A spatio-temporal characterization of tree development based on patchiness patterns

Pierre Fernique[1,2,6] Anaëlle Dambreville[2,3] Jean-Baptiste Durand[1,4]
Christophe Pradal[1,2] Pierre-Éric Lauri[5] Frédéric Normand[3] Yann Guédon[1,2]

[1]**Inria**, Virtual Plants

[2]**CIRAD**, UMR AGAP

[3]**CIRAD**, UPR HortSys

[4]**Université Grenoble Alpes**, LJK & **Inria**, Mistis

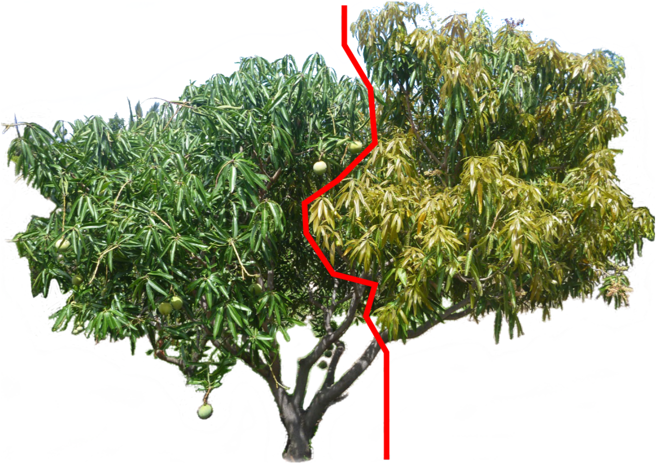
[5]**INRA**, UMR SYSTEM

[5]**INRA**, UMR MISTEA

2016/11/11

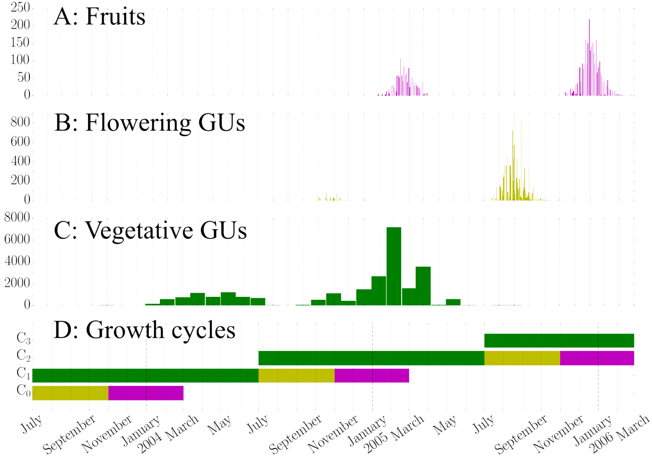
This presentation focuses on mango tree application

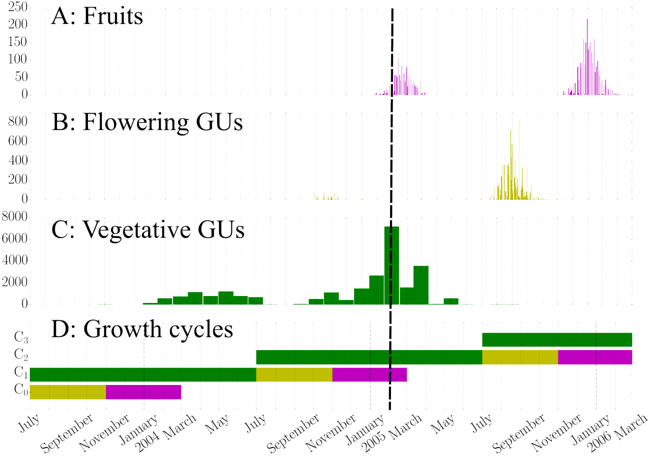
This presentation focuses on mango tree application

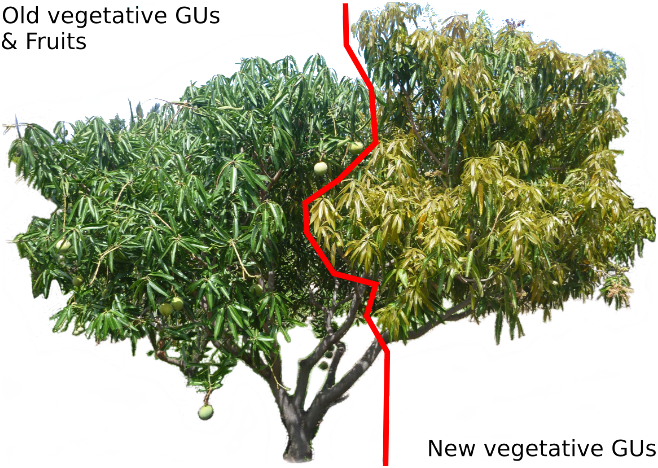


Introduction

¬ Mango tree




A: Fruits
250
200
150
100
50
0
B: Flowering GUs
800
600
400
200
0
C: Vegetative GUs
8000
6000
4000
2000
0
D: Growth cycles
C_3
C_2
C_1
C_0
July
September
November
January
2004 March
May
July
September
November
January
2005 March
May
July
September
November
January
2006 March


Old vegetative GUs
& Fruits
New vegetative GUs

Patchiness is characterized by clumps of either vegetative or reproductive GUs within the canopy [Chacko, 1986].

Concerns more or less large branching systems and entails various agronomic problems [Ramírez and Davenport, 2010].

Patchiness is characterized by clumps of either vegetative or reproductive GUs within the canopy [Chacko, 1986].
Concerns more or less large branching systems and entails various agronomic problems [Ramírez and Davenport, 2010].
The objective is unfold as follows:

1. Identifying patches components within a tree.
2. Identifying patches identities within a forest.
3. Identifying patchiness dynamics within a forest.

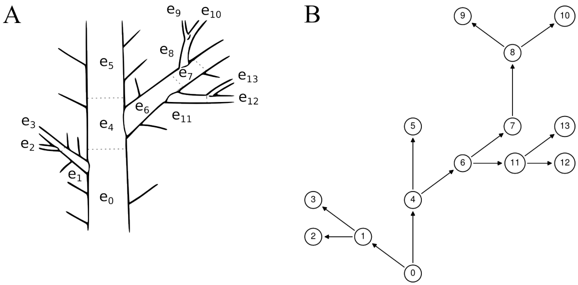
The objective is unfold as follows:

1. Identifying patches components within a tree.
2. Identifying patches identities within a forest.
3. Identifying patchiness dynamics within a forest.

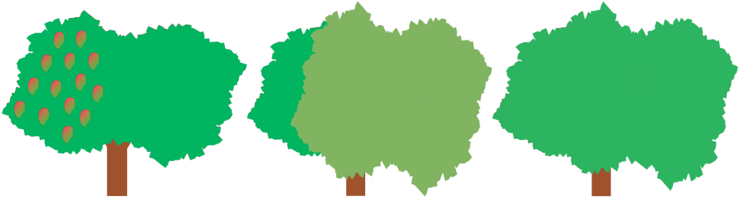
The experimental orchard was located at the Cirad research station in Saint-Pierre, Réunion Island [Dambreville et al., 2013].
7 cultivars, 5 mango trees by cultivar.
Described at the GU scale for 2 complete growth cycles.

Vertexes represent botanical entities.

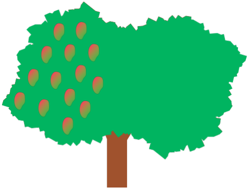
Edges encode either the temporal precedence of two botanical entities produced by the same meristem or the branching relationship between two botanical entities.

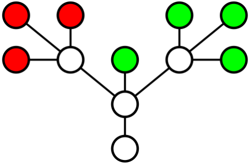


- *Patchiness* is varying according to time.

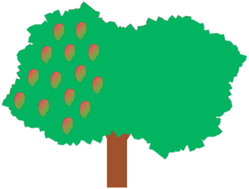


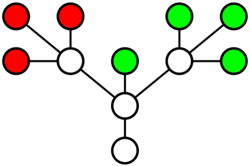
- *Patchiness* is varying according to time.
- *Patchiness* is mostly the result of what is observed on the canopy (i.e., tree-graph leaves) **at a given time**.



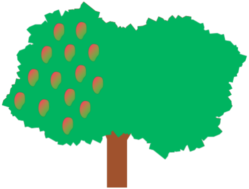


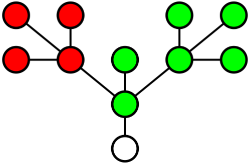
- *Patchiness* is varying according to time.
- *Patchiness* is mostly the result of what is observed on the canopy (i.e., tree-graph leaves) **at a given time**.
- *Patchiness* is induced by the presence of homogoneous tree subgraphes in a tree graph **at a given time**.



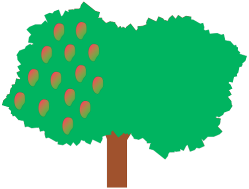


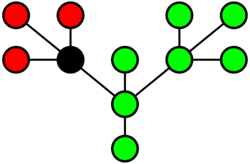
- *Patchiness* is varying according to time.
- *Patchiness* is mostly the result of what is observed on the canopy (i.e., tree-graph leaves) **at a given time**.
- *Patchiness* is induced by the presence of homogoneous tree subgraphes in a tree graph **at a given time**.



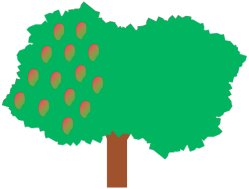


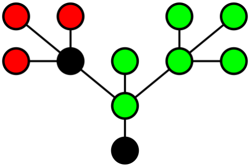
- *Patchiness* is varying according to time.
- *Patchiness* is mostly the result of what is observed on the canopy (i.e., tree-graph leaves) **at a given time**.
- *Patchiness* is induced by the presence of homogoneous tree subgraphes in a tree graph **at a given time**.



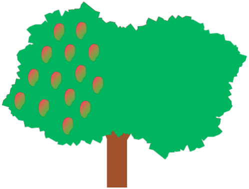


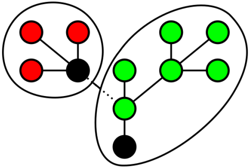
- *Patchiness* is varying according to time.
- *Patchiness* is mostly the result of what is observed on the canopy (i.e., tree-graph leaves) **at a given time**.
- *Patchiness* is induced by the presence of homogoneous tree subgraphes in a tree graph **at a given time**.



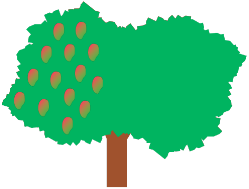


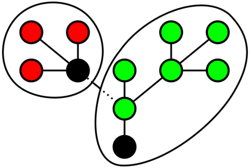
- *Patchiness* is varying according to time.
- *Patchiness* is mostly the result of what is observed on the canopy (i.e., tree-graph leaves) **at a given time**.
- *Patchiness* is induced by the presence of homogoneous tree subgraphes in a tree graph **at a given time**.





Identifying patches of mango trees ≡ Finding an **optimal tree quotienting** of labeled tree graphs.





Identifying patches of mango trees $\equiv$ Finding an **optimal tree quotienting** of labeled tree graphs.

A tree quotienting of a tree graph $T = (\mathcal{V}, \mathcal{E})$ is defined by:

- A set of vertices $\mathcal{P} \subseteq \mathcal{V}$.
- The function Π

$$\begin{aligned}\Pi : \mathcal{P} &\to \mathfrak{P}(\mathcal{V}) \\ p &\mapsto \left\{u \in \mathcal{V} : u \in De(p) \setminus \cup_{q \in \mathcal{P} \cap de(p)} De(q)\right\},\end{aligned}$$

where $de(v)$ is the descendant set of vertex v, $De(v) = de(v) \cup \{v\}$ and $\mathfrak{P}(\mathcal{V})$ is the power set of $\mathcal{V}$.

$\mathcal{P}$ denotes the set of change-points.

Identifying patches of mango trees $\equiv$ Finding an **optimal tree quotienting** of labeled tree graphs.

For $\mathcal{S} \subset \mathbb{N}$, let $(X_{i,j,k})_{k \in \mathcal{V}_{i,j}}$ be the $\mathcal{X}$-valued random process indexed by the random tree $i \in \mathcal{I} = \{0, \cdots, I\}$ observed at time $j \in \mathcal{J} = \{0, \cdots, J\}$ denoted by $T_{i,j} = (\mathcal{V}_{i,j}, \mathcal{E}_{i,j})$.

As for sequences, it's assumed that

$$\forall (i, i', j, j') \in \mathcal{I}^2 \times \mathcal{J}^2 \ : \ i \neq i' \vee j \neq j', \, (X_{i,j,k})_{k \in \mathcal{V}_{i,j}} \perp\!\!\!\perp (X_{i',j',k})_{k \in \mathcal{V}_{i',j'}} \, ,$$

$$\forall \left(p, p'\right) \subseteq \mathcal{P}^2 : p \neq p', \, (X_k)_{k \in \Pi(p)} \perp\!\!\!\perp (X_k)_{k \in \Pi(p')} \, ,$$

$$\forall p \subseteq \mathcal{P}, \, (X_k)_{k \in \Pi(p)} \sim \mathrm{Categorical}\left(\boldsymbol{\theta}_p\right) \text{ (i.i.d).}$$

Thus, log-likelihood $\mathcal{L}\left((x_k)_{k \in \mathcal{V}} ; \mathcal{P}, (\theta_p)_{p \in \mathcal{P}}\right)$ is easily computed for a given $\mathcal{P}$.

Finding an **optimal tree quotienting** defined by $\widehat{\mathcal{P}}$ of a labeled tree graph reduces to model selection using a penalized log-likelihood criterion.

$$\widehat{\mathcal{P}} = \underset{P \in \mathfrak{P}(\mathcal{V})}{\arg\max} \left\{ \mathcal{L}\left((x_k)_{k \in \mathcal{V}} ; \mathcal{P}, (\theta_p)_{p \in \mathcal{P}} \right) - \mathrm{pen}\left(|\mathcal{P}|\right) \right\},$$

To compute the penalization $\mathrm{pen}\left(|\mathcal{P}|\right)$ we used slope heuristic methods [Baudry et al., 2012] with

$$\mathrm{pen}\left(|\mathcal{P}|\right) = 2\,\widehat{\kappa}\,\log \frac{|\mathcal{V}|^{|\mathcal{P}|}}{|\mathcal{P}|!}.$$

Contrarilly to sequences, given the number of quotients $|\mathcal{P}|$, the inference of $\mathcal{P}$ cannot be done with exact methods.

By definition:

$$\mathcal{P}^{(0)} = \{r\},$$

with r the root of the tree graph, and

$$\mathcal{P}^{(1)} = P^{(0)} \cup \left\{ \arg\max_{v \in \mathcal{V}} \left\{ \mathcal{L}\left((x_k)_{k \in \mathcal{V}} ; \mathcal{P} \cup \{v\}, (\theta_p)_{p \in \mathcal{P} \cup \{v\}} \right) \right\} \right\},$$

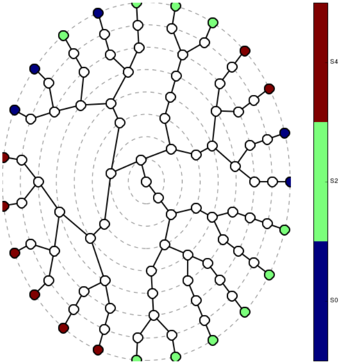
is optimal.

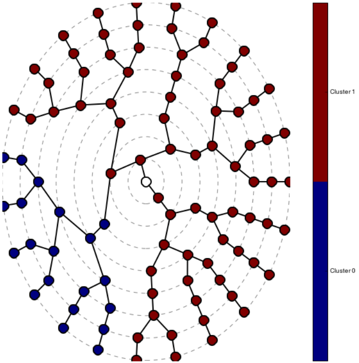
Identifying patches components within a tree

¬ Inference of tree-quotienting models

A split approach:

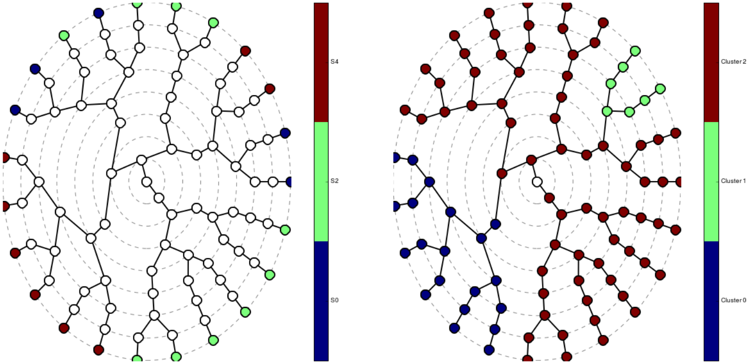
$$\mathcal{P}^{(k)} = P^{(k-1)} \cup \left\{ \underset{t \in \mathcal{T}}{\arg\max} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k-1)} \cup \{t\} \right), \theta_{\nu\left(\mathcal{P}^{(k-1)} \cup \{t\}\right)} \right) \right\} \right\},$$





A split approach:

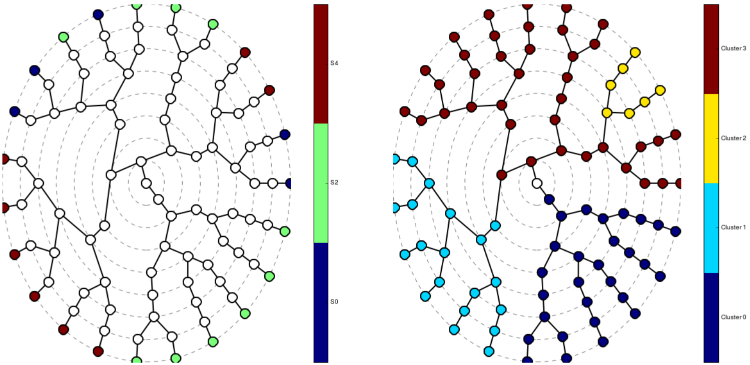
$$\mathcal{P}^{(k)} = P^{(k-1)} \cup \left\{ \underset{t \in \mathcal{T}}{\arg\max} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k-1)} \cup \{t\}\right), \theta_{\nu\left(\mathcal{P}^{(k-1)} \cup \{t\}\right)}\right) \right\} \right\},$$



Identifying patches components within a tree

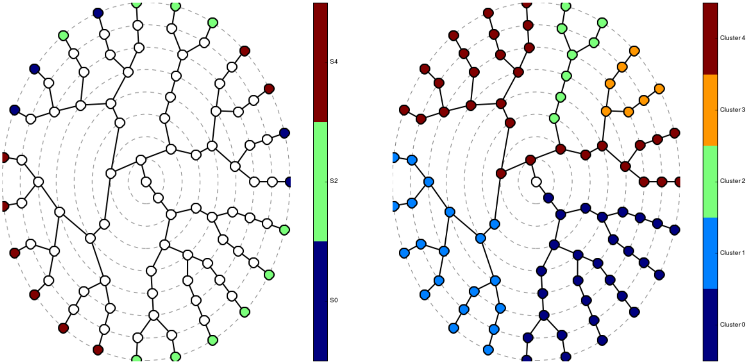
A split approach:

$$\mathcal{P}^{(k)} = P^{(k-1)} \cup \left\{ \underset{t\in\mathcal{T}}{\arg\max} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k-1)} \cup \{t\}\right), \theta_{\nu\left(\mathcal{P}^{(k-1)}\cup\{t\}\right)}\right)\right\}\right\},$$



A split approach:

$$\mathcal{P}^{(k)} = P^{(k-1)} \cup \left\{ \underset{t \in \mathcal{T}}{\arg\max} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k-1)} \cup \{t\} \right), \theta_{\nu\left(\mathcal{P}^{(k-1)} \cup \{t\}\right)} \right) \right\} \right\},$$

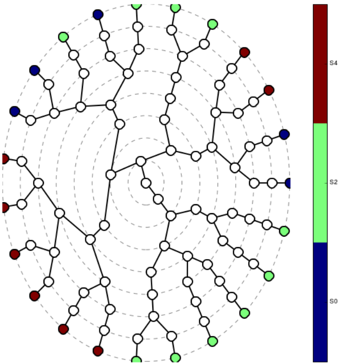


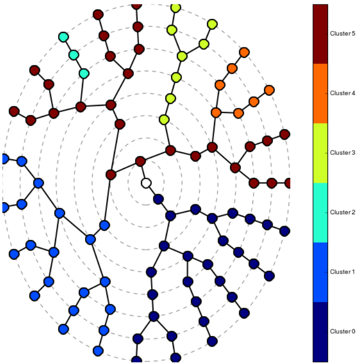
Identifying patches components within a tree

¬ Inference of tree-quotienting models

A split approach:

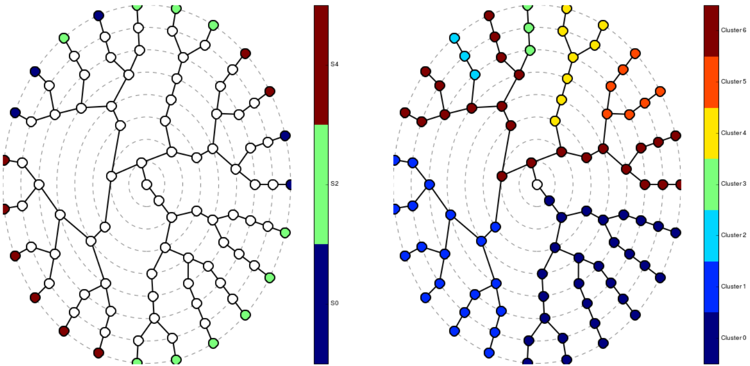
$$\mathcal{P}^{(k)} = P^{(k-1)} \cup \left\{ \underset{t\in\mathcal{T}}{\arg\max} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k-1)} \cup \{t\}\right), \theta_{\nu\left(\mathcal{P}^{(k-1)}\cup\{t\}\right)}\right)\right\}\right\},$$





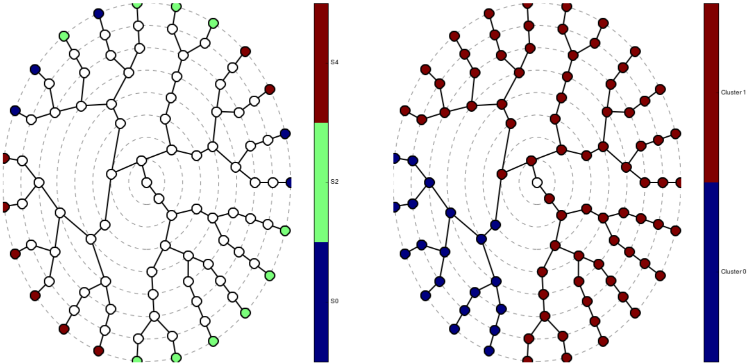
A split approach:

$$\mathcal{P}^{(k)} = P^{(k-1)} \cup \left\{ \arg\max_{t \in \mathcal{T}} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k-1)} \cup \{t\}\right), \theta_{\nu\left(\mathcal{P}^{(k-1)} \cup \{t\}\right)}\right) \right\} \right\},$$



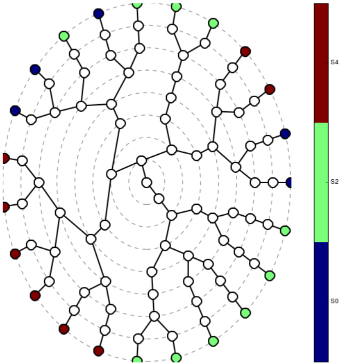
A merge approach:

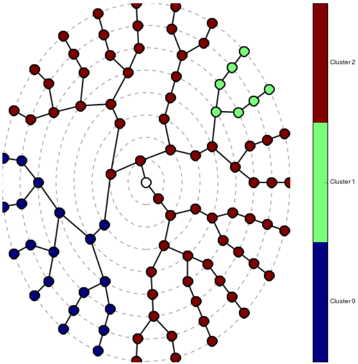
$$\mathcal{P}^{(k-1)} = P^{(k)} \setminus \left\{ \underset{t\in\mathcal{P}^{(\|)}}{\arg\max} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k)} \setminus \{t\}\right), \theta_{\nu\left(\mathcal{P}^{(k)}\setminus\{t\}\right)}\right) \right\} \right\},$$



A merge approach:

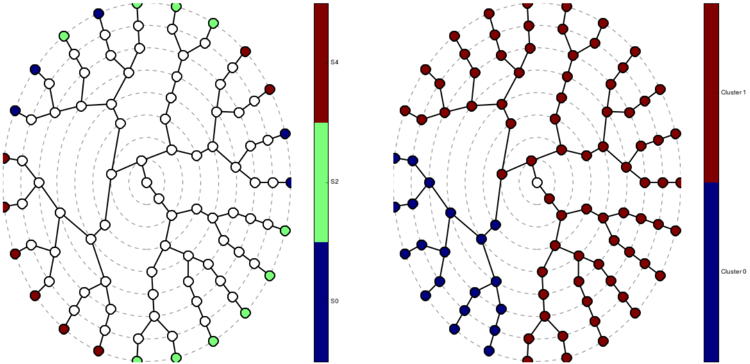
$$\mathcal{P}^{(k-1)} = P^{(k)} \setminus \left\{ \underset{t \in \mathcal{P}^{(\|)}}{\arg\max} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k)} \setminus \{t\} \right), \theta_{\nu\left(\mathcal{P}^{(k)} \setminus \{t\}\right)} \right) \right\} \right\},$$





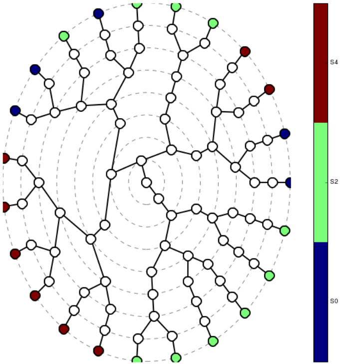
A merge approach:

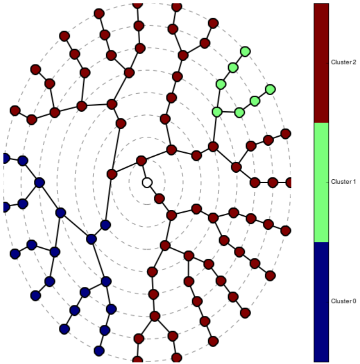
$$\mathcal{P}^{(k-1)} = P^{(k)} \setminus \left\{ \operatorname*{arg\,max}_{t \in \mathcal{P}^{(\|)}} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k)} \setminus \{t\} \right), \theta_{\nu\left(\mathcal{P}^{(k)} \setminus \{t\}\right)} \right) \right\} \right\},$$



A merge approach:

$$\mathcal{P}^{(k-1)} = P^{(k)} \setminus \left\{ \underset{t \in \mathcal{P}^{(\|)}}{\arg\max} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k)} \setminus \{t\} \right), \theta_{\nu\left(\mathcal{P}^{(k)} \setminus \{t\}\right)} \right) \right\} \right\},$$

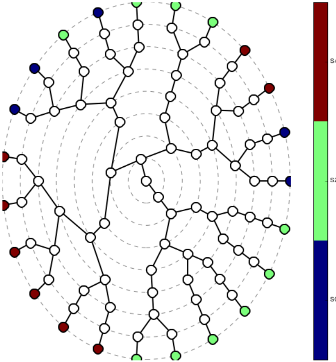


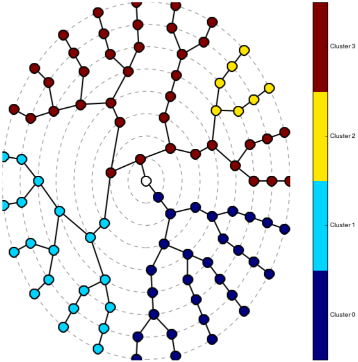


Identifying patches components within a tree

A merge approach:

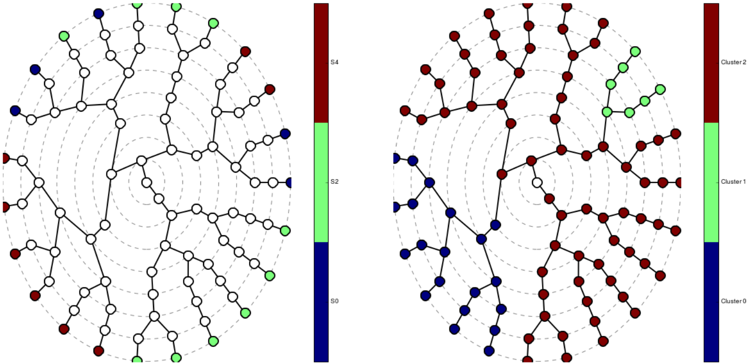
$$\mathcal{P}^{(k-1)} = P^{(k)} \setminus \left\{ \underset{t \in \mathcal{P}^{(\|)}}{\arg\max} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k)} \setminus \{t\} \right), \theta_{\nu\left(\mathcal{P}^{(k)} \setminus \{t\}\right)} \right) \right\} \right\},$$



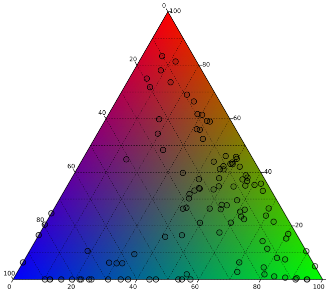


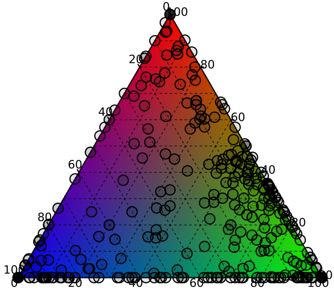
A merge approach:

$$\mathcal{P}^{(k-1)} = P^{(k)} \setminus \left\{ \underset{t \in \mathcal{P}^{(\|)}}{\arg\max} \left\{ \mathcal{L}\left(\bar{x}; \nu\left(\mathcal{P}^{(k)} \setminus \{t\} \right), \theta_{\nu\left(\mathcal{P}^{(k)} \setminus \{t\}\right)} \right) \right\} \right\},$$



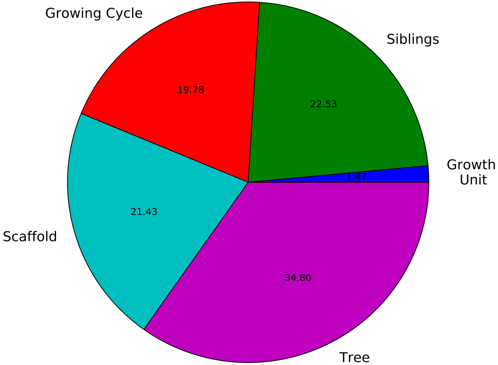
Composition of trees (left-hand circles) against the composition of patches (right-hand circles).





The RGB poles respectively represent trees that purely reproductive tree, purely vegetative and purely quiescent.

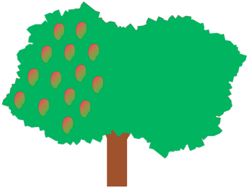
Scale of patch expression in reference to the different biological quotienting

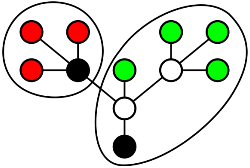


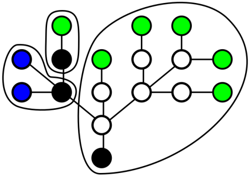
Identifying patches identities with a forest

¬ A reduction for interpretation

Manipulating tree quotienting of labeled trees is not an easy task.





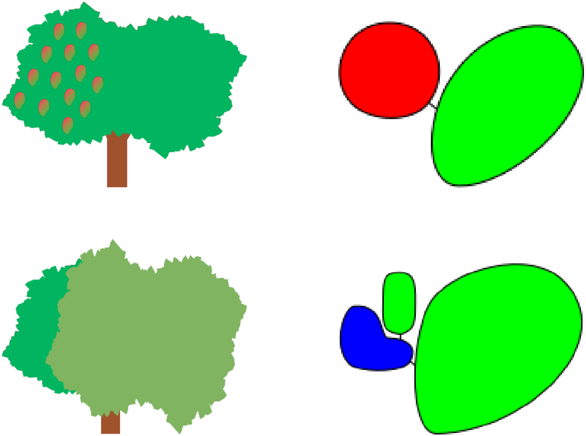


But manipulating labeled tree is.

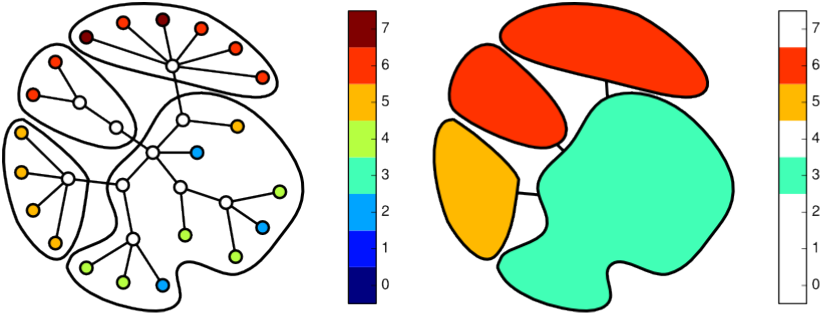
Identifying patches identities with a forest

¬ A reduction for interpretation

Manipulating tree quotienting of labeled trees is not an easy task.

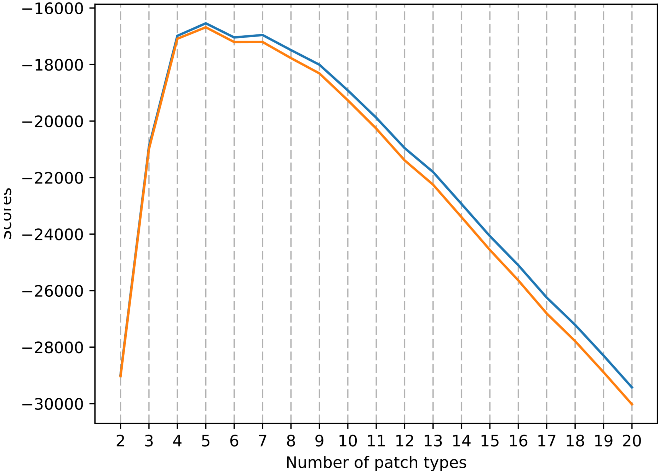


But manipulating labeled tree is.



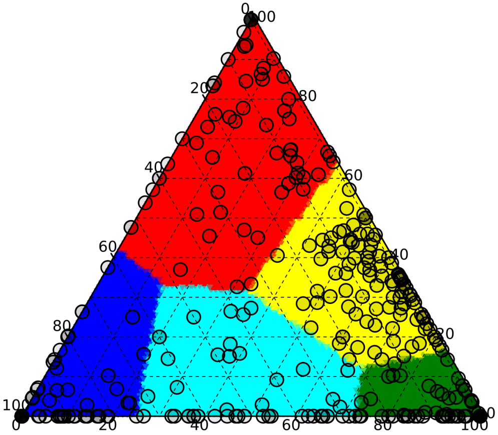
Using EM algorithm and MAP (Maximum A Posteriori) assignment of quotients of standard mixture models [McLachlan and Peel, 2000] such that vertices in same quotient are assigned to the same component [Picard et al., 2005].

5 different patch types found:



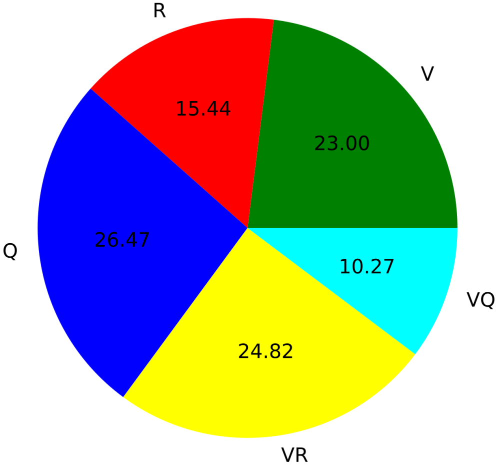
Using BIC, ICL criterion.

5 different patch types found:



The RGB poles respectively represent trees that purely reproductive tree, purely vegetative and purely quiescent.

5 different patch types found:

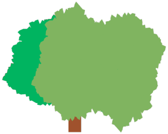


R (resp. V, Q) for reproductive (resp. vegetative, quiescent) patches.

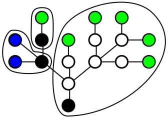
Identifying patchiness dynamics with a forest

¬ Problems

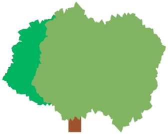
Identifying patches of mango trees ≡ Finding an optimal tree quotienting of labeled tree graphs.



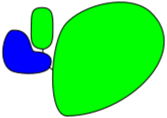
≡



Characterizing patches of mango trees ≡ Reducing an optimal tree quotienting into a labeled tree graph.

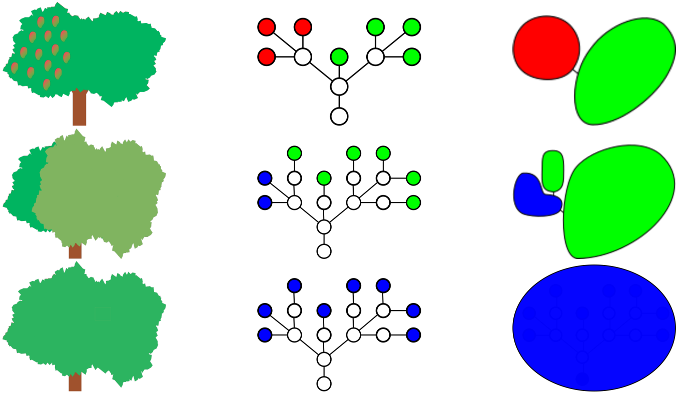


≡



This considerations are purely spatial !

How to take into account that a mango tree i has been observed a different given dates $j \in \{0, \cdots, J\}$?



What is the outcome of a patch ? What is the source of a patch ?

Let $H_i = (\mathcal{V}_i, \mathcal{E}_i)$ be a directed acyclic graph such as

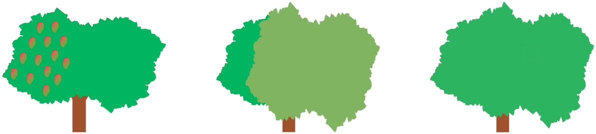
- Its vertex set is defined as follows

$$\mathcal{V}_i = \{(j, p) \in \mathcal{J} \times \mathcal{P}_{i,j}\}.$$

- Its edge set is defined as follows

$$\mathcal{E}_i = \left\{ \left((j, p), \left(j+1, p'\right)\right) \in \mathcal{V}_i^2 : \Pi_{i,j}(p) \cap \Pi_{i,j+1}\left(p'\right) \neq \emptyset \right\}.$$

H_i represents the patchiness dynamic of the mango tree i.



Let $H_i = (\mathcal{V}_i, \mathcal{E}_i)$ be a directed acyclic graph such as

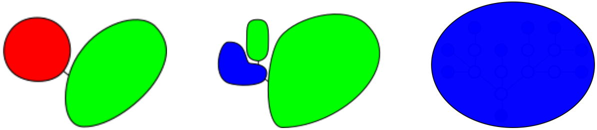
- Its vertex set is defined as follows

$$\mathcal{V}_i = \{(j, p) \in \mathcal{J} \times \mathcal{P}_{i,j}\}.$$

- Its edge set is defined as follows

$$\mathcal{E}_i = \left\{ \left((j, p), (j+1, p')\right) \in \mathcal{V}_i^2 : \Pi_{i,j}(p) \cap \Pi_{i,j+1}(p') \neq \emptyset \right\}.$$

H_i represents the patchiness dynamic of the mango tree i.



Let $H_i = (\mathcal{V}_i, \mathcal{E}_i)$ be a directed acyclic graph such as

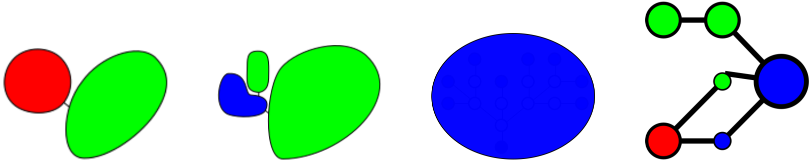
- Its vertex set is defined as follows

$$\mathcal{V}_i = \{(j,p) \in \mathcal{J} \times \mathcal{P}_{i,j}\}\,.$$

- Its edge set is defined as follows

$$\mathcal{E}_i = \left\{\left((j,p)\,,(j+1,p')\right) \in \mathcal{V}_i^2 : \Pi_{i,j}(p) \cap \Pi_{i,j+1}(p') \neq \emptyset\right\}.$$

H_i represents the patchiness dynamic of the mango tree i.



Let $H_i = (\mathcal{V}_i, \mathcal{E}_i)$ be a directed acyclic graph such as

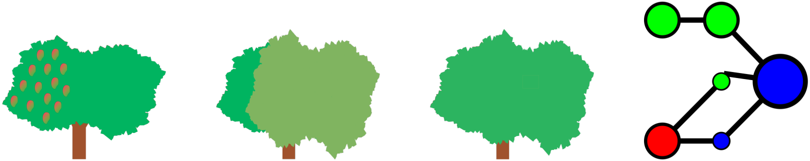
- Its vertex set is defined as follows

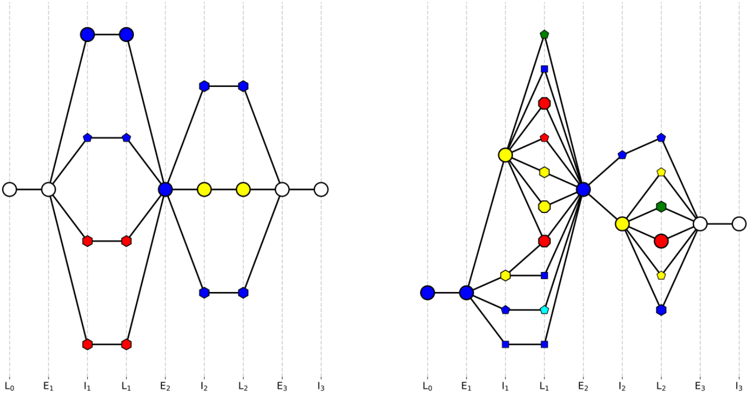
$$\mathcal{V}_i = \{(j, p) \in \mathcal{J} \times \mathcal{P}_{i,j}\}.$$

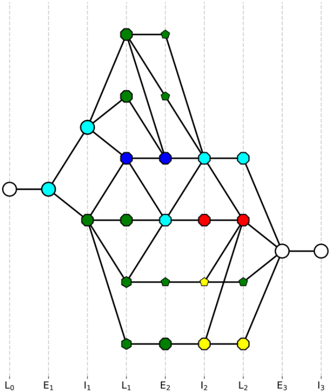
- Its edge set is defined as follows

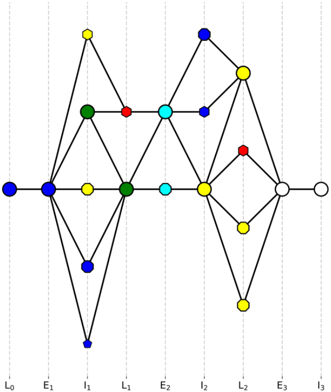
$$\mathcal{E}_i = \left\{ \left((j, p), (j+1, p')\right) \in \mathcal{V}_i^2 : \Pi_{i,j}(p) \cap \Pi_{i,j+1}(p') \neq \emptyset \right\}.$$

H_i represents the patchiness dynamic of the mango tree i.

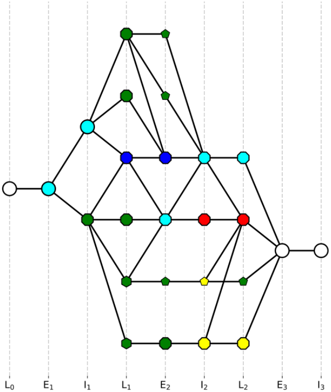



L0 E1 I1 L1 E2 I2 L2 E3 I3
L0 E1 I1 L1 E2 I2 L2 E3 I3


L_0
E_1
I_1
L_1
E_2
I_2
L_2
E_3
I_3


L_0
E_1
I_1
L_1
E_2
I_2
L_2
E_3
I_3

$$\mathcal{E}_i = \left\{ \left((j,p), \left(j+1,p'\right) \right) \in \mathcal{V}_i^2 : \Pi_{i,j}(p) \cap \Pi_{i,j+1}\left(p'\right) \neq \emptyset \right\}.$$

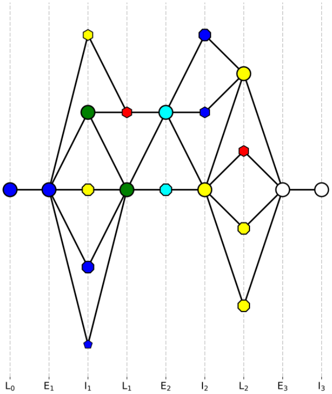


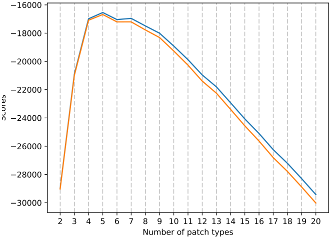
$$\mathcal{E}_i = \left\{ \left((j,p), \left(j+1,p'\right) \right) \in \mathcal{V}_i^2 : \left| \Pi_{i,j}(p) \cap \Pi_{i,j+1}\left(p'\right) \right| \geq \alpha \left| \Pi_{i,j}(p) \right| \right\}.$$

Identifying patchiness dynamics with a forest

¬ Discussion

MAP from mixture models for quotiented tree.





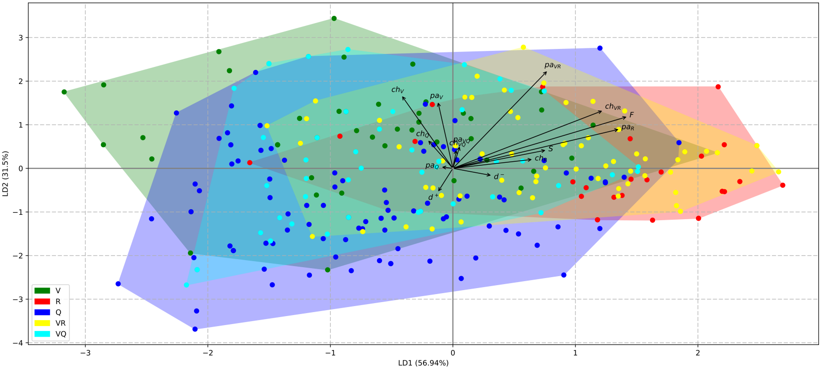
MAP from Markov models for quotiented tree such as the state duration is at most 1.

35 DAGs with a depth of 9.

Difficulties to detect particular events.

We used supervised classification tools on patch types given some summary statistics:

- The number of children denoted by d^+.
- The number of children that are of a given patch type denoted by $ch_.$.
- The number of parent denoted by d^-.
- The number of parent that are of a given patch type denoted by $pa_.$.
- The patch scale of expression denoted by S.



LD1 (56.94%)
LD2 (31.5%)
V
R
Q
VR
VQ

35 DAGs with a depth of 9.

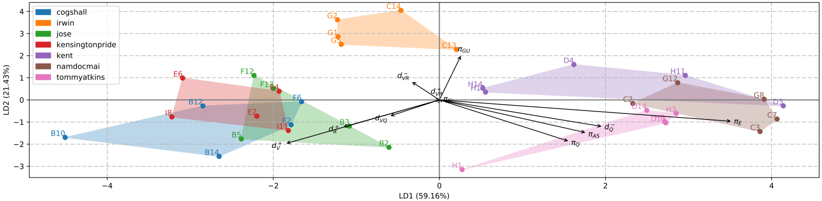
Difficulties to compare them.

We used supervised classification tools on cultivars given some summary statistics:

- The average number of children of a given patch type denoted by $d_{\cdot}^{-}$.
- The average number of parent of a given patch type denoted by $d_{\cdot}^{+}$.
- The proportion of patch in each scale of expression or each period denoted by $\pi_{\cdot}$.

Identifying patchiness dynamics with a forest

¬ Discussion



Such approaches are limited, we would rather perform unsupervised classification using models for random graphs.

Mixture model for random graphs [Daudin et al., 2008] but:

- undirected against directed edges (NoP),
- unlabeled against labeld vertices (NoP),
- $\mathcal{V}$ known against random but with $\mathcal{J}$ known.

Using ARMA-like models for the distribution of vertices along $\mathcal{J}$ [Weiß, 2008].

Limited number of vertices/states could lead to forward-backward algorithms for exact inference.

Conclusion

- Generalized the multiple change-point model from path-indexed data to tree-indexed data in order to detect homogeneous zones in tree-indexed data.

Conclusion

- Generalized the multiple change-point model from path-indexed data to tree-indexed data in order to detect homogeneous zones in tree-indexed data.
- Summarized time series of tree-indexed data into directed acyclic graphs.

Conclusion

- Generalized the multiple change-point model from path-indexed data to tree-indexed data in order to detect homogeneous zones in tree-indexed data.
- Summarized time series of tree-indexed data into directed acyclic graphs.
- This has been illustrated with the problem of patchiness within mango trees. Nevertheless, the methodology is not restricted to categorical data on tree-indexed data.

Conclusion

- Patches seem to be well detected with our heuristic.

Conclusion

- Patches seem to be well detected with our heuristic.
- Some work is remaining concerning interpretability of directed acyclic graphs.

Conclusion

- Patches seem to be well detected with our heuristic.
- Some work is remaining concerning interpretability of directed acyclic graphs.
- Apply this methodology to study patchiness in other species or other phenomena (mango tree asynchronisms)?

References I

Baudry, J.-P., Maugis, C., and Michel, B. (2012).
Slope heuristics: overview and implementation.
Statistics and Computing, 22(2):455–470.

Chacko, E. (1986).
Physiology of vegetative and reproductive growth in mango (*Mangifera indica* L.) trees.
In *Proceedings of the First Australian Mango Research Workshop*, volume 1, pages 54–70. CSIRO Australia, Melbourne.

Dambreville, A., Lauri, P.-É., Trottier, C., Guédon, Y., and Normand, F. (2013).
Deciphering structural and temporal interplays during the architectural development of mango trees.
Journal of experimental botany, 64(8):2467–2480.

References II

Daudin, J.-J., Picard, F., and Robin, S. (2008).
A mixture model for random graphs.
Statistics and computing, 18(2):173–183.

McLachlan, G. and Peel, D. (2000).
Finite mixture models.
Wiley New York.

Picard, F., Robin, S., Lavielle, M., Vaisse, C., and Daudin, J.-J. (2005).
A statistical approach for array CGH data analysis.
BMC bioinformatics, 6(1):27.

Ramírez, F. and Davenport, T. L. (2010).
Mango (*Mangifera indica* L.) flowering physiology.
126(2):65–72.

References III

Weiß, C. H. (2008).
Thinning operations for modeling time series of counts—a survey.
AStA Advances in Statistical Analysis, 92(3):319–341.