Session: How will big data help in biodiversity conservation? → just counting trees and species?

... is not a trivial task
... at least in high diverse forests

Key questions envolved:

- What is the presumable number of undetected species?

- What is the survey reliability?

- What is the conservation status or value of any site or region (diversity repositor, seed provider, protected area, stepping stone for restoration...) ??

Inventário Florístico Florestal de Santa Catarina

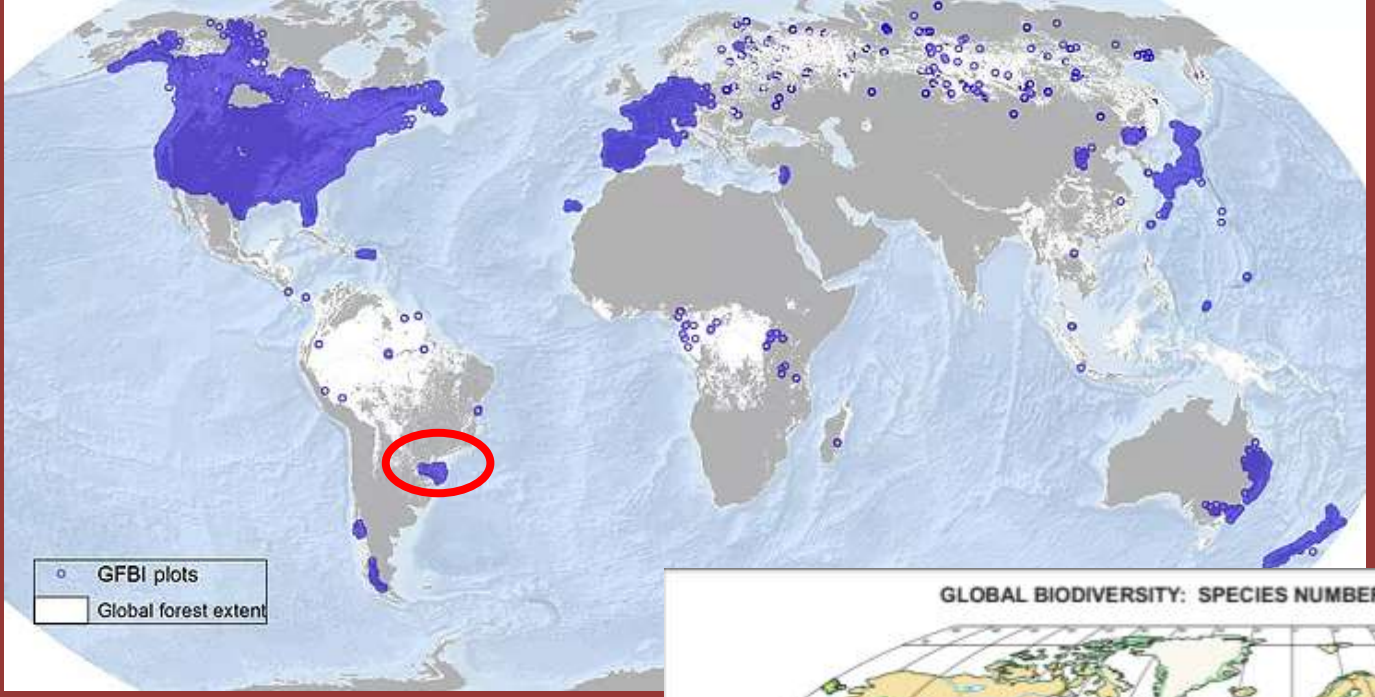
Global Forest Biodiversity Initiative
Conference Beijing - September 2017

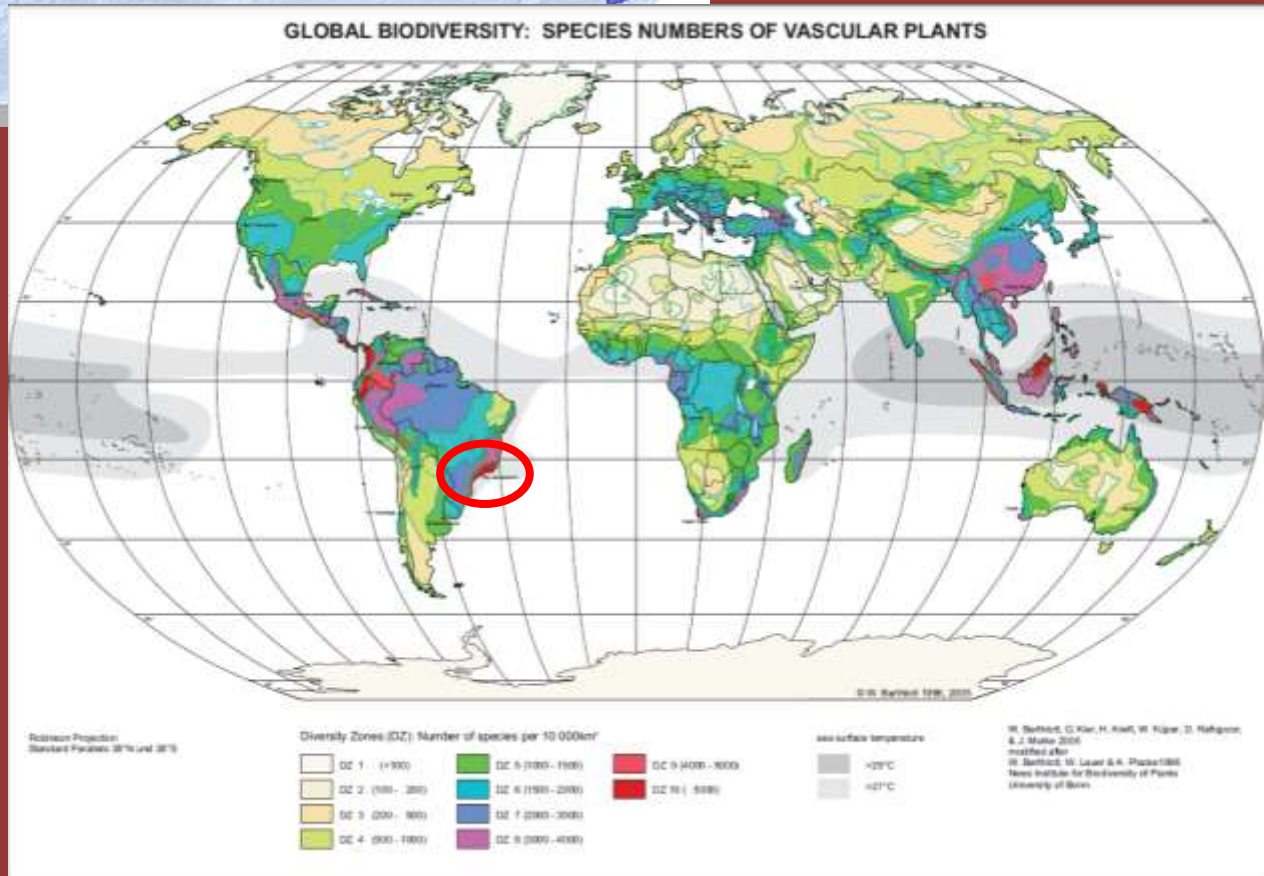
Nonparametric species-richness estimators enable accurate estimates of gamma-diversity in (sub) tropical forests?

Laio Z. Oliveira, P. Moser, A. Gasper, G. Piazza, A. Oliveira, Alexander C. Vibrans
Universidade Regional de Blumenau-Santa Catarina, Universidade Federal de Minas Gerais, Brasil

acv@furb.br
iff.sc.gov.br

FURB FAPESC GOVERNO DE SANTA CATARINA SERVIÇO FLORESTAL BRASILEIRO FAO Fepema



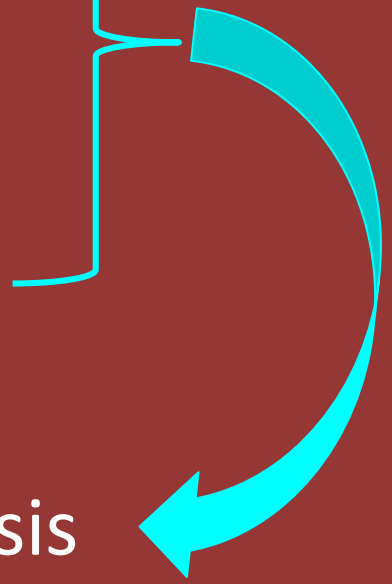


Barthlott et al. 2005

IF data are consistent...

... large scale inventories with systematic sampling can provide....

..... unbiased estimations of population parameters
.....binary species data (presence/absence)
.....species density / basal area data
.....data on:
- common species
- abundant species
- infrequent species
- rare species



Meta analysis

Hypothesis:

It is expected that estimators' performance may vary according to community structure and species' spatial distribution (patchiness), as well as to sample design/size

Assumptions

Diversity is linked to rarity, or better: to number of rare species, frequently in high number and fundamental elements in tropical forests

Recorded rare species may give informations about number of missing ones!!

Richness estimators consider rare species but have different properties and give different weight to rare species

Aims:

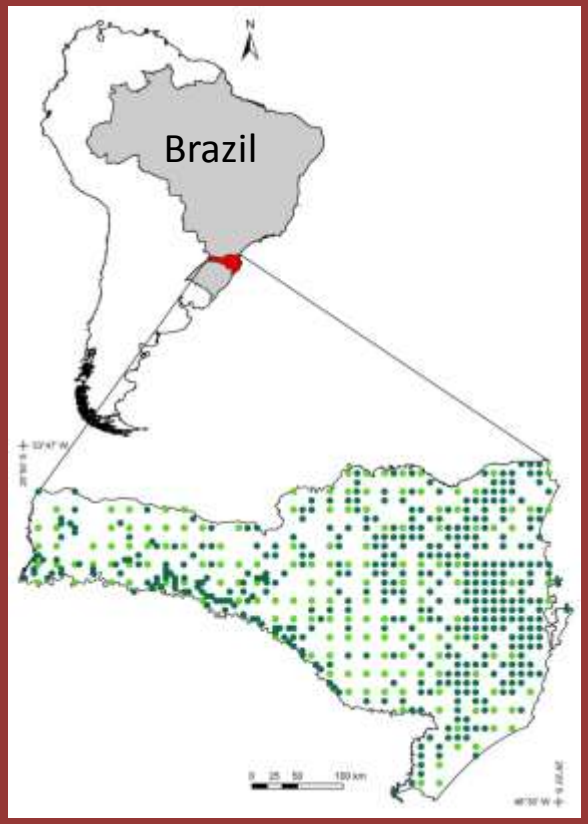
i) Assess the performance of estimators (incid./abundance based)

ii) Evaluate the effect of species patchiness (spatial distribution)

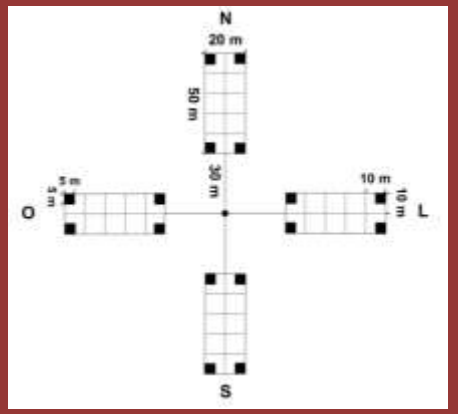
iii) Evaluate the effect of sample intensity

Methods – Study area + sample design in accordance to NFI Brazil

- sistematic sampling (permanent plots), following a nationwide grid



- 4,000m² clusters with 4 crosswise subplots (1,000m²) and 40 subunits (100m²); x,y coord. every tree DBH ≥10cm



→ **1073 permanent plots**
→ **539 forest plots**

→ **1st cycle (2007-2011)**
→ **2nd cycle (2014-2018)**

Santa Catarina State
95.000 km² (1.1% of Brazil)
26° - 29° S
Biodiversity hotspot (atlantic forest)
High endemism of vascular plants
~ 860 tree and shrub species

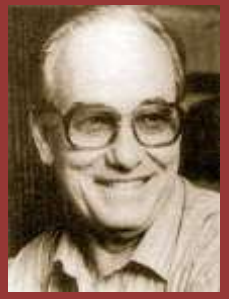
Methods:

Computing bias and precision compared to "ground truth" richness (theoretical total species richness – TTSR or S_{real})

..... S_{real} - a rare condition!

Flora Ilustrada Catarinense (1965-) 183 fasc. + NeoTropTree Data Base (Oliveira-Filho, 2014)

(Klein & Reitz 1978; 1981)



Roberto M. Klein
1923 - 1992

Raulino Reitz
1919 - 1990

www. http://prof.icb.ufmg.br/treeatlan/

$$S_{Chao\,1} = S_{obs} + \frac{F_1^2}{2F_2}$$

$$S_{ACE} = S_{abund.} + \frac{S_{rare}}{C_{ACE}} + \frac{F_1}{C_{ACE}}\gamma^2_{ACE}$$

Abundance based

Magurran 2004
Measuring biol. diversity

$$S_{Chao\,2} = S_{obs} + \frac{Q_1^2}{2Q_2}$$

$$S_{ICE} = S_{freq.} + \frac{S_{infr}}{C_{ICE}} + \frac{Q_1}{C_{ICE}}\gamma^2_{ICE}$$

$$S_{Jack\,1} = S_{obs.} + Q\left(\frac{m-1}{m}\right)$$

$$S_{Jack\,2} = S_{obs.} + \left[\frac{Q_1(2m-3)}{m} + \frac{Q_2(m-2)^2}{m(m-1)}\right]$$

Incidence based

EstimateS 8.2

Sample-based Extrapolation (2x)
(Colwell et al. 2012).

Michaelis-Menten $S(n) = \frac{S_{\max}.n}{B+n}$

$S_{obs.}$ = n° of observed species

F_1= n° of singletons (1 ind.)

F_2= n° of doubletons (2 ind.)

Q_1= n° of uniques (in 1 plot)

Q_2= n° of duplicates (in 2 plots)

S_{rare} = n° of rare species (< 10 ind.)

S_{abund}= n° of abundant species (≥ 10 ind.)

S_{infr} = n° of infrequent species (< 10 plots)

$S_{freq.}$ = n° of common species (≥ 10 plots)

N_{rare} = n° of individuals of rare species

m_{infr} = n° of sample plots with at least 1 infrequent species

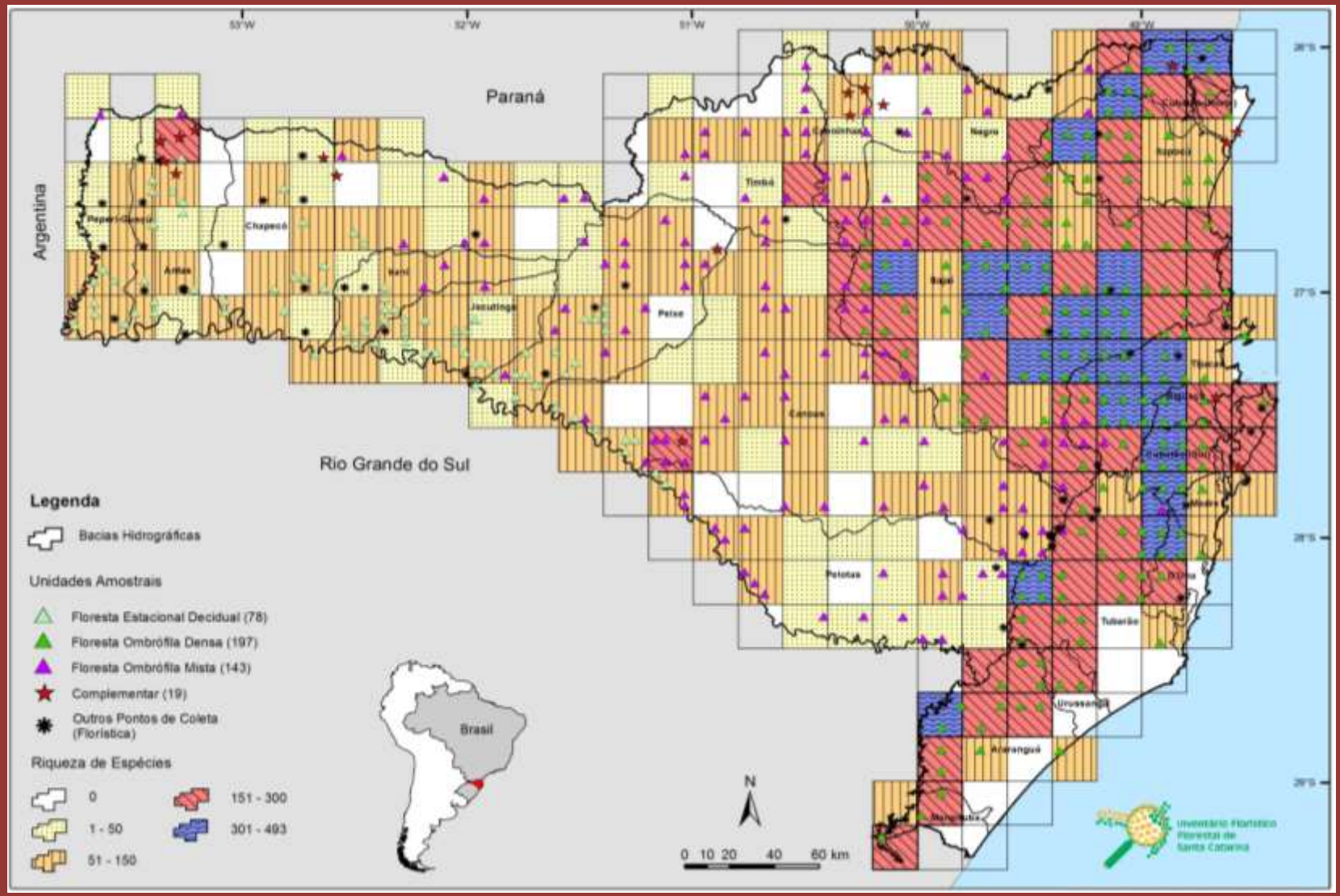
$C_{ACE} = 1-F_1/N_{rare}$

$C_{ICE} = 1-Q_1/m_{infr}$

$\gamma^{ACE/ICE}$ = estimated coefficient of variation of F and Q

i) Assess the performance of estimators (incidence/abundance based)

(Observed) Species richness map (400km² grid cells)



Semi Decid. Forests Araucaria Forests Evergreen Rainforest

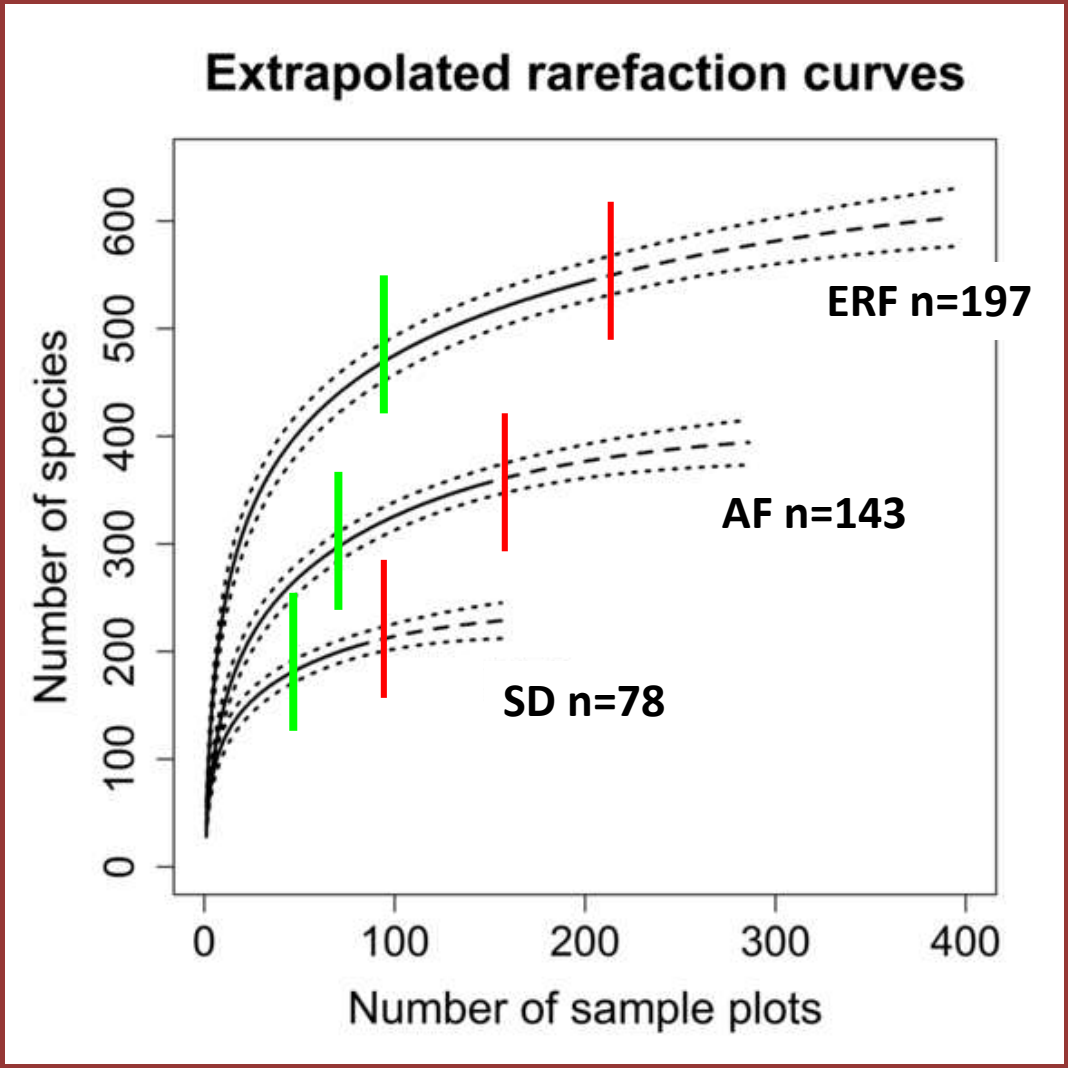
Rarity category	SF ($n = 78$)	%	AF ($n = 143$)	%	ERF ($n = 197$)	%
Singleton	33	16.0	47	13.3	59	10.9
Doubleton	48	23.3	26	7.3	93	17.2
≤10 recorded individuals	96	46.6	159	44.9	222	41.0
Unique	43	20.9	75	21.2	97	17.9
Duplicate	69	33.5	52	14.7	145	26.8
≤10 sample plots	124	60.2	234	66.1	301	55.5
Total species richness (S_{obs})	206	–	354	–	542	–

n sample size, *ERF* evergreen rainforest, *AF Araucaria* forest, *SF* semi-deciduous forest, % is the proportion of S_{obs}

What is the real species richness??
...according to Flora + Neotroptree

	S_{obs}	S_{real}	
ERF - Evergreen rainforest	542	708	+ 30%
AF - Araucaria Forest	354	463	+ 30%
SD - Semi-decid. Forest	204	307	+ 50%
Total	620	859	+ 38%

.....what do the estimators preview?

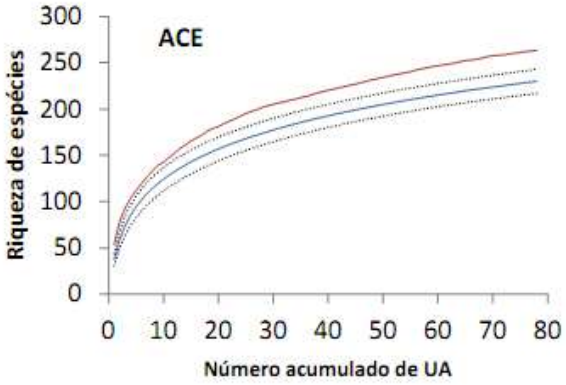


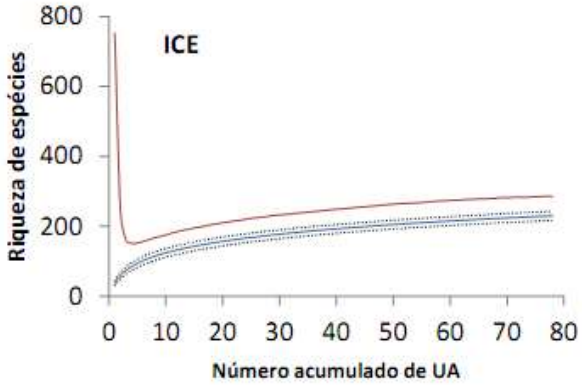
Species number x sampling effort

- **80% of total species number with 50% of sampling effort**

ERF – Evergreen rainforest 85%
AF – Araucaria Forest 83%
SD – Semi deciduous Forest 83%

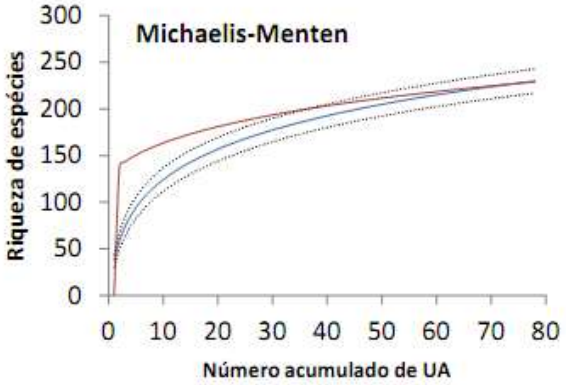
---- extrapolation (2x)
(Colwell et al. 2012)

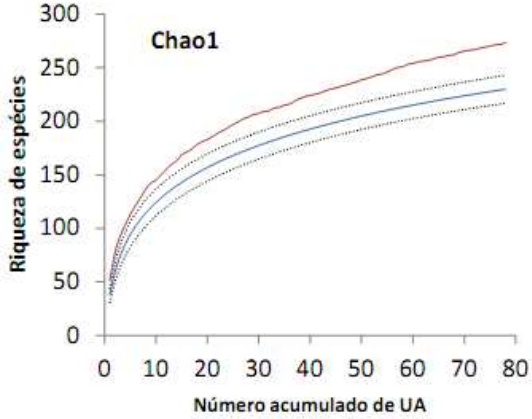


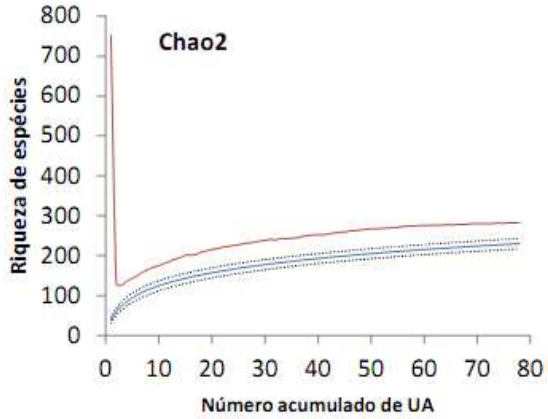


Different estimators...

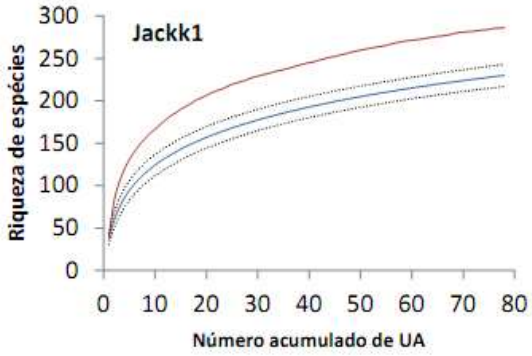
→ ...different results

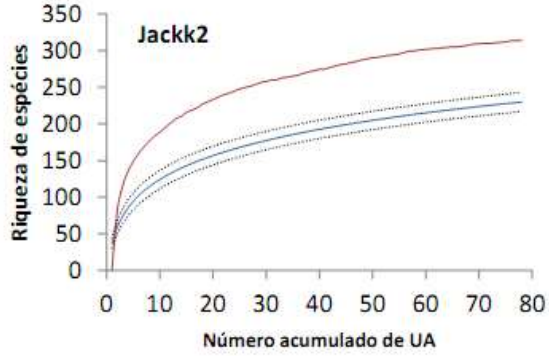






...who is "right"??





Rarefaction curve and estimators for **Semi- deciduous Forests** (SF), based on 78 IFFSC sample plots.

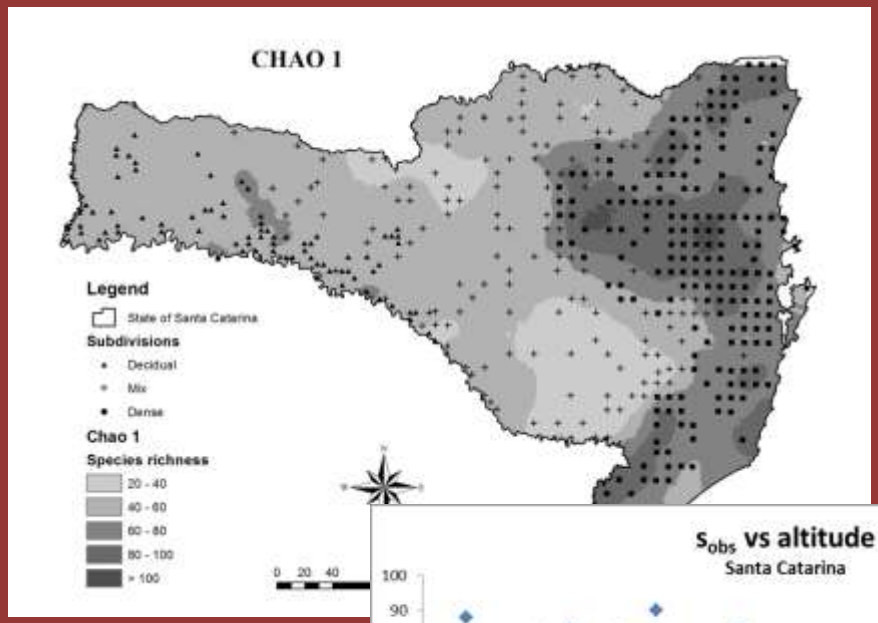
$\hat{S}_{real}$ x S_{real}

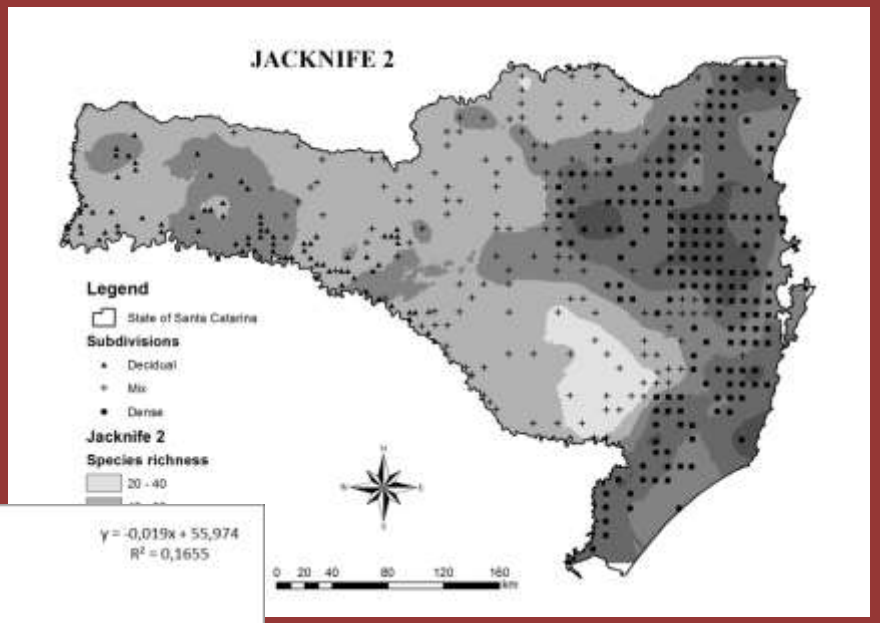
Table 2 Performance measures of nonparametric species-richness estimators applied to IFFSC's species-richness data; the estimators are presented in order of increasing bias

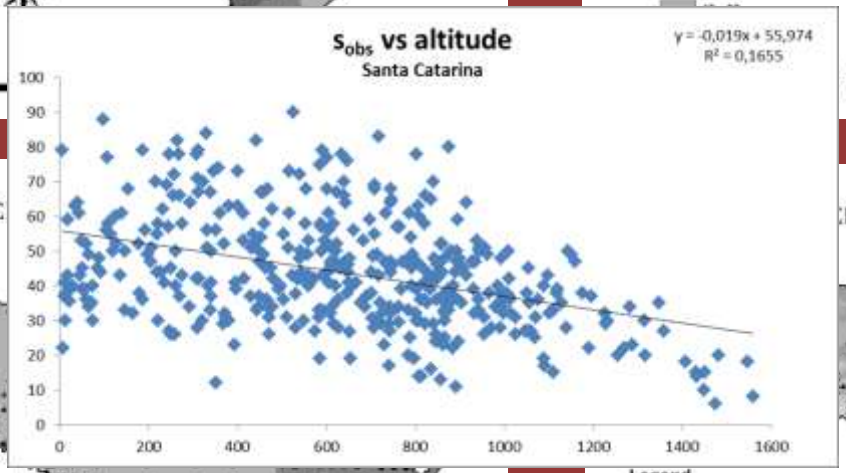
Forest type	Estimator	$\hat{S}_{real}$	Bias	Precision
ERF	Jackknife2	687.2	−0.1700	0.1769
$n = 197$	Jackknife1	638.5	−0.2355	0.1854
$S_{obs} = 542$	Chao2	636.5	−0.2405	0.1764
$S_{real} = 708$	ICE	617.8	−0.2460	0.1527
	Michaelis–Menten	533.8	−0.2533	0.1166
	Extrapolation (2×)	603.4	−0.2784	0.1877
	Chao1	590.9	−0.3001	0.1930
	ACE	578.1	−0.3154	0.1938
	S_{obs} (rarefaction)	542.0	−0.3730	0.2152

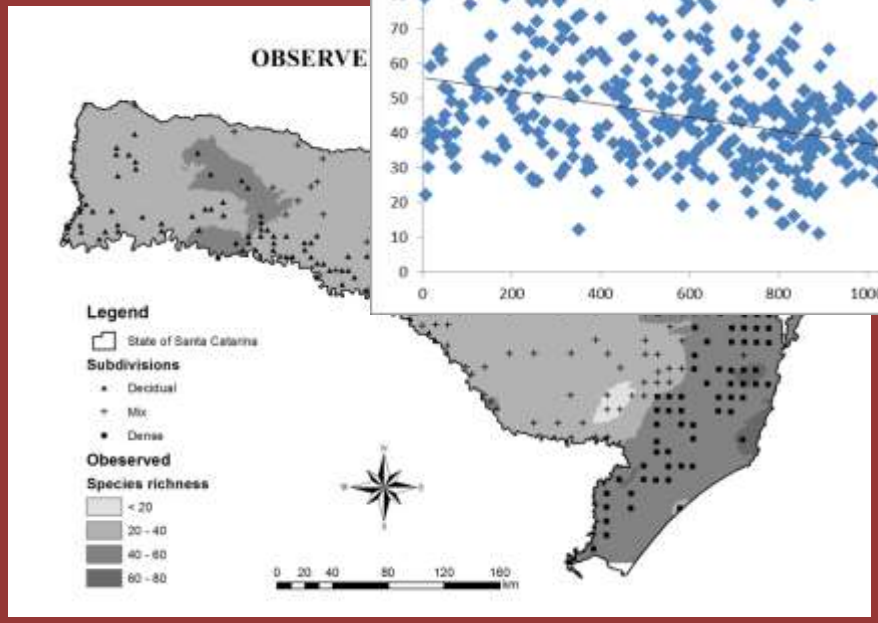
S_{obs} observed species-richness, S_{real} total theoretical species richness, $\hat{S}_{real}$ total estimated species richness, n sample size, *ERF* evergreen rainforest, *AF Araucaria* forest, *SF* semi-deciduous forest

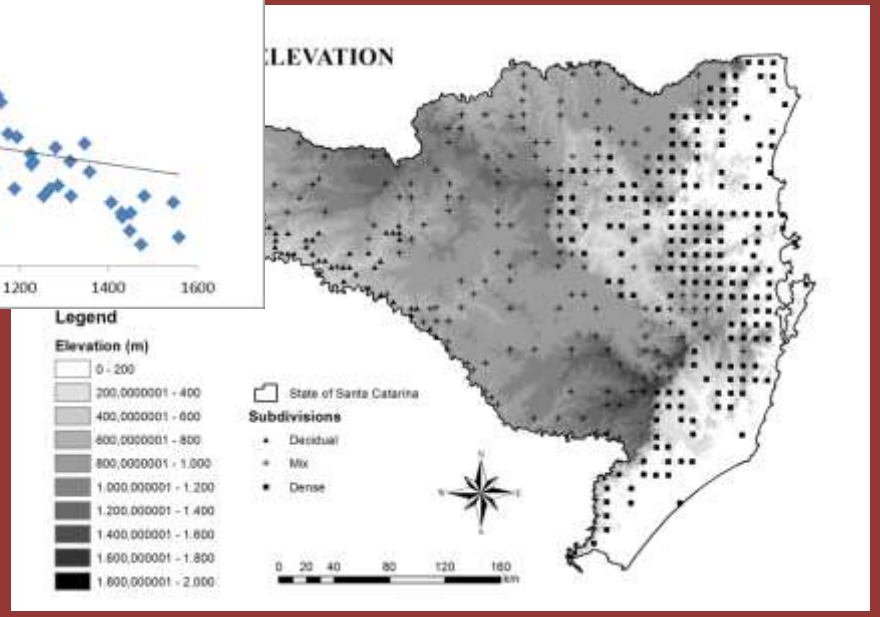
Interpolated species richness maps (kriging)











ii) Evaluate the effect of species patchiness (spatial distribution)

Patchiness

Does spatial correlation occur within a 10km sample grid? May be...

Considering spatial correlation:

We tested three degrees of patchiness (denoted, A) for the forest type richness estimation:

i) a randomized spatial distribution of species (A=0),

ii) an intermediate degree of patchiness (A=0.5) and

iii) a high degree of patchiness (A=0.75)

(Chazdon et al. (1998) and Colwell (2013)

Incidence based estimatores (Chao 2, Jack 1+2, ICE) are more robust i.e. less sensible regarding species patchiness + bias

Precision decreases in all cases

Forest type	Estimator	Bias			Precision		
		$A = 0$	$A = 0.5$	$A = 0.75$	$A = 0$	$A = 0.5$	$A = 0.75$
ERF	Chao1	−0.244	−0.301	−0.341	0.113	0.175	0.232
$n = 197$	Chao2	−0.247	−0.233	−0.222	0.122	0.150	0.185
$S_{obs} = 542$	Jackknife1	−0.234	−0.233	−0.236	0.129	0.170	0.218
$S_{real} = 708$	Jackknife2	−0.194	−0.172	−0.152	0.127	0.165	0.208
	ACE	−0.257	−0.311	−0.357	0.111	0.177	0.239
	ICE	−0.254	−0.233	−0.218	0.111	0.195	0.196
	M.-Menten	−0.256	−0.268	−0.297	0.112	0.139	0.169
	Extrap. (2×)	−0.268	−0.275	−0.284	0.134	0.174	0.217
	S_{obs} (Raref.)	−0.264	−0.294	−0.326	0.157	0.196	0.243

$A = 0$ randomized spatial distribution of species, $A = 0.5$ intermediate degree of patchiness, $A = 0.75$ high degree of patchiness, S_{obs} observed species richness, S_{real} total theoretical species richness, $\hat{S}_{real}$ total estimated species richness, n sample size, SF Semi-deciduous Forest, AF *Araucaria* Forest, ERF evergreen rainforest

Values denoted with by * were obtained using the Chao2 classic formula

iii) Evaluate the effect of sample intensity

Species richness – Evergreen Rainforest (ERF)
Comparing estimations with S_{obs} and S_{real} at different sample sizes

10 km
S_{obs} = 542
+ 36%
n=197

20km
S_{obs} = 398
n=46

Estimators

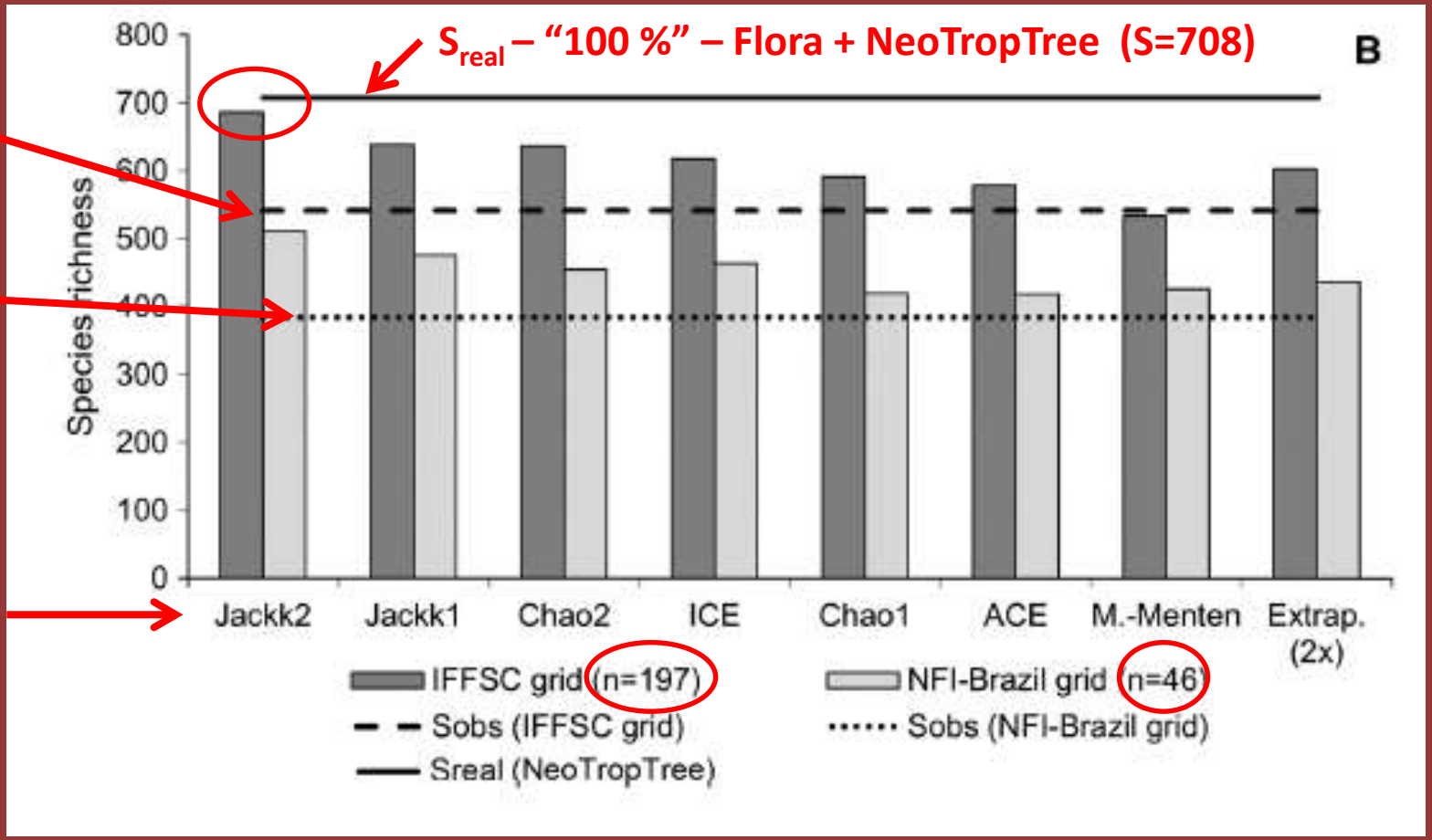


Fig. 4 Species-richness estimates by nonparametric estimators based on data from IFFSC's 10-km grid and the same grid adapted to NFI-Brazil 20-km grid. a *Araucaria* Forest, b evergreen rainforest

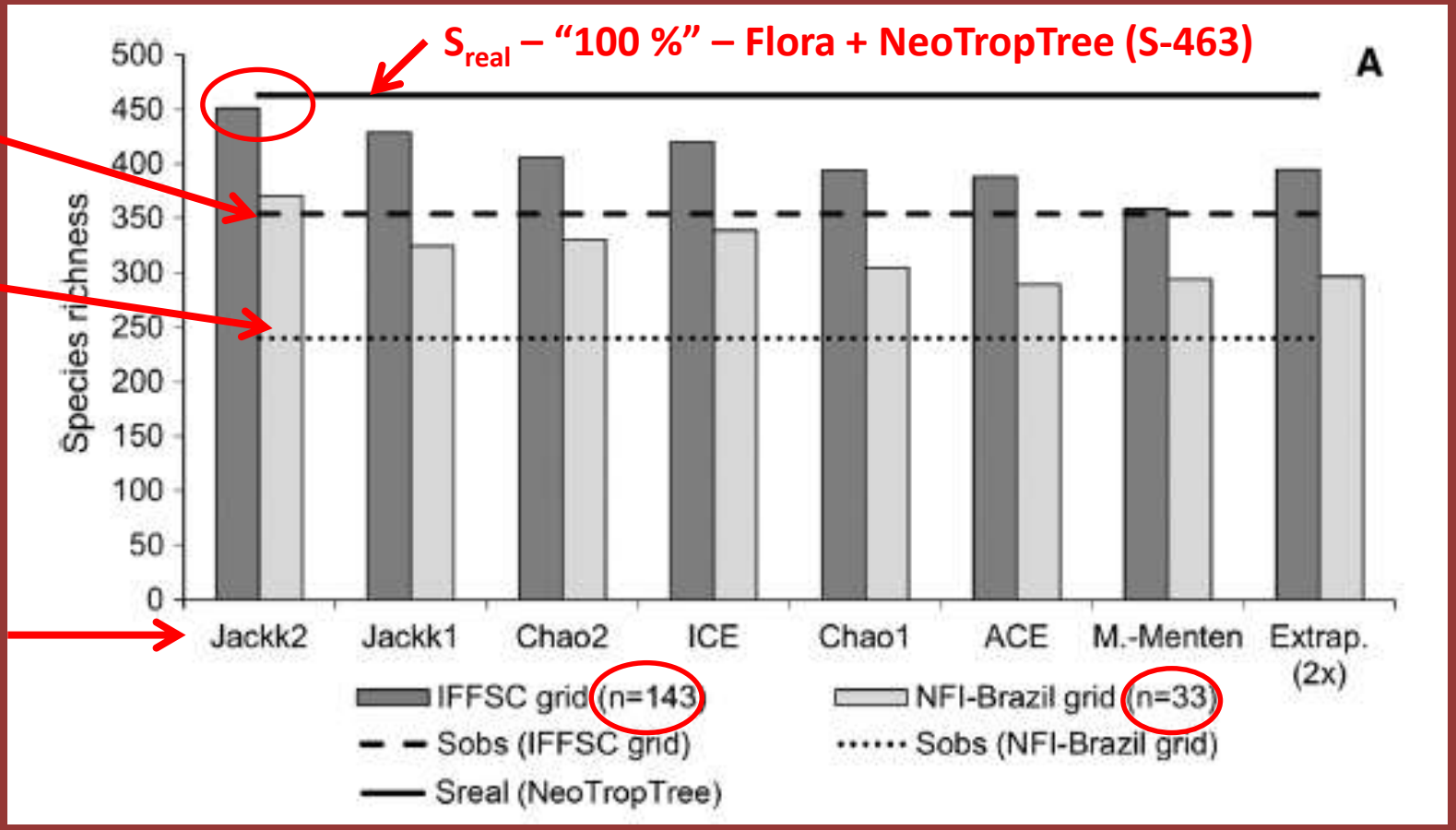
Oliveira, LZ, Moser, P, Vibrans, AC, Piazza, G, Gasper, A, Oliveira-Filho, A (2016) Insights for selecting the most suitable nonparametric species richness estimators for subtropical Brazilian Atlantic Forests, Braz. J. of Botany 39: 593-603

Species richness – Araucaria Forest (AF)
Comparing estimations with S_{obs} and S_{real} at different sample sizes

10 km
S_{obs} = 354
+ 44%
n=197

20km
S_{obs} = 245
n=46

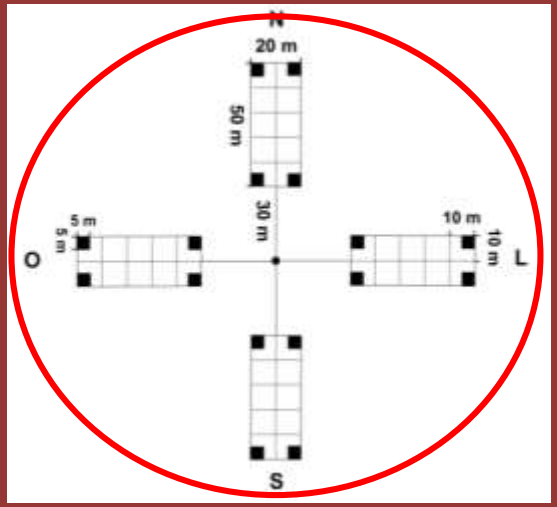
Estimators



Oliveira, LZ, Moser, P, Vibrans, AC, Piazza, G, Gasper, A, Oliveira-Filho, A. Insights for selecting the most suitable nonparametric species richness estimators for subtropical Brazilian Atlantic Forests, Braz. J. of Botany 39: 593-603

“Extra floristics”

Collect all fertile plants inside and around the plot and on the access way



Increase of S_{obs} ~ 10%

(presence/absence only)

Conclusions

- No richness estimator exceeds theoretical (real) richness (less bad....)
- Under the given conditions (proportion of rare and common species) the Incidence based estimators performed better ($S_{Jack\,2} = S_{obs.} + \left[\frac{Q_1(2m-3)}{m} + \frac{Q_2(m-2)^2}{m(m-1)}\right]$ uniq/dupl/infr
- Incidence based estimators showed to be less sensitive (↓bias) to "patched" species spatial distribution
- These estimators allow to compute n° of missing species and therefore the inventory reliability
- With increasing sample intensity (rare) species detection increases (36-44%)
- Extra floristics + understory species may close the gap between observed and real species richness in large area inventories

make this exercise with your data!
quality control + modelling potencial richness and gaps!

Braz. J. Bot
DOI 10.1007/s40415-016-0269-8

Insights for selecting the most suitable nonparametric species-richness estimators for subtropical Brazilian Atlantic Forests

Laio Z. Oliveira[1] · Paolo Moser[1] · Alexander C. Vibrans[1] · Gustavo A. Piazza[2] · André L. de Gasper[3] · Ary T. Oliveira Filho[4]

Inventário Florístico Florestal de Santa Catarina

Xie-Xie!!

acv@furb.br
http://www.iff.sc.gov.br/
http://www.furb.br/sinflor/

Universidade Regional de Blumenau - FURB Santa Catarina – Brasil

Welcome to XXV IUFRO World Congress Curitiba 2019

FURB FAPESC GOVERNO DE SANTA CATARINA SERVIÇO FLORESTAL BRASILEIRO FAO Fepema