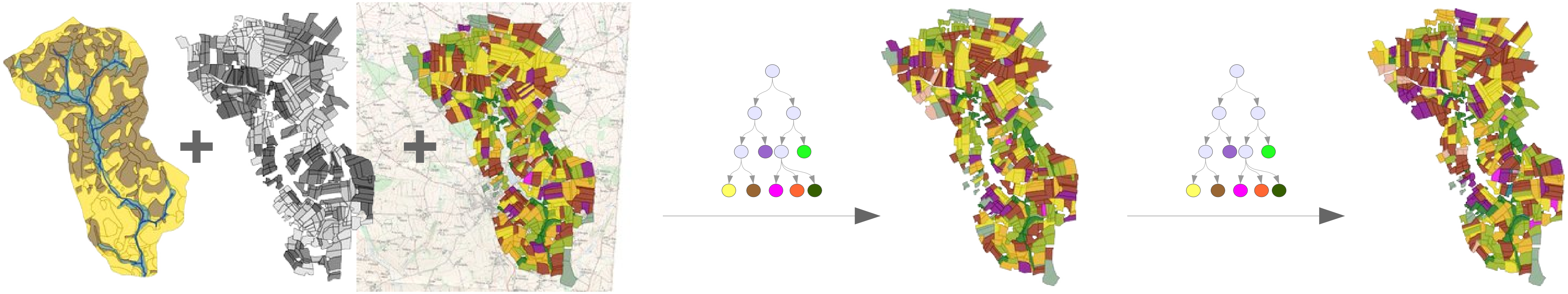
RTP
PAYSAGE
ENVIRONNEMENT



Stochastree,

a crop transition model based on stochastic decision trees, which integrates agronomic constraints.

L. Sorel, C. Walter, P. Durand, **V. Viaud (speaker)**

Rationale 1/2

Environmental assessment of agrosystems

Assessment of alternate agricultural practices:

- landcover rules the location of practices over the landscape
- the spatial distribution of practices influences the ecological response of the landscape
- organic matter transfer and transformation processes have dynamics spanning over decades in 1st order catchments

-> scenario-based analyses requires **spatialised** and **long-term** datasets and modelling approaches

Rationale 2/2

Crop transition modelling

Several approaches exist, but…

- most models focus either on crop succession or spatial structure
- their tuning often require strong expertise, large datasets, arbitrary decision
- (complexity and validation?)

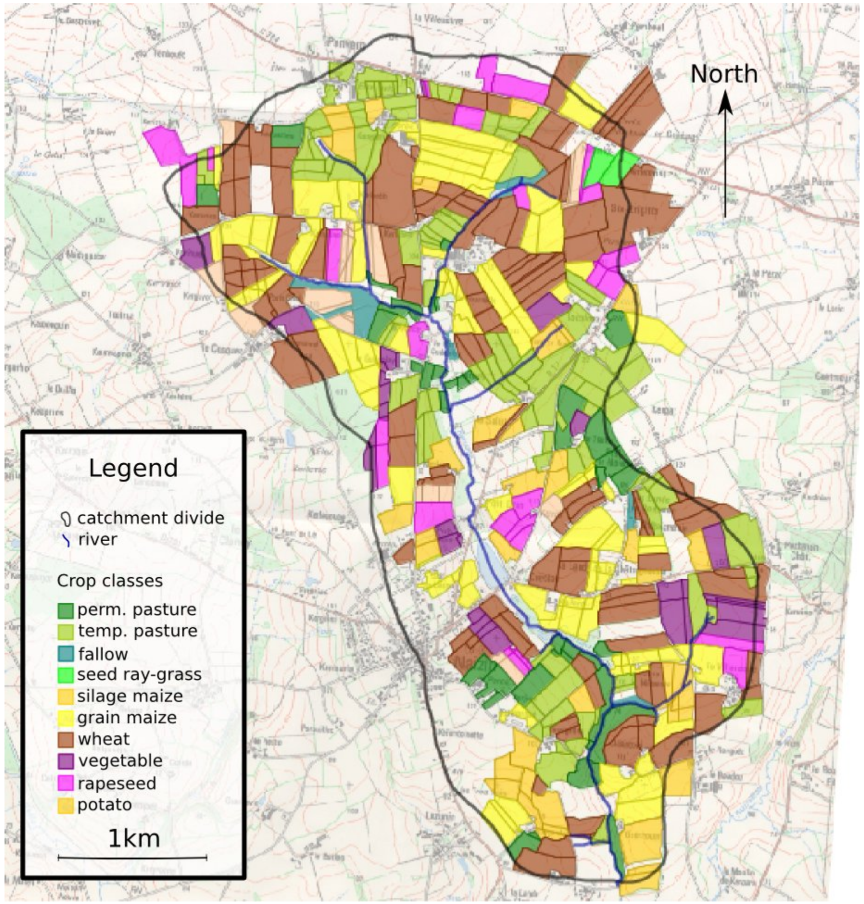
Objective

To stochastically simulate crop transition while accounting for agronomic, spatial and temporal driving factors

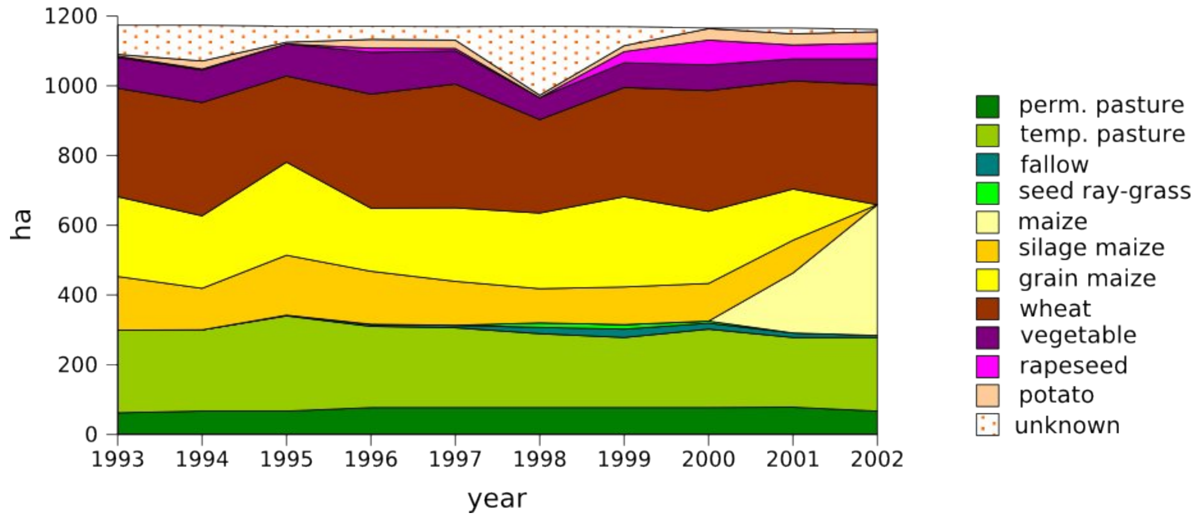
- datamining approach to identify the drivers (1993-98 dataset)
- validation of the model and simulations (1998-2002 dataset)
- comparison with a 1st order Markov chain model (transition probability matrix)

The case study landscape – landcover change drivers

Landcover dynamics



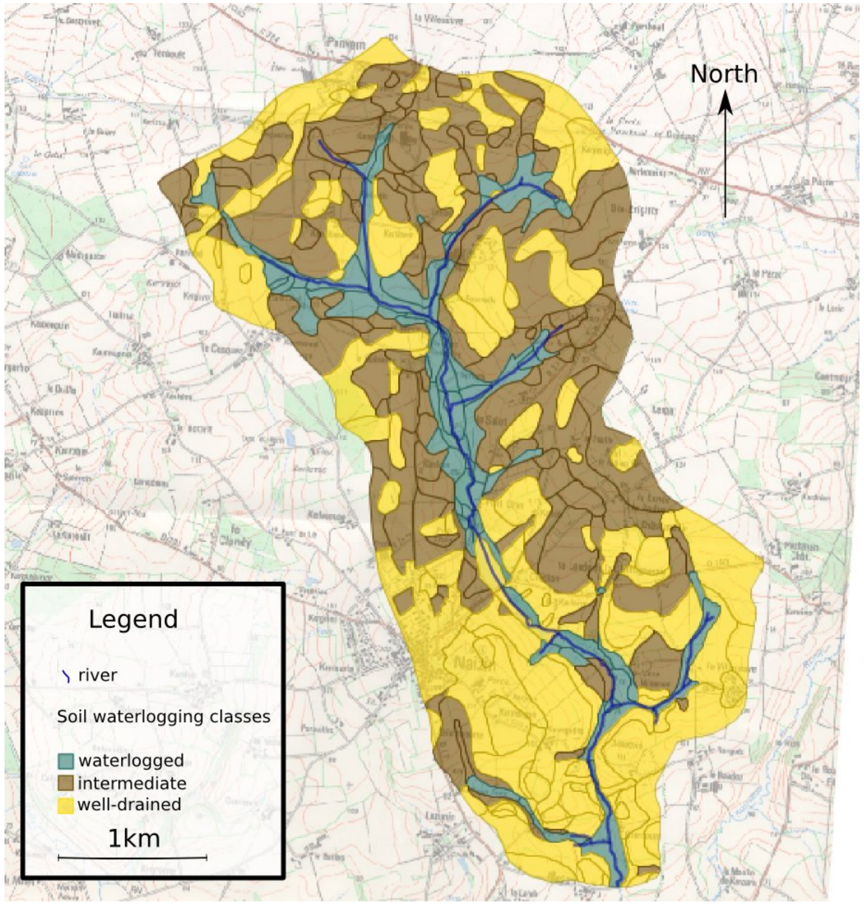
(year 2000)

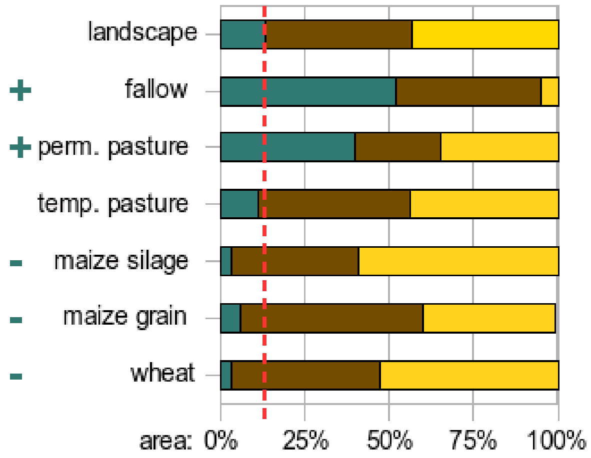


Naizin catchment (Morbihan, France)

- 12km², 80% dedicated to intensive farming
- farm-types (pigs, dairy cows, mix) are characterised by specific crop area proportions (≈production objectives)
- cultivated areas are stationary
- field patterns are assumed static, few exchanges of fields between farmsteads

Soil waterlogging

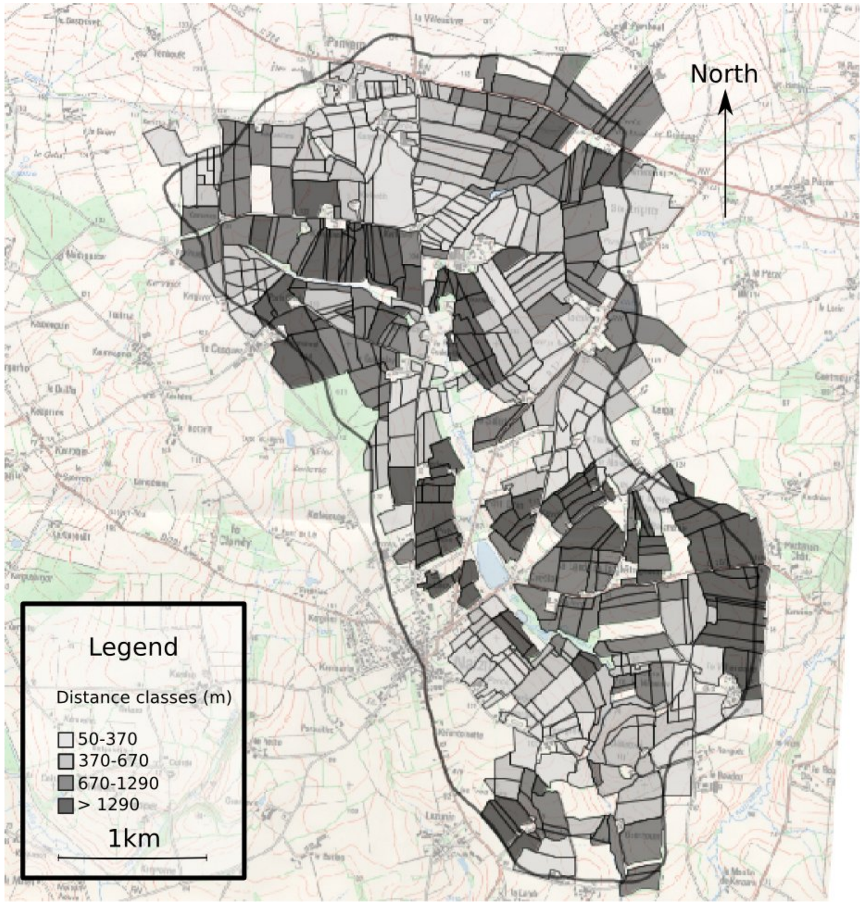




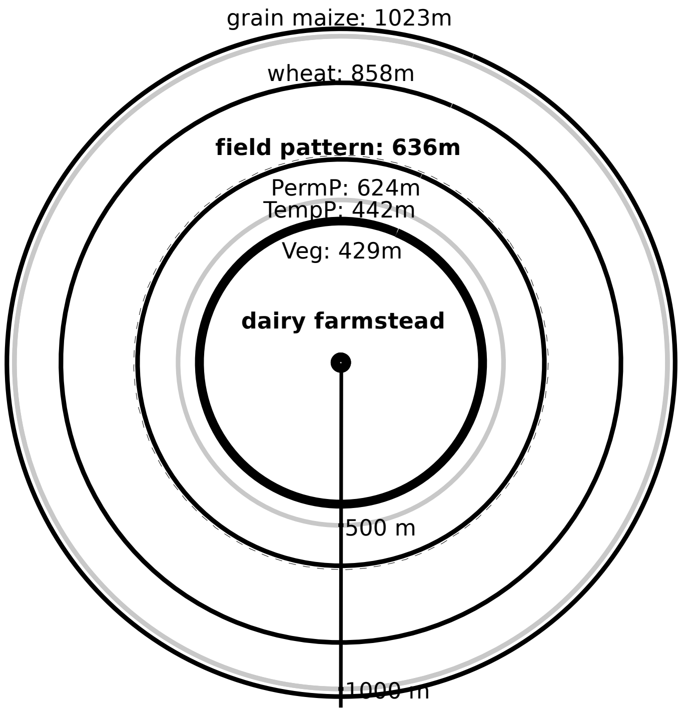
Crop allocation on soil waterlogging classes

- perm. pastures and fallows on waterlogged soils
- cereals on well-drained soils

Spatial distribution of crops around the farmstead



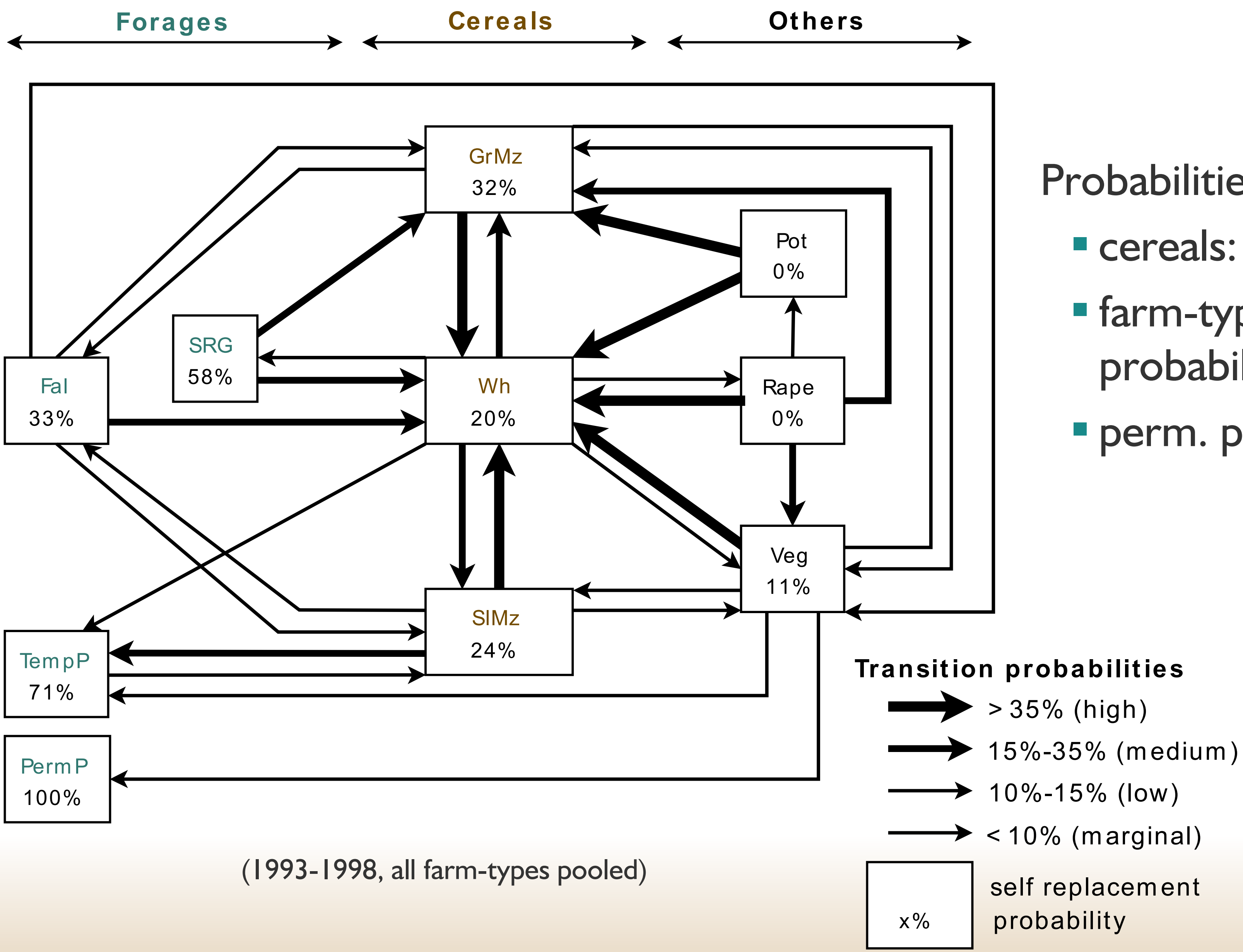
(year 2000)



Area-weighted distance (AWD) between the fields and the farmsteads of a given farm-type

- dairy field pattern is strongly spatially structured
- pastures are close to the farmstead, cereals are further

Effect of preceding crop on the transition process



(1993-1998, all farm-types pooled)

Probabilities learned over 1993-98

- cereals: main transition target
- farm-types have specific probabilities
- perm. pastures=growing state !

The *Rotomatrix* model

An empirical construction (1993-1998 annual transitions)

- generic transition probability matrix

landuse N	landuse N+1 *perm. pasture*	*temp. pasture*	*fallow*	*seed ray-grass*	*silage maize*	*grain maize*	*wheat*	*vegetable*	*rapeseed*	*potato*
perm. pasture	1									
temp. pasture		0.71	0.01		0.12	0.03	0.08	0.05		0.01
fallow			0.33		0.13	0.13	0.27	0.13		
seed ray-grass				0.58		0.25	0.17			
silage maize		0.2	0.02		0.24	0.03	0.4	0.09	0.01	0.01
grain maize		0.03	0.02		0.02	0.32	0.47	0.12		0.02
wheat		0.13	0.01	0.01	0.15	0.33	0.2	0.14	0.01	0.04
vegetable	0.02	0.1	0.01		0.14	0.14	0.45	0.11		0.03
rapeseed						0.3	0.4	0.2		0.1
potato		0.08			0.1	0.36	0.38	0.08		

- dairy transition probability matrix (farm-type specific matrix)

landuse N	landuse N+1 *perm. pasture*	*temp. pasture*	*silage maize*	*grain maize*	*wheat*	*vegetable*	*rapeseed*	*potato*
perm. pasture	1							
temp. pasture		0.73	0.15	0.01	0.05	0.05	0.01	0.01
silage maize		0.25	0.25	0.03	0.39	0.06	0.01	
grain maize		0.09	0.17	0.04	0.57	0.04		0.09
wheat		0.33	0.32	0.08	0.23	0.03		0.02
vegetable	0.03	0.31	0.23	0.09	0.31	0.03		
rapeseed					0.75			0.25
potato		0.29	0.07		0.64			

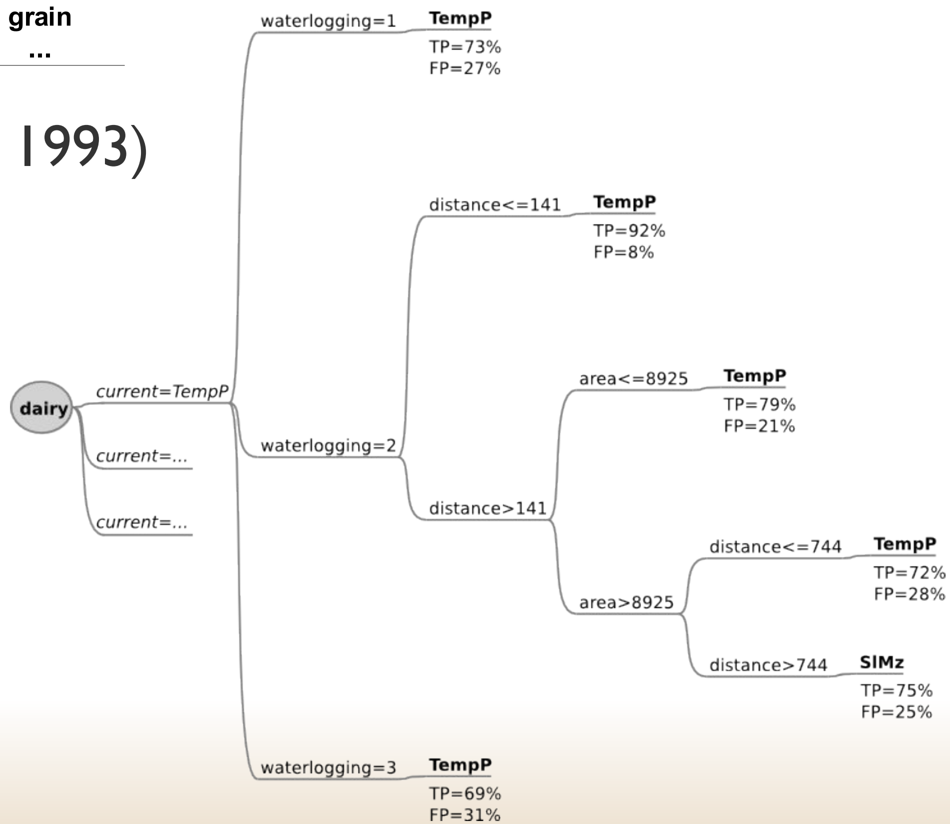
The *Stochastree* model

A datamining construction (1993-1998 annual transitions)

- structure of the learning dataset:

Field characteristics					
farm-type	*area (m²)*	*distance (m)*	*waterlogging*	*current crop*	***next crop***
dairy	9570	183	2	grain	**temppasture**
dairy	19722	561	1	maizesilage	**grain**
dairy	13155	578	2	temppasture	**grain**
dairy	47005	198	1	permpasture	**permpasture**
dairy	19664	5211	2	maizegrain	**potato**
pig	6739	1776	3	fallow	**grain**
...	...	...	...	...	**...**

- deterministic tree (C4.5 – Quinlan, 1993) (T/FP: true/false positive instances)



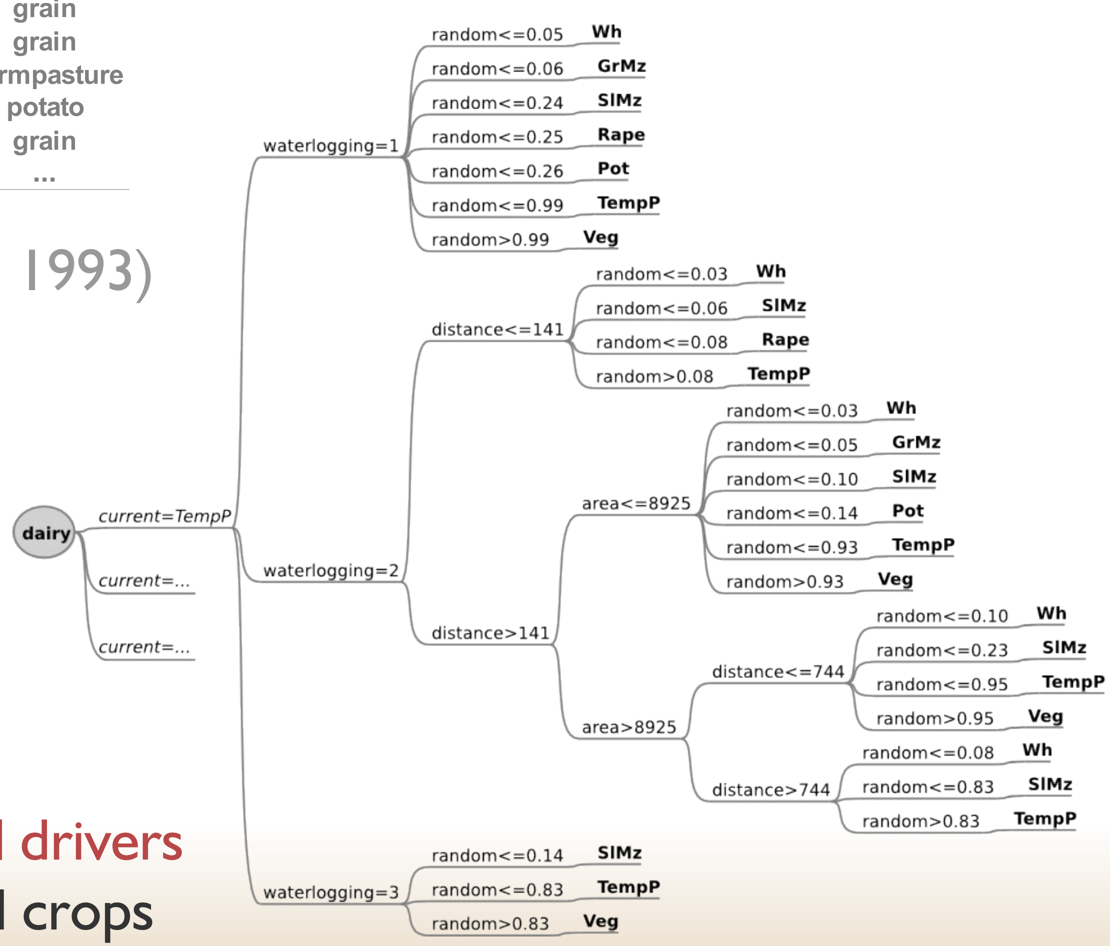
The *Stochastree* model

A datamining construction (1993-1998 annual transitions)

- structure of the learning dataset:

Field characteristics					
farm-type	*area (m²)*	*distance (m)*	*waterlogging*	*current crop*	***next crop***
dairy	9570	183	2	grain	**temppasture**
dairy	19722	561	1	maizesilage	**grain**
dairy	13155	578	2	temppasture	**grain**
dairy	47005	198	1	permpasture	**permpasture**
dairy	19664	5211	2	maizegrain	**potato**
pig	6739	1776	3	fallow	**grain**
...	...	...	...	...	**...**

- deterministic tree (C4.5 – Quinlan, 1993)
- TP and FP proportions at tree leaves were used as probabilities



- "current crop" is always the first tested attribute (temporal driver)
- tree structure integrates the spatial drivers between the current and predicted crops on a case per case basis

Model comparison (short term)

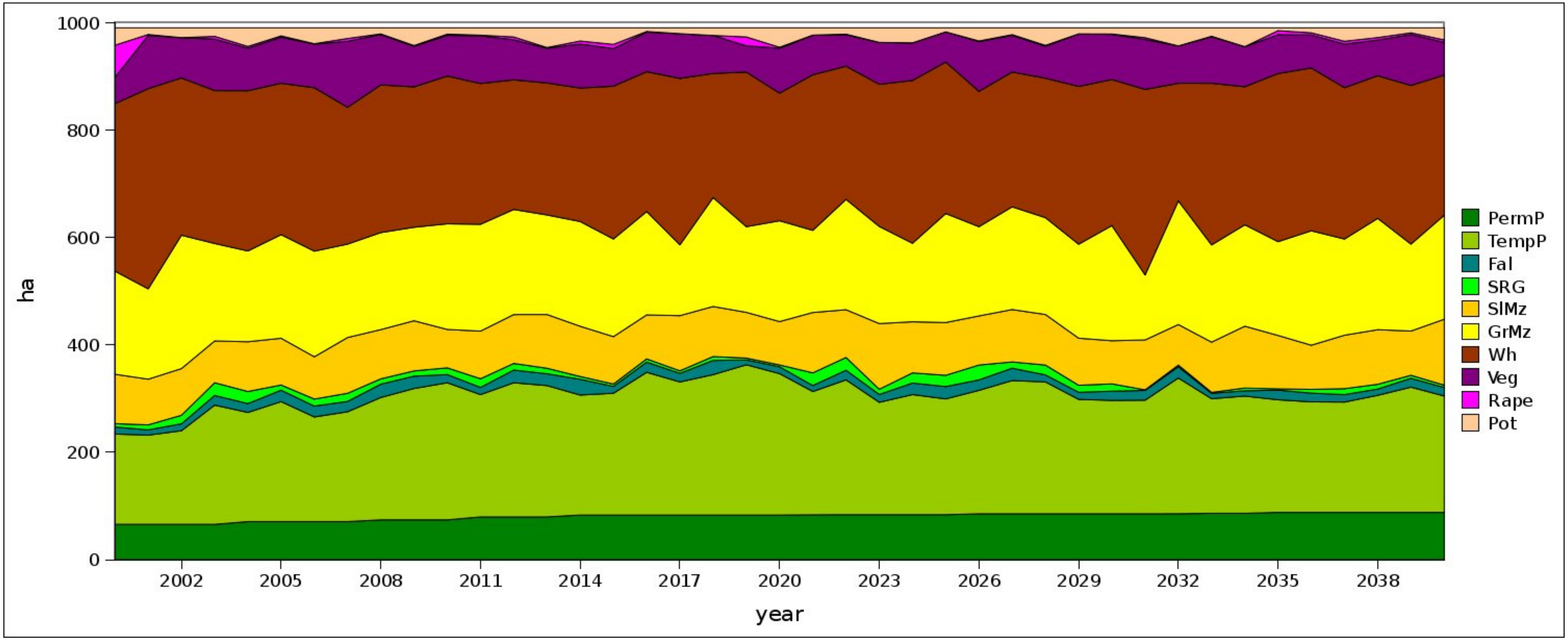
Prediction rate on a validation dataset (1998-2002 annual transition)

	simulation method				
	Rotomatrix		*Stochastree*		*random*
	generic	*specific*	*generic*	*specific*	
expected crop	**successfull prediction rates (%)**				
perm. pasture	97	97	97	97	10
temp. pasture	52	54	52	55	10
fallow	21	31	22	32	10
silage maize	16	18	14	17	9
grain maize	30	32	34	34	11
wheat	36	39	38	41	9
vegetable	13	10	12	12	10
rapeseed	0	0	1	0	10
potato	3	5	4	7	10
(all)	**38**	**40**	**39**	**41**	**10**

- stochastic trees and matrices perform better than pure random classifier
- farm-type specific matrices/trees perform usually better than generic ones (temp. pastures, fallow, cereals, potato)
- minor crops have low prediction rates
- no significant difference between *Rotomatrix* and *Stochastree*

Long term simulations

Landcover area dynamics (farm-type specific decision trees)



- initial state = year 2000, duration = 40 years
- reproduction of stationary crop areas
- similar ability of *Rotomatrix* and *Stochastree* to reproduce farm-type production objectives

Long term simulations

Crop allocation on the waterlogged soil class (13% landscape area)

crop class	observed area % on waterlogged soils	difference between observed and simulated area %			
		Rotomatrix		Stochastree	
		generic	specific	generic	specific
fallow	*52%*	**-43%**	**-39%**	**-24%**	**-9%**
temp. pasture	*8%*	-1%	**-3%**	-1%	**-4%**
silage maize	*3%*	2%	**2%**	1%	1%
grain maize	*9%*	2%	1%	2%	1%
wheat	*4%*	**3%**	**3%**	1%	**3%**

Global trend to homogenize area % to the landscape waterlogged soil area (13%)

Concerning fallows:

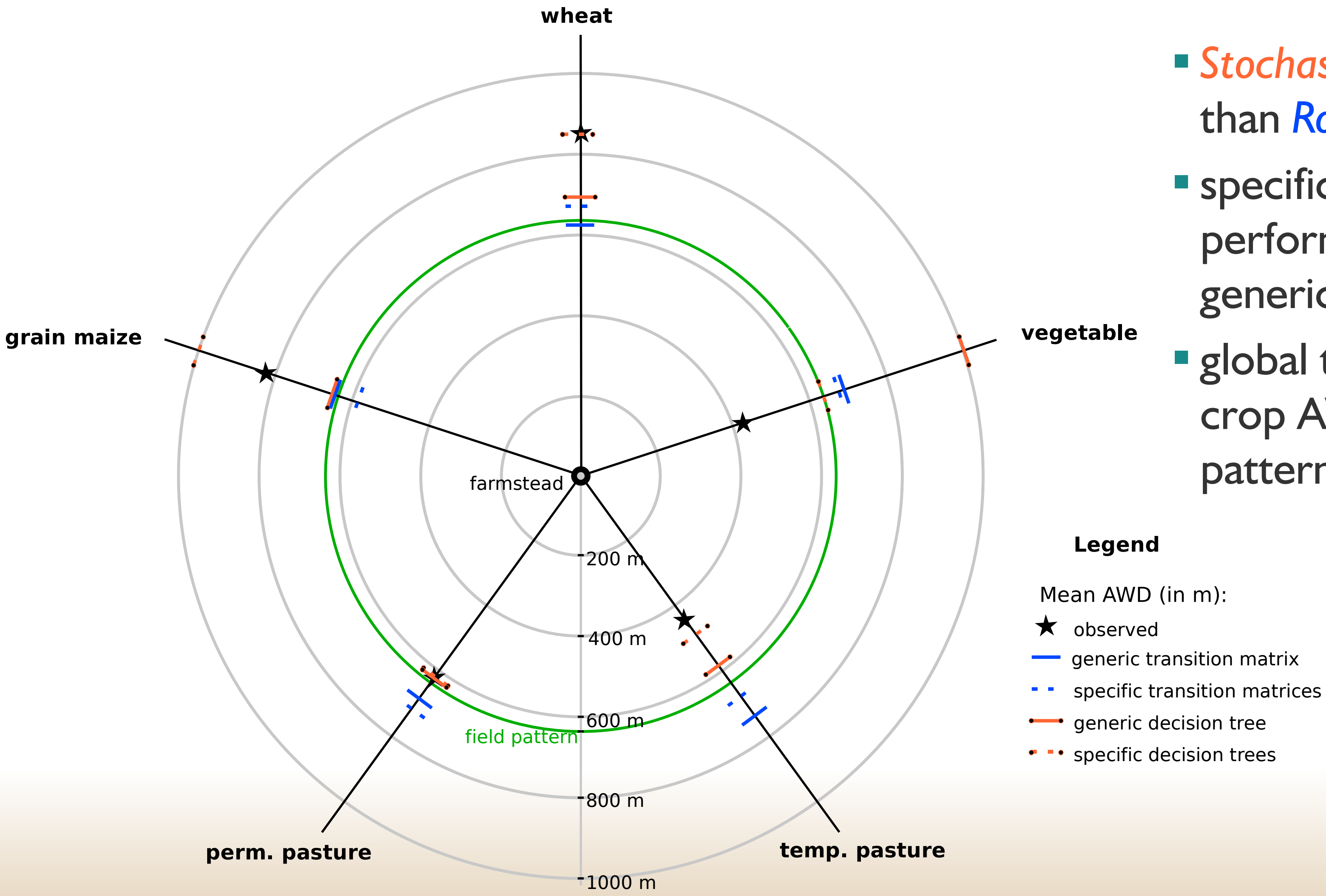
- *Stochastree* performed better than *Rotomatrix*
- specific trees/matrices performed better than generic ones

Other crops:

- similar ability of *Rotomatrix* and *Stochastree* to comply with crop water sensitivity

Long term simulations

Spatial distribution of crops around the dairy farmstead



- *Stochastree* performed better than *Rotomatrix*
- specific trees/matrices performed better than generic ones
- global trend to homogenize crop AWD to the field pattern AWD

Conclusion and perspectives

Results

- model construction based on datamining (not expertise)
- validation based confrontation with observed data
- structural analogy between the decision trees and the transition matrices
- transition matrices may implicitly integrate some soil waterlogging constraints
- decision trees maintain the spatial distribution of crops around the farmstead

Limits

- drivers stationarity is required during the construction of the trees/matrices, and in the comparison between simulated and observed data
- consequences of alternate trees are hard to forecast (conditional probabilities)

Perspectives

- model coupling: with distributed biophysical models (crop growth, water and nutrient fluxes), gene flow models, etc.
- use of decision trees to simulate an initial landcover state from a blank field pattern, farm-type, distance and waterlogging information