Evaluating spatial and temporal fragmentation of a categorical variable using new metrics based on entropy: example of vegetation land cover

entropy 2021 The Scientific Tool of the 21st Century
5–7 May 2021, ONLINE

Didier G. Leibovici [1,2*], Christophe Claramunt [3] and Shaun Quegan [1]

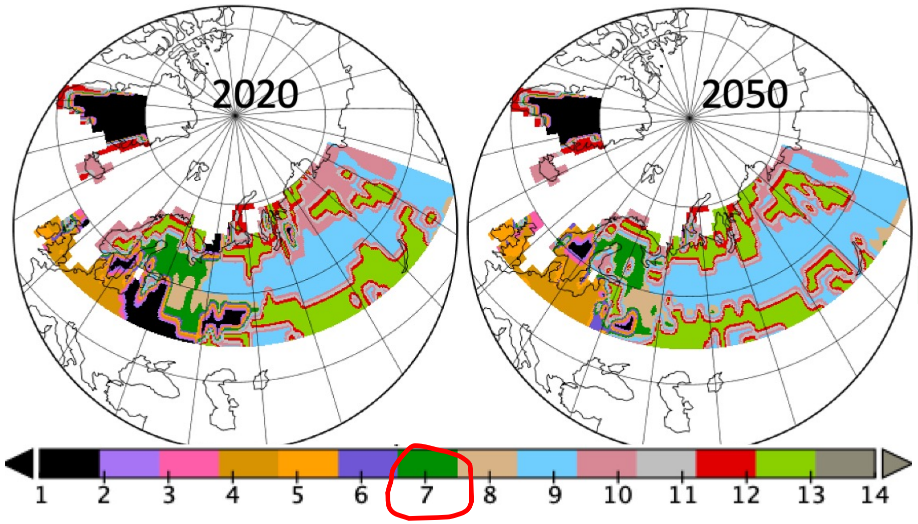
1 School of Mathematics and Statistics, University of Sheffield, Sheffield, UK;

2 GeotRYcs Cie, Montpellier, France;

3 Naval Academy Research Institute, Brest cedex 9, France

Abstract: Associated with climate change and/or land use pressure, forest fragmentation is a spatio-temporal shrinking process that reduces the sizes of forest patches. This breaks up forest patches so increasing their number before the small ones progressively disappear. Fragmentation can be assessed spatially as a level of the current status of the fragmented spatial configuration and temporally as the level of the speed of the fragmentation process itself. Among the different landscape metrics based on patches as indicative measures for fragmentation, the Shannon entropy of the observed spatial distribution of categories has been of particular interest. Based on a recently suggested spatio-temporal entropy framework focusing on patch size and shape distributions, this paper shows how to derive useful fragmentation metrics at local and global levels, spatially, temporally or both. Moreover, it shows that using fully symmetric approaches between space, time and category within this framework, can lead to more sensitive fragmentation metrics as well as providing complementary local approach for cartographic representation. Land cover data simulations from land surface modelling to a 2100 horizon are used to illustrate the proposed fragmentation metrics.

Dominant *vegetation*



7 boreal needleleaf evergreen

Keywords: Shannon entropy; spatio-temporal information; fragmentation; spatio-temporal process; categorical variable; vegetation; land cover; climate change; land use change

* Correspondence: d.leibovici@sheffield.ac.uk; geotrycs@gmail.com

Spatially: Is this area fragmented?

(habitat) **fragmentation**:

- Reduction in the total area **T+++ S+**
- Decrease of the interior **T+++ S++**
- Isolation of one area (fragment) from other areas **T+ S++++**
- Breaking up of one patch into several smaller patches **T+++ S+++**
- Decrease in size patch **T++++ S++**

WIKIPEDIA
The Free Encyclopedia

Temporally: How much more (fragmented) since last year?

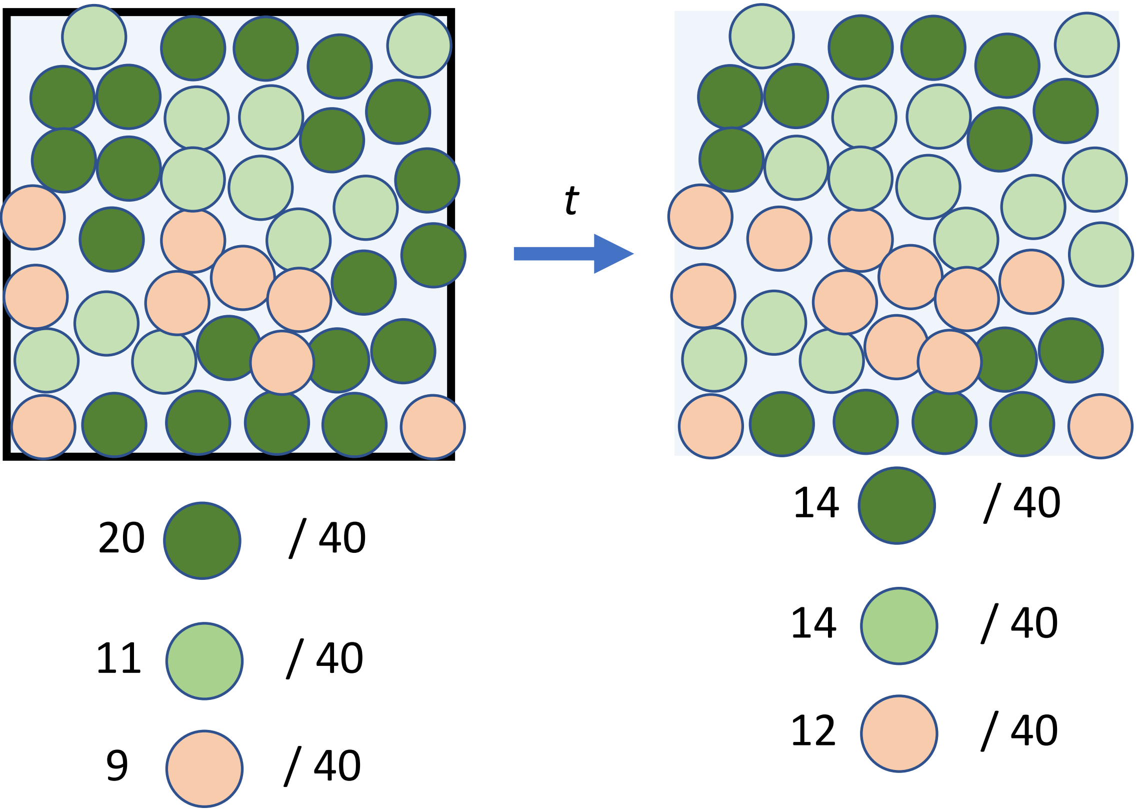
GeotRYcs

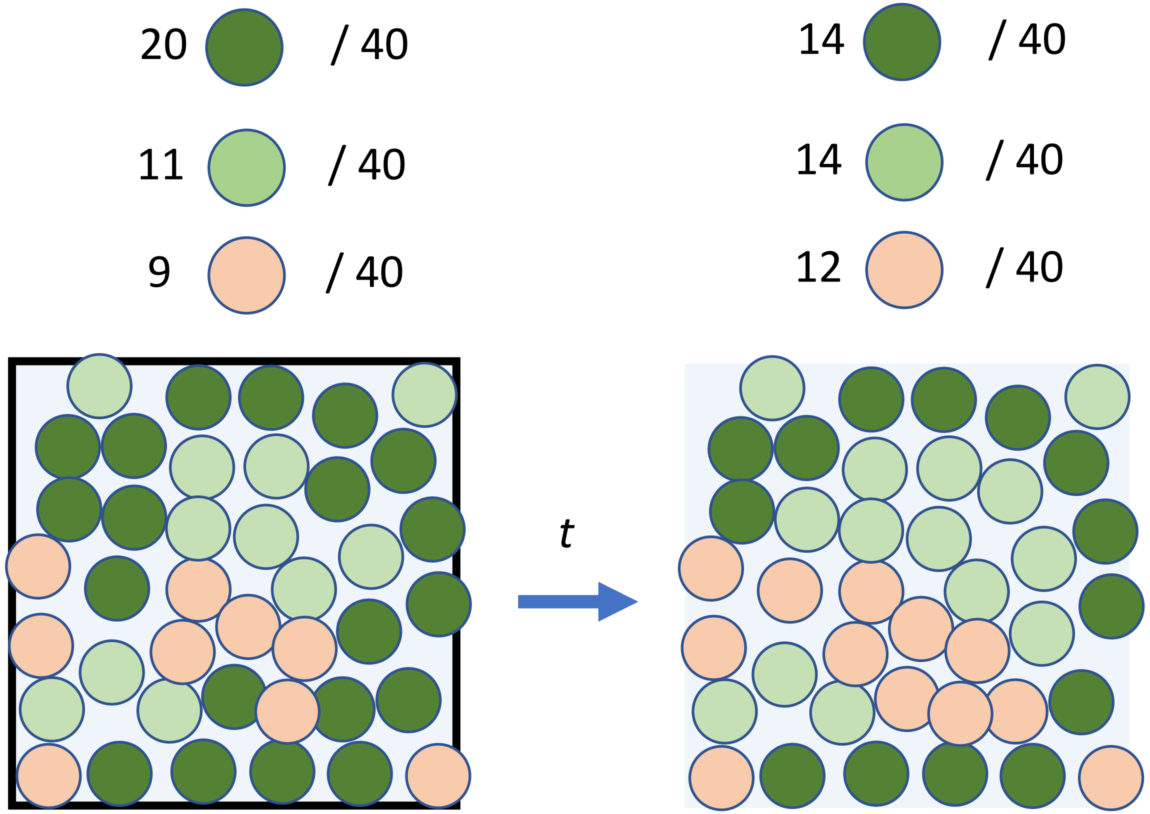
ECOLE NAVALE

The University Of Sheffield.

•Reduction in the total area **T+++ S+**

Focusing on forest





•Isolation of one area (fragment) from other areas **T+ S++++**

•Breaking up of one patch into several smaller patches **T+++ S+++**

•Decrease of the interior **T+++ S+**
(edge density)

•Decrease in size patch **T++++ S++**

Spatial entropy?

$H(C) = -\sum_c p_c log(p_c)$

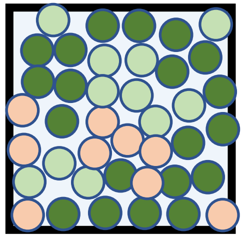
C

$H(C) = log(3)$

$max_{\{p_c, c=1,\dots,|C|\}}(H(C)) = log(|C|)$

uniform distribution, $p_c = 1/|C|$

occurrences distribution



20 / 40

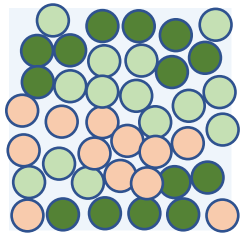
11 / 40

9 / 40

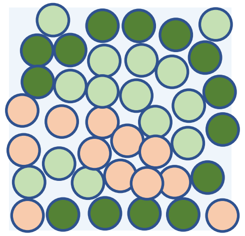
H(C) = 1.037217

H^u(C)= 0.9441156

isolation & breaking



= ?



size patch & interior

14 / 40

14 / 40

12 / 40

H(C) = 1.096067

H^u(C)= 0.9976835

Focusing on forest ?

$H^u(C) = -1/log(|C|) \sum_c p_c log(p_c) = 1$

(nearly) uniform distribution

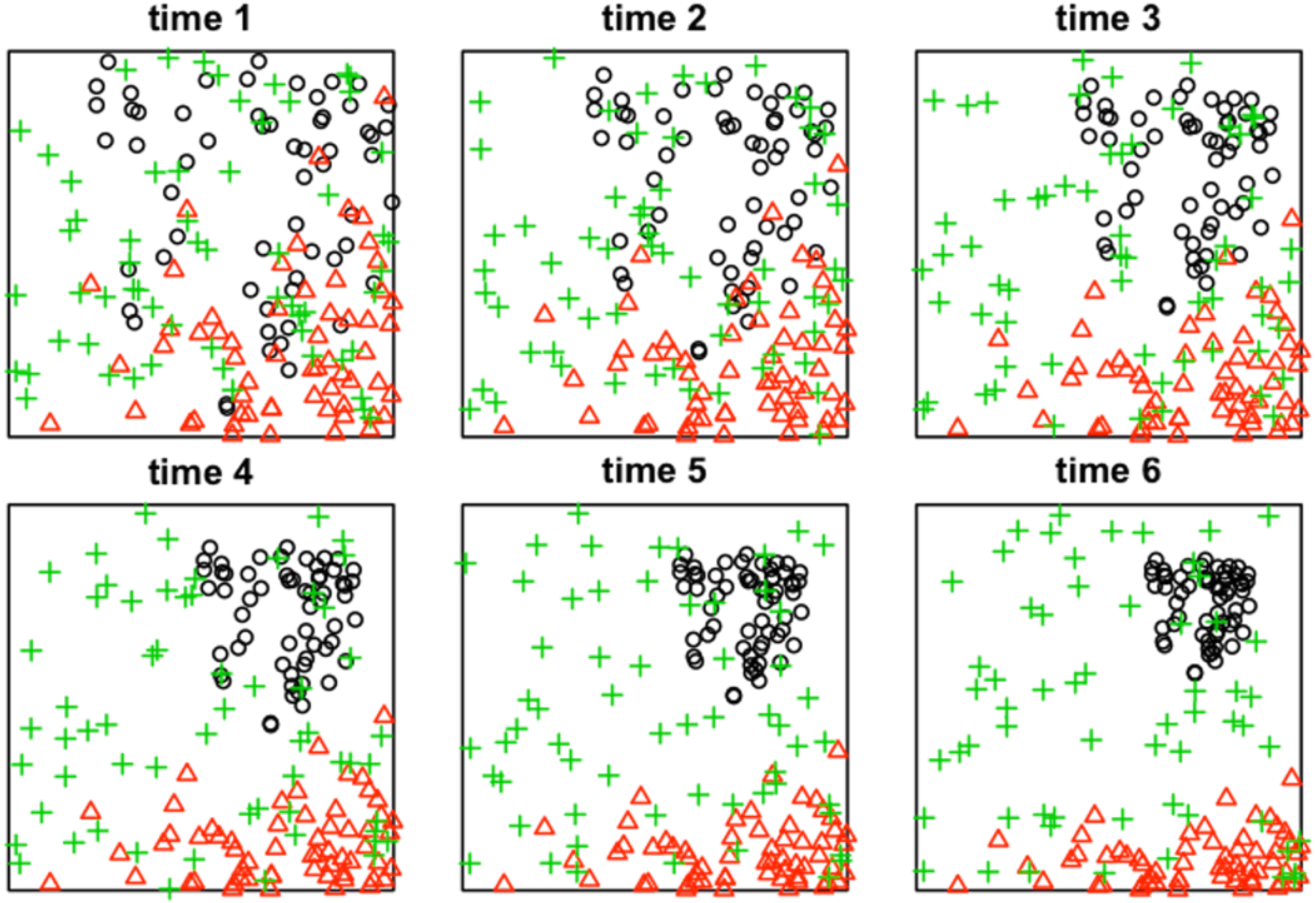
different meanings but no difference in entropy!

Spatial entropy?

occurrences distribution

60 points per class: o, + and △ C

$H(C) = log(3)$



$H^u(C) = -1/log(|C|) \sum_c p_c log(p_c) = 1$

uniform distribution (not spatially)

different patterns but no difference in entropy!

The University Of Sheffield.

Spatial entropy?

route >>> categorical to spatial

route >>> spatial to categorical

Spatial entropy?

See for example [Leibovici DG](), Claramunt C, Le Guyader D and Brosset D (2014) Local and global spatio-temporal entropy indices based on distance-ratios and co-occurrences distributions. **International Journal of Geographical Information Science, 28(5): 1061-1084**

See [Leibovici DG]() and Claramunt C (2019) On Integrating Patch Size and Shape Distributions into a Spatio-Temporal Information Entropy Framework **Entropy, 21(11):1112 (special issue)**

The University Of Sheffield.

ECOLE NAVALE

Spatial entropy !

route >>> categorical to spatial

1 neighbouring
- co-occurrences distribution
- spatial weights

2 global methods
- spatial entropy
- spatio-temporal entropy

SOOk

selSOOk

CAkOO

3 localising
- zone mapping (extra structure)
- scan statistics
- density
- multiway analysis

scankOO

dkOO

route >>> spatial to categorical

A patch structures
- patch size distribution
- patch shape distribution

B decomposition
- conditional entropy
- tensor decomposition
- non-negative approximation

PsishENT

C re-localising
- conditional mapping
- relative intensity
- multiway analysis

1 A 2 C B 3

nnFCAk

nnCAkOO

kOO framework ⊂ PsishENT framework

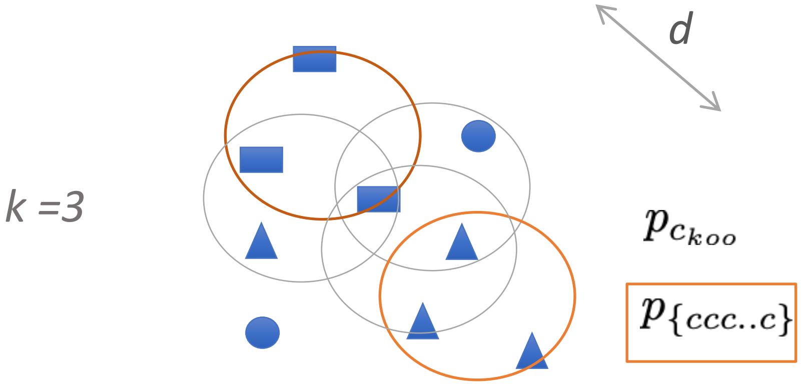
GeotRYcs

ECOLE NAVALE

The University Of Sheffield.

Spatial entropy !

co-occurrences distribution

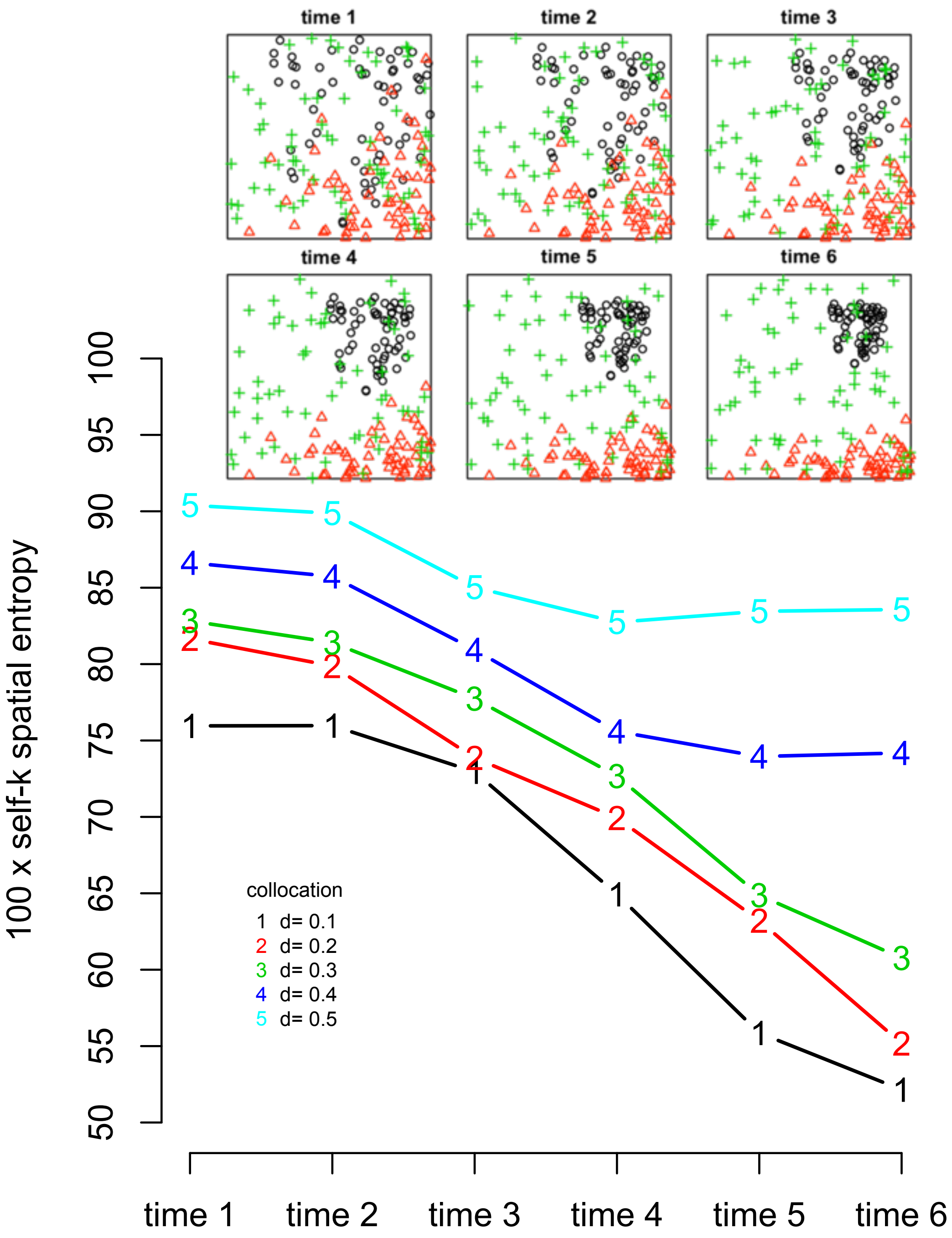


k-spatial entropy

$$H_S^u(C,k) = -1/log(|C_{koo}|) \sum_{c_{koo}} p_{c_{koo}} log(p_{c_{koo}})$$

Self-k-spatial entropy

$$H_{sS}^u(C,k) = -1/log(|C_{\{ccc..c\}}|) \sum_{c} p_{\{ccc..c\}} log(p_{\{ccc..c\}})$$



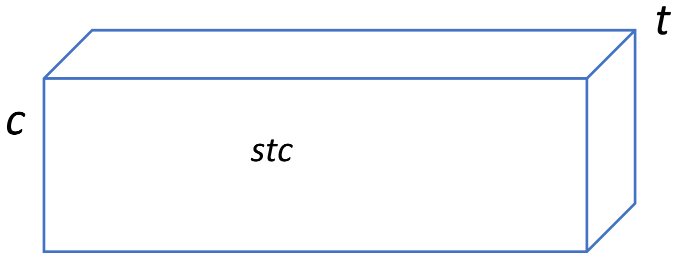
The University Of Sheffield.

Spatial entropy !

size & shape

patch size distribution

category



time patch size "group"

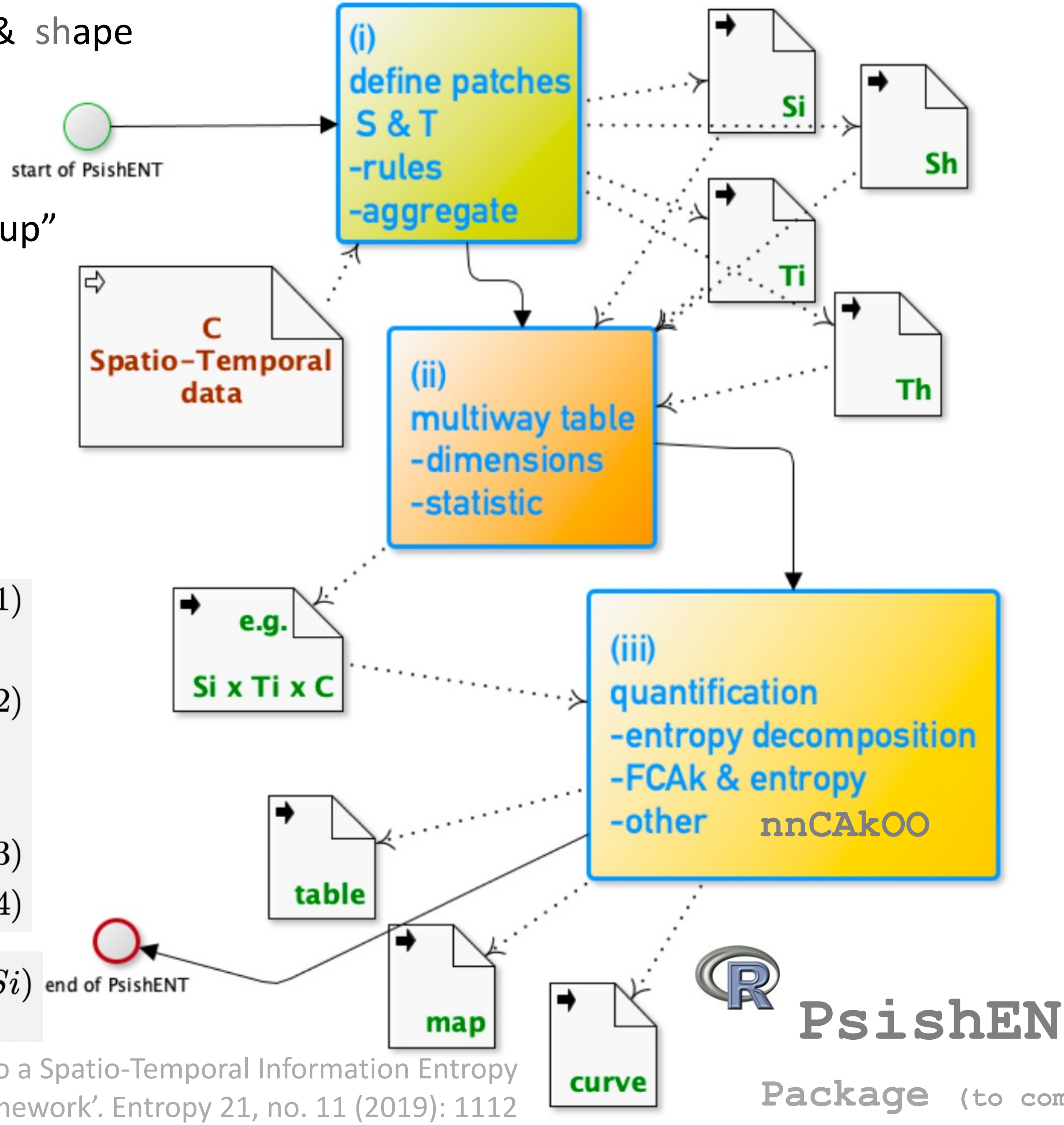
spatial patch size "group"

$$Si \times Ti \times C$$

count of occurrences

$$\begin{aligned}
H(C,Si,Ti) &= H(C) + H(Si) + H(Ti) - MI(C,Si,Ti) && (1)\\
&= H(Si,Ti) + H(C \mid Si,Ti) = H(C) + H(Si,Ti \mid C)\\
&= H(Si) + H(Ti \mid Si) + H(C \mid Si,Ti) && (2)\\
&= H(Si) + H((C,Ti) \mid Si)\\
&= H(Si) + H(Ti \mid Si) - H(Si \mid Ti) + H(C,Si \mid Ti)\\
&= MI(Si,Ti) + H(Ti \mid Si) + H(C,Si \mid Ti) && (3)\\
&= MI(Si,Ti) + H(Si \mid Ti) + H(C,Ti \mid Si) && (4)
\end{aligned}$$

$$D_{KL}(p_{SiTiC} \mid p_{Si} \otimes p_{Ti} \otimes p_C) = \sum_{stc} p_{stc} log(p_{stc}/(p_s p_t p_c)) =^{def} MI(C,Ti,Si)$$



Leibovici, DG, and Claramunt. 'On Integrating Size and Shape Distributions into a Spatio-Temporal Information Entropy Framework'. Entropy 21, no. 11 (2019): 1112

The University Of Sheffield.

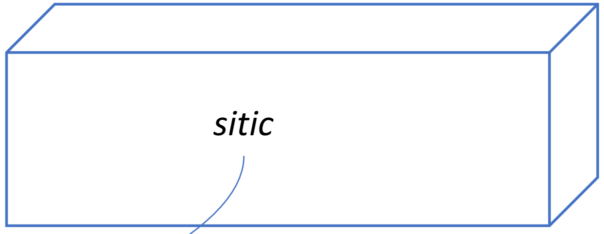
Figure 1. *Modus Operandi* of the patch size and shape entropy (PsishENT) framew

Spatial entropy !

patch size distribution

category

c



ti time patch size “group”

$$Si \times Ti \times C$$

si spatial patch size “group”

***(si)(ti)(c)*count of occurrences**

or

count of co-occurrences

$$o_1, o_2, o_3 \in \mathcal{O}_{stc} \text{ in co-occurrence of order } k = 3 \text{ for } C = c,$$
$$\text{iif } max_{o,o' \in \{o_1,o_2,o_3\}} d(o, o') \leq d_\epsilon$$
$$\text{where } d \text{ distance in } \mathcal{S} \times \mathcal{T}$$

or

local statistic (*e.g.*, *distance-ratio*)

$$d_{stc}^{ratio} =^{def} \frac{mean_{(o_1,o_2) \in W_c} d(o_1, o_2)}{mean_{(o_1,o_2) \in B_c} d(o_1, o_2)}$$

$$\text{where } W_c = \{(o_1, o_2) \in \mathcal{O}_{st} \times \mathcal{O}_{st} \mid C(o_1) = c, C(o_2) = c\}$$
$$\text{and } B_c = \{(o_1, o_2) \in \mathcal{O}_{st} \times \mathcal{O}_{st} \mid C(o_1) = c, C(o_2) \neq c\}$$

The University Of Sheffield.

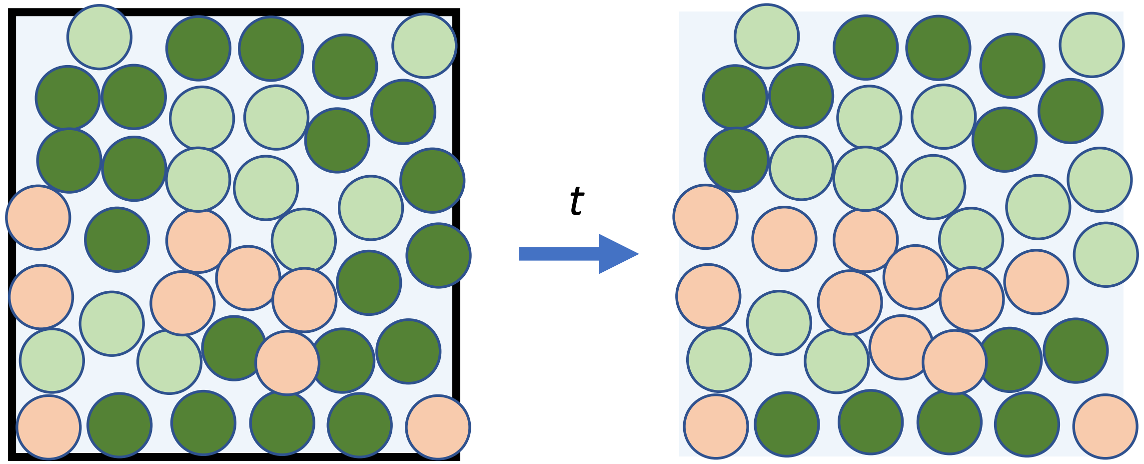
Leibovici, DG, and Claramunt. ‘On Integrating Size and Shape Distributions into a Spatio-Temporal Information Entropy Framework’. Entropy 21, no. 11 (2019): 1112

If no or not much spatial fragmentation Si concentrated in large patches (*i.e., dominance*)

Spatially: Is this area fragmented?

size & shape

with the `PsishENT` *framework*



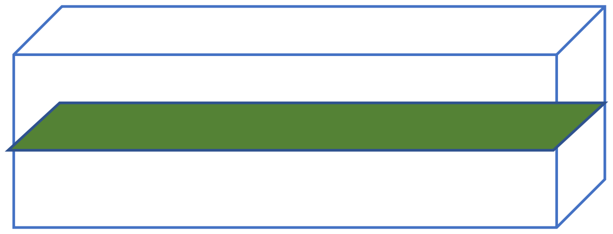
Focusing on forest

e.g.,

$H(Si \mid C = c)$ $H(Ti \mid C = c)$

$Si \times Ti$

c = Forest



$$H(Si, Ti) = H(Si) + H(Ti) - MI(Si, Ti) \quad \mid C = c$$

but also $D_{KL}(Si_{(C=c)} | Si_{(C \neq c)})$ *idem with* T_i *and* $D_{KL}(Si_{(C=c,t=t2)} | Si_{(C=c,t=t1)})$

Temporally: How much more (fragmented) since last year?

If no or not much time process Ti concentrated in large patches (*e.g.,* all period)

many Decompositions & spatial Maps

Topological characteristics

spatial patch — time patch — Categorical (fuzzy)

$$Si \quad Ti \quad C$$

$$\mathcal{S} \quad \mathcal{T} \quad \mathcal{C}$$

o

spatial location — time — categorical

- as part of global statistics

$$E_{stc} \subset E_{sitic} = \{o \in \mathcal{S} \times \mathcal{T} \times \mathcal{C} \mid \mathcal{C}(o) = c, Si(o) = si, Ti(o) = ti\}$$

$$E_{stc} = \{o \in \mathcal{S} \times \mathcal{T} \times \mathcal{C} \mid \mathcal{C}(o) = c, \mathcal{S}(o) = s, \mathcal{T}(o) = t\}$$

- as local statistics

(habitat) **fragmentation**:

WIKIPEDIA
The Free Encyclopedia

- Reduction in the total area **T+++ S+**
- Decrease of the interior **T+++ S++**
- Isolation of one area (fragment) from other areas **T+ S++++**
- Breaking up of one patch into several smaller patches **T+++ S+++**
- Decrease in size patch **T++++ S++**

Canonical characteristics

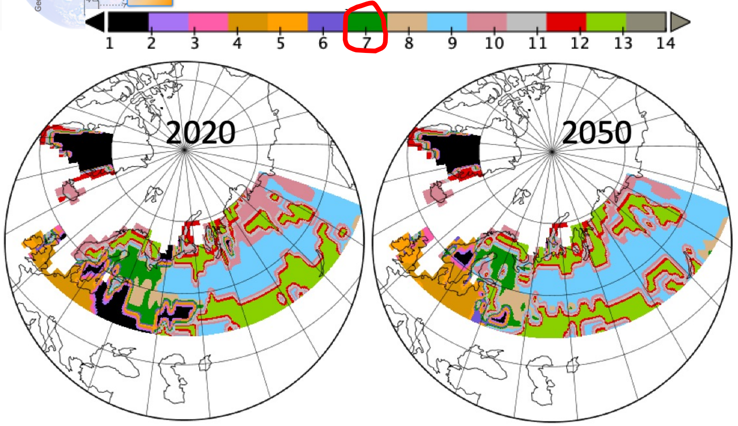
7 boreal needleleaf evergreen

```
pft7 (Si 1 %15) i
```

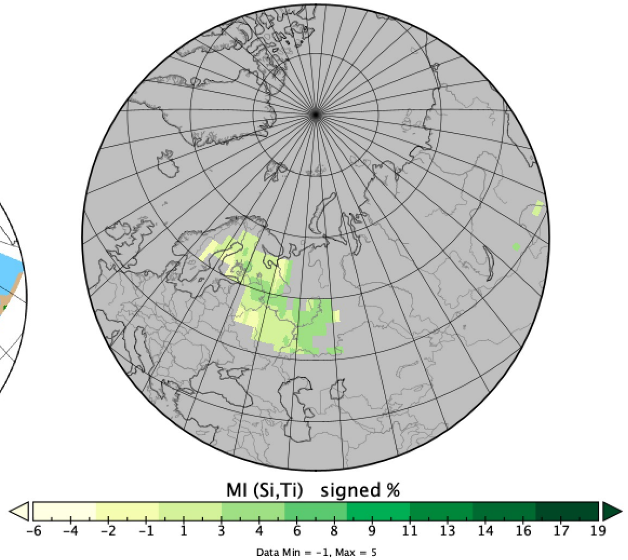
	2020	2050	2080 (+-2)
Si 1	5	32	7
Si 2	20	6	6
Si >2	0	0	16
Si >7	8	53	51
Si >25	0	0	143
Si >50	257	371	0
Si >100	208	0	0
	498	462	223

GeotRYcs

Dominant *vegetation*



2020 C= boreal needeleaf evergreen

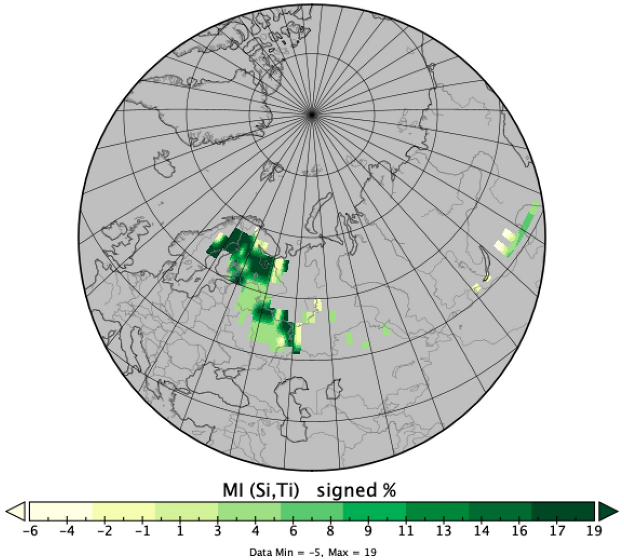


an example ...

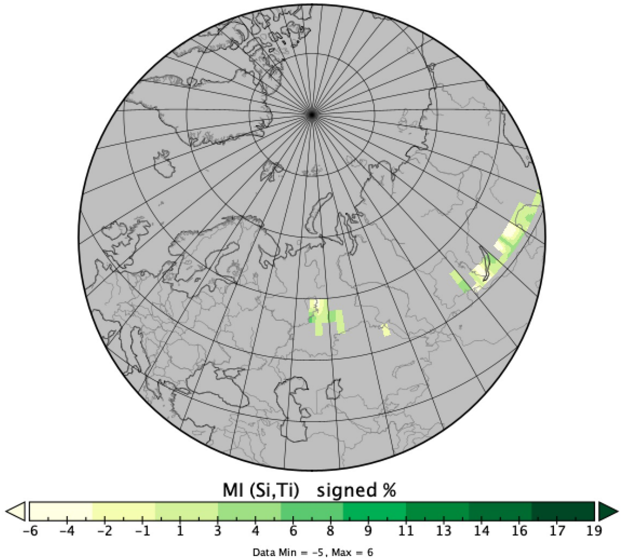
7 boreal needleleaf evergreen

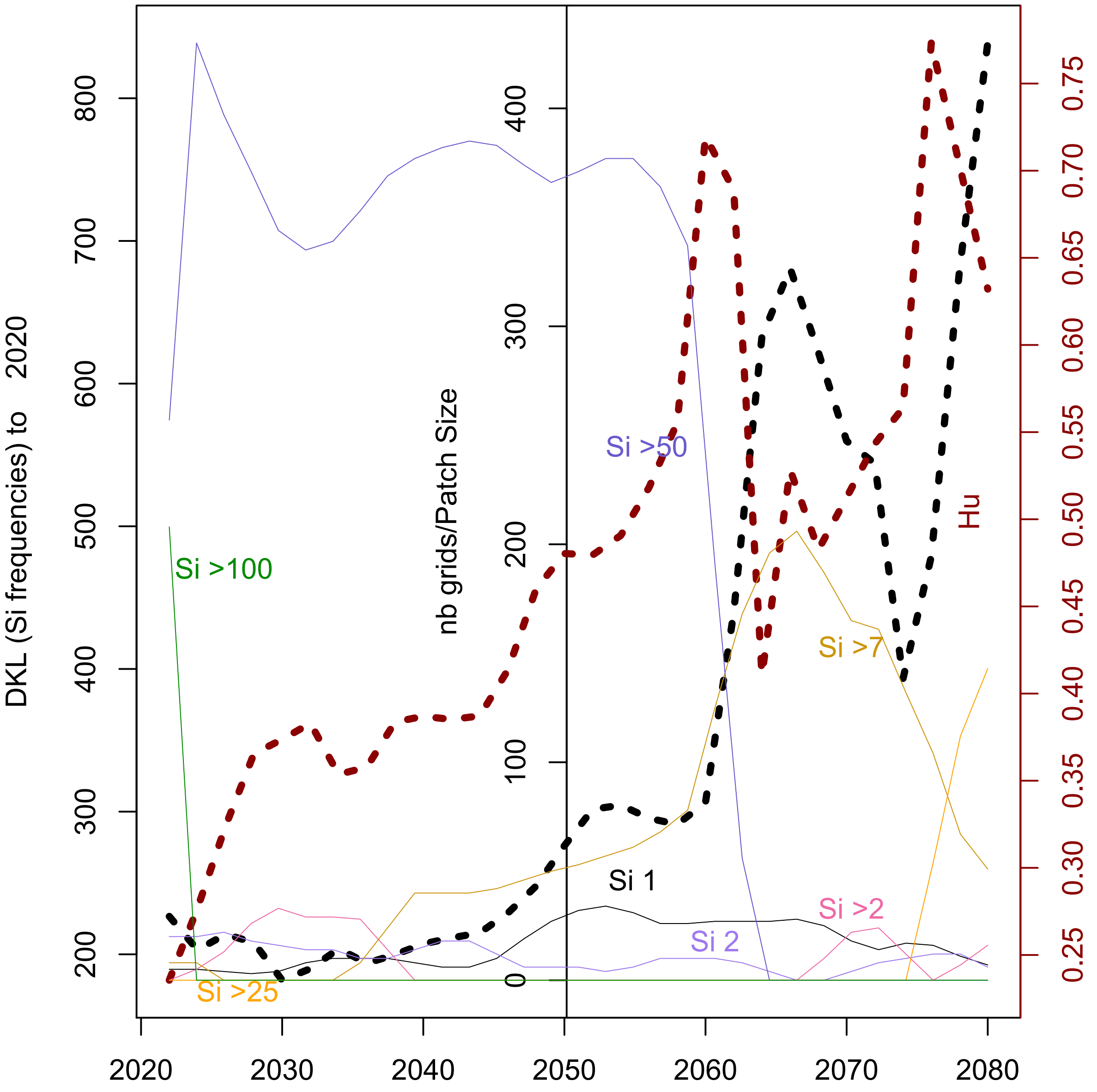
$H(Si \mid C = c)$

2050 C= boreal needeleaf evergreen



2080 C= boreal needeleaf evergreen





$D_{KL}(S_{i_{(C=c,t=t2)}} | S_{i_{(C=c,t=t1)}})$

7 boreal needleleaf evergreen

$MI(Si, Ti) \quad | C = c$

The University Of Sheffield.

ECOLE NAVALE

GeotRYcs

ECOLE NAVALE

fully symmetric co-occurrences

$$o_1, o_2, o_3 \in E_{stc}$$

are in co-occurrence of order $k = 3$,

if $max_{o,o' \in \{o_1,o_2,o_3\}} d(o,o') \leq d_\epsilon$

where $d()$ being the distance in $\mathcal{S} \times \mathcal{T} \times \mathcal{C}$

and d_ϵ a chosen collocation distance parameter

$$E_{stc} = \{o \in \mathcal{S} \times \mathcal{T} \times \mathcal{C} \mid \mathcal{C}(o) = c, \mathcal{S}(o) = s, \mathcal{T}(o) = t\}$$

$$E_{stc} \subset E_{sitic} = \{o \in \mathcal{S} \times \mathcal{T} \times \mathcal{C} \mid \mathcal{C}(o) = c, Si(o) = si, Ti(o) = ti\}$$

The University Of Sheffield.

Conclusion

- Fragmentation / entropy measures
- local and global / spatial & temporal
- route >>> categorical to spatial
- route >>> spatial to categorical

Spatial entropy is required

- **PsishENT** … add-on package (to come!)
- decompositions & geographical maps

My ResearchGate
My LinkedIn

https://wwwGeotRYcs.com

Appendix

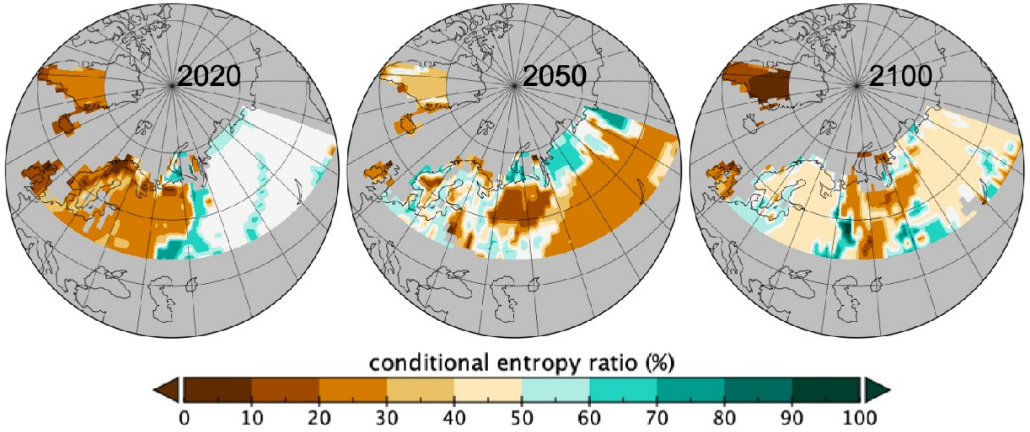


Figure 6. Map of the ratios to conditional entropy $H^u(C \mid Si)$ of Table 1 from occurring local patch sizes (ranges: 2020 2%–77%, 2050 2%–80%, 2100 0%–92%).

Similarly, in Figure 7 is represented the conditional entropy ratio for $H^u(Si \mid C)$ where local patches of C values were used to map the local effect. Overall over the 2020-2100 period, there was an increase in homogeneity as the conditional entropy is decreasing (see Table 1). Spatially there is an increase in homogeneity of patch sizes given the involved *ptfs* (C) in all areas, so either less variation in *pfts* or in their patch sizes.

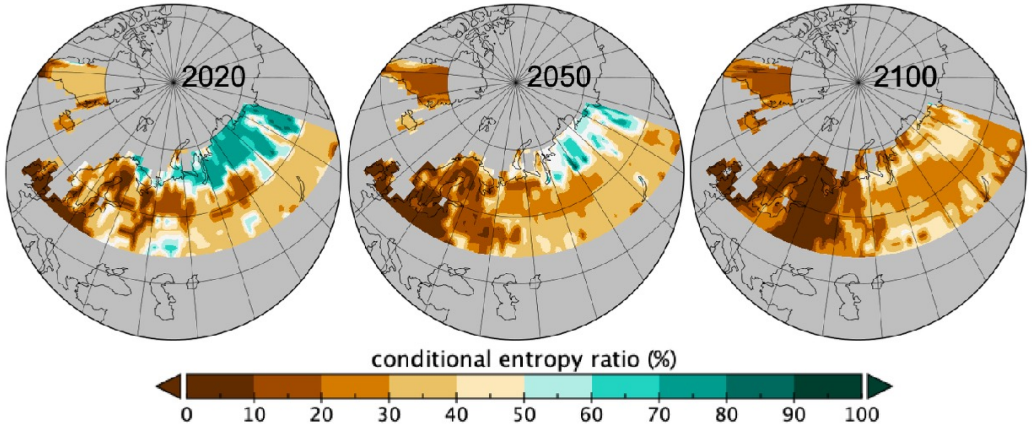
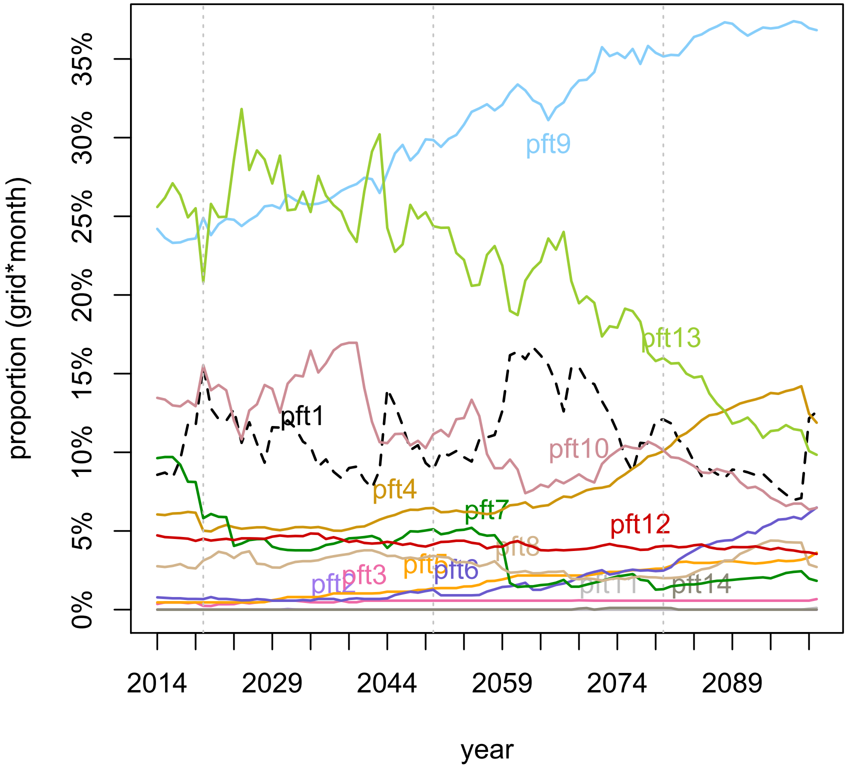


Figure 7. Map of the ratios to conditional entropy $H^u(Si \mid C)$ of Table 1 from occurring local patches of C (ranges: 2020 1%–87%, 2050 0%–89%, 2100 1%–67%).

Land Surface Modelling (ORCHIDEE scenario RCP85) fraction of

1 bare ground
2 tropical broadleaf evergreen
3 tropical broadleaf raingreen
4 temperate needleleaf evergreen
5 temperate broadleaf evergreen
6 temperate broadleaf summergreen
7 boreal needleleaf evergreen
8 boreal broadleaf summergreen
9 boreal needleleaf summergreen
10 C3 grass
11 C4 grass
12 NonVascular moss&lichen
13 boreal broadleaf shrubs
14 C3 arctic grass



Appendix

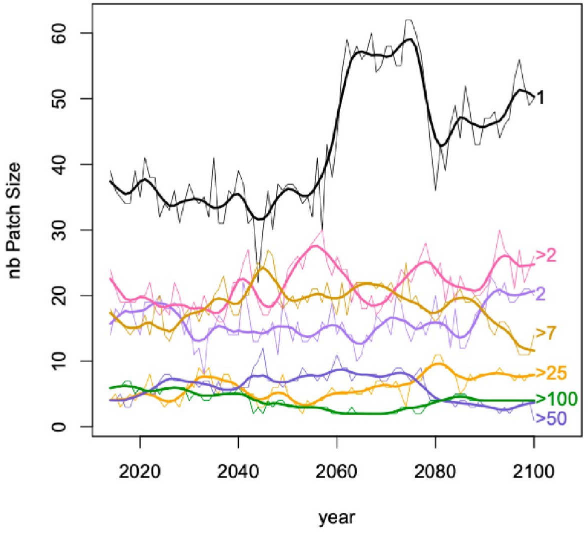
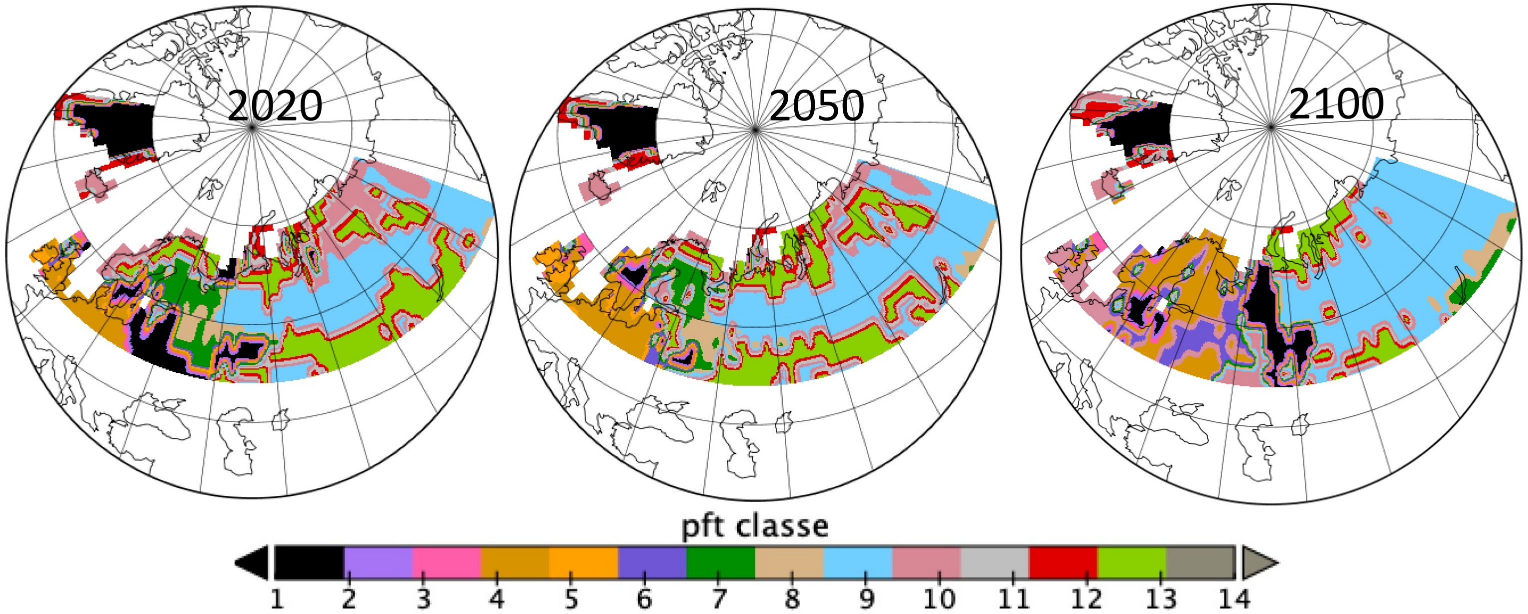


Figure 5. Frequencies of the 7 classes of spatial patches over 846 inland grid cells for all *pfts* where a 1 patch is a grid-cell with fraction > 15% (wider solid lines are smoother fit of the time series) in thinner lines.



Dominant *pft* per grid cell

The University Of Sheffield.

$$
\begin{aligned}
D_{KL}(p_{SiTi} \mid p_{Si} \otimes p_{Ti}) &= \sum_{st} p_{siti} log(p_{siti}/(p_{si}p_{ti})) =^{def} MI(Si,Ti) && (1)\\
&= H(Si) - H(Si \mid Ti) && (2)\\
&= \sum_{ti} p_{ti}(H(Si) - H(Si \mid Ti = ti)) && (3)\\
&= \sum_{si} p_{si}(H(Ti) - H(Ti \mid Si = si)) && (4)
\end{aligned}
$$

from frequencies DKL (1)

S7 \ T7	T<=2	T>2	T>5	T>15	T>25	T>35	T>45	T>55	T>65	T>75
S1	-0.02	-0.11	3.44	-0.47	-0.71	-0.97	-0.81	0.21	0.00	4.72
S2	0.85	0.00	-0.18	-0.16	4.85	0.00	-0.95	0.08	-0.18	2.40
S>5	0.74	-0.43	-1.32	1.88	0.68	-2.28	-4.21	5.51	7.36	4.75
S>15	-0.05	-0.19	-1.03	-0.74	-1.30	4.11	-1.91	0.22	6.02	3.59
S>25	-0.11	0.00	-2.08	3.89	1.24	1.57	-5.42	-0.49	4.38	2.68
S>35	0.00	-0.06	7.27	2.60	0.01	-0.47	-0.62	0.00	0.00	-0.18
S>45	0.31	3.31	6.74	2.12	-0.29	-0.52	-0.67	0.00	0.00	-0.22
S>55	0.00	-0.08	-0.35	-0.05	0.25	-0.07	1.55	0.02	-0.14	0.00
S>65	-0.13	-1.12	-4.69	-1.87	3.35	5.06	19.08	-0.60	-1.79	0.00
S>75	0.00	-0.75	-2.59	-0.80	1.26	3.75	9.20	-0.56	-1.16	0.00
S>85	0.00	0.28	0.84	-0.12	-0.28	0.14	0.27	-0.07	-0.09	-0.09
S>95	0.08	-0.11	4.66	-0.30	-0.62	0.05	0.12	-0.16	-0.20	-0.19
S>105	0.25	5.75	10.22	-0.94	-1.70	-0.52	-0.89	-0.47	-0.53	-0.52

from above (3) and (4)

#%

T<=2	T>2	T>5	T>15	T>25	T>35
0.73	2.87	-5.82	3.05	11.71	22.31
T>45	T>55	T>65	T>75		
39.33	3.89	12.97	8.95		

#%

S1	S2	S>5	S>15	S>25	S>35
2.77	4.19	-6.08	2.74	-5.18	6.64
S>45	S>55	S>65	S>75	S>85	S>95
5.43	3.13	45.66	25.63	1.48	5.03
S>105					
8.58					