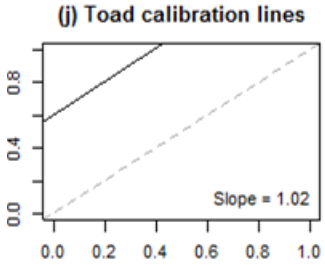
1998

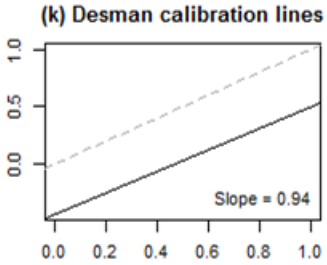


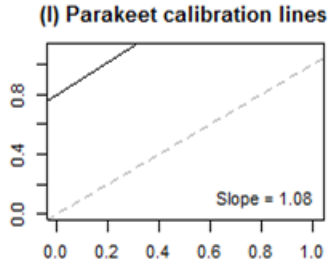
2008











presences not necessarily well classified

but very strong relationship with occurrence frequency, for species of all sorts (expanding, contracting, invasive)

if we request predictions of **“presence**” or “**absence**”, we’re bound for failure; we are trying to categorically predict something that is **not categorically predictable**

reality is not so categorical, and models approach the actual (probabilistic) reality, even when they’re not so good at classifying presence/absence

why not assume that **reality is better reflected by probabilities** than by actual observations, and **treat them with quantum-analogous methods**?

Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2016) **25,** 489–501

MACROECOLOGICAL METHODS

Delineating probabilistic species pools in ecology and biogeography

Dirk Nikolaus Karger[1,2*], Anna F. Cord[3], Michael Kessler[2], Holger Kreft[4], Ingolf Kühn[5,6,7], Sven Pompe[5,8], Brody Sandel[9], Juliano Sarmento Cabral[4,7], Adam B. Smith[10], Jens-Christian Svenning[9], Hanna Tuomisto[1], Patrick Weigelt[4,11] and Karsten Wesche[12,7]

Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2015, **6,** 853–858 doi: 10.1111/2041-210X.12372

APPLICATION

`fuzzySim`: applying fuzzy logic to binary similarity indices in ecology

A. Márcia Barbosa*

using **continuous** rather than categorical presence

improves (macro)ecological analyses

and is **feasible** on more occasions than thought

Conservation Letters

A journal of the Society for Conservation Biology

VIEWPOINT

Quantum Conservation Biology: A New Ecological Tool

Joseph W. Bull

(1) Quantum mechanics is already being applied in understanding biological phenomena at lower levels of ecological organization;
(2) Conservationists increasingly focus upon probabilities of species occurrence, not observed occurrence, presenting striking analogies to quantum mechanics.

unpredictability is intrinsic

to species distributions

concepts and methods from **quantum mechanics**

may improve understanding and prediction of species distributions

species distribution models may provide the "wavefunctions" behind presence probability

Points of View

Species Distributions, Quantum Theory, and the Enhancement of Biodiversity Measures

Raimundo Real ✉, A. Márcia Barbosa, Joseph W. Bull

Systematic Biology, Volume 66, Issue 3, May 2017, Pages 453–462, https://doi.org/10.1093/sysbio/syw072

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Ecophysics reload—exploring applications of theoretical physics in macroecology

Sidney F. Gouveia[a,*], Juan G. Rubalcaba[b], Vladislav Soukhovolsky[c,d], Olga Tarasova[d], A. Márcia Barbosa[e], Raimundo Real[f]

PRACTICAL

- we'll quickly make some **presence-(pseudo)absence** example models (GLM, GAM)
- we'll get the **presence probability** predicted by those models
- for comparison, we'll also make **Maxent** suitability models, which require the **same data** (presences + the rest of the background) but are not appropriate for fuzzy logic

note this is **not a course on modelling!**

PRACTICAL

```
dat <- gridRecords(rst = variables,
pres.coords = occurrences[ , c("lon",
"lat")], species = occurrences$spcode)

vars_sel <- corSelect(data = dat, sp.cols =
spc, var.cols = var_cols, cor.thresh = 0.7)

mod_GLM <- glm(formula = species ~ var1 +
var2 + var3, family = binomial, data = dat)

pred_GLM <- predict(model, newdata = dat,
type = "response")
```