Kriging-based metamodeling with qualitative variables to design grassed buffer zones in small agricultural catchment

Claire Lauvernet, C. Helbert, C. Catalogne

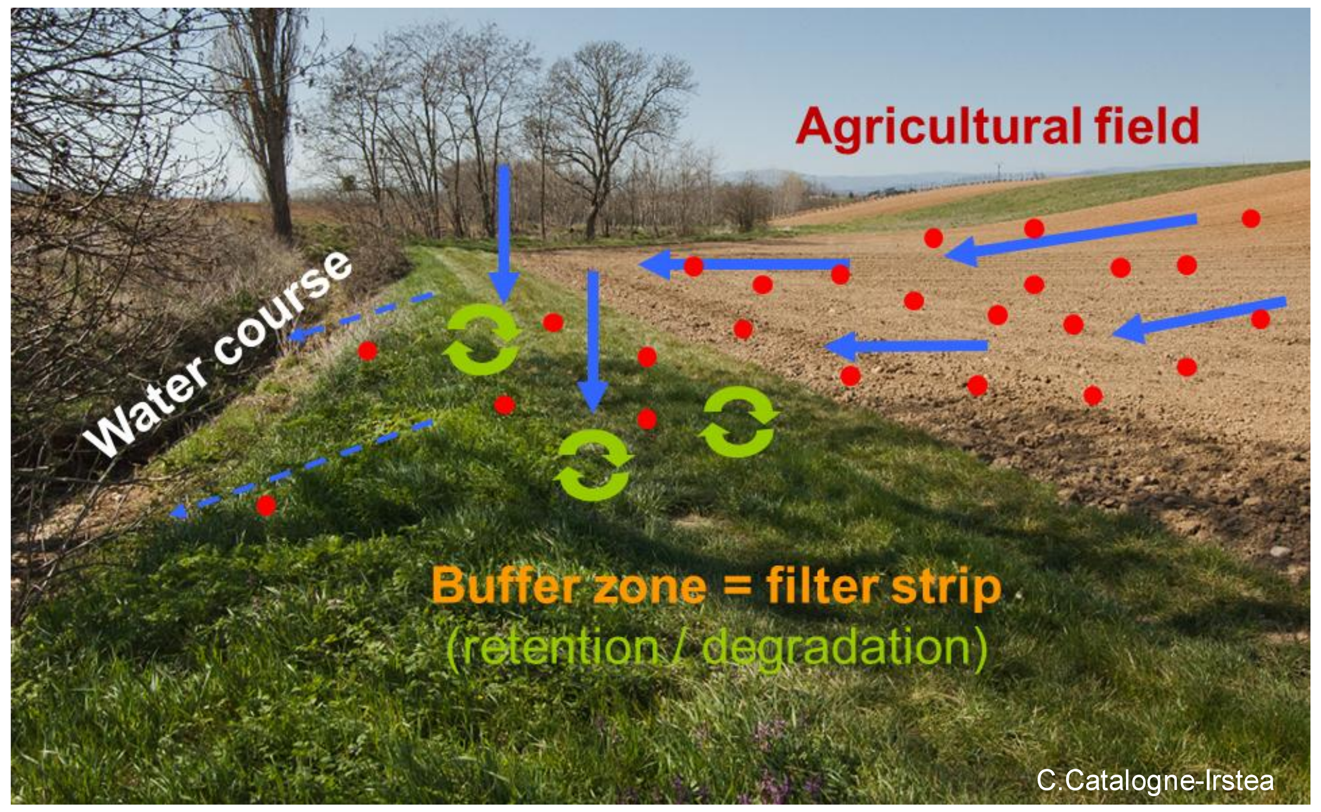
Kriging-based metamodeling with qualitative variables to design grassed buffer zones in small agricultural catchment

Lauvernet Claire[1], Helbert Céline[2], Catalogne Clotaire[1]

[1] RIVERLY, Irstea centre de Lyon-Villeurbanne, France.

[2] Univ Lyon, UMR CNRS 5208, École Centrale de Lyon, France.

Contact : claire.lauvernet@irstea.fr



Objectives

- Buffer strips are identified as the BMP of Choice for Runoff Mitigation to limit contamination of surface water by pesticides.
- Their efficiency strongly depends on soil, agronomic and climatic conditions and they need to be optimized by considering appropriate sizing.

➔ Irstea developed a complete/complex toolkit to design site-specific VFS by simulating their efficiency to limit runoff transfers : BUVARD (Carluer et al., 2017).

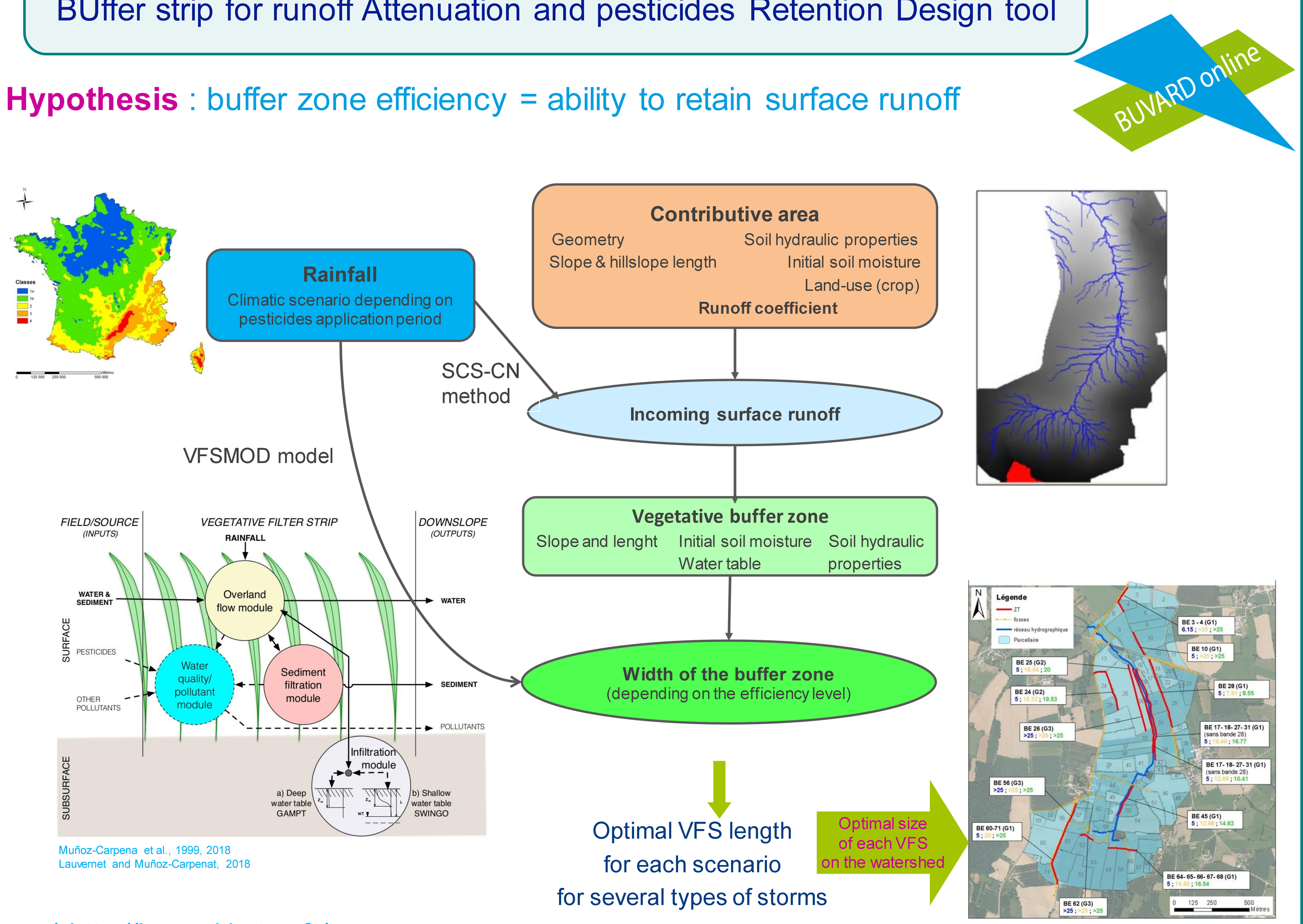
- This tool is based on quantitative *and* qualitative variables
- Need for a simpler and efficient tool for end-users

➔ A metamodel of BUVARD = methodological challenges for operational purposes

The toolkit BUVARD*

BUffer strip for runoff Attenuation and pesticides Retention Design tool

Hypothesis : buffer zone efficiency = ability to retain surface runoff

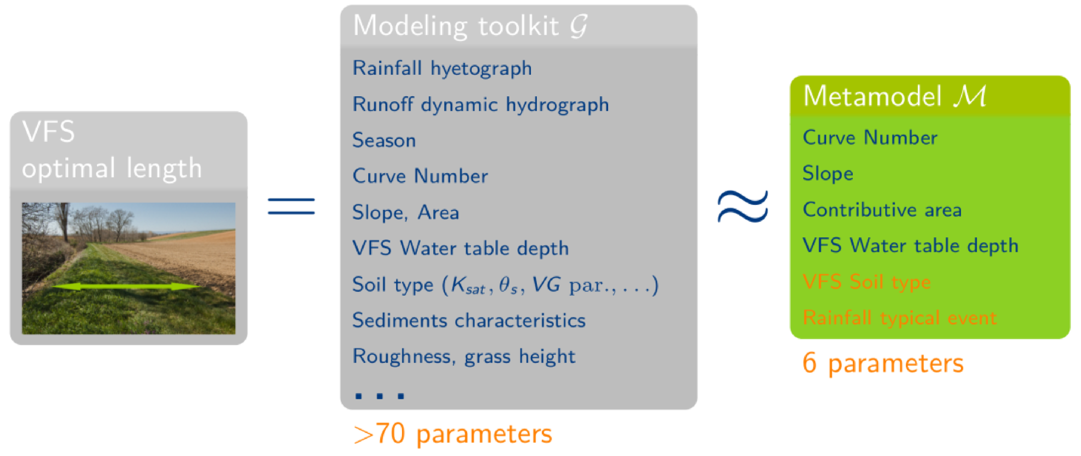


* http://buvard.irstea.fr/

Methods for building the metamodel

Why metamodeling BUVARD?

- Simple to use but still based on physics
- Able to evaluate an output of the toolchain at any point of the domain
- Allows evaluating sensitivity indices at smaller numerical cost
- Can be easily coupled/integrated in hydrological modeling frameworks



BUT: How to deal with qualitative/categorical variables (type of soil, type of rainfall event), that are typical in water quality modeling?

Stochastic method ?
Gaussian Processes / Kriging
The relation between points is expressed by a covariance structure between the obs.
➔ Need for an adapted cov. kernel

Deterministic method ?
GAM (Generalized Additive Model)
The relation between points is assumed to be deterministic and explicit
➔ Ok with qualit. var.

Gaussian Processes with mixed variables:

- Hyp. = the deterministic output of the model is the realization of a GP
- The GP Z is conditioned by points from the model simulations (still a GP)
- Several options to deal with categ. var.:
  - one GP per modalities couple
  - adapting the covariance kernel by progressive complexity

$$Z(x) = m(x) + W(x)$$
$$Cov(W(x), W(x')) = c(x, x') = \sigma^2 R(\theta, x, x')$$
$$R(\theta, x, x') = \prod R_k(\theta_k, x_k, x'_k)$$
$$R(\theta, x, x') = \prod_{j=1}^{J} \tau_j \prod_{k=1}^{K} R_k(\theta_k, x_k, x'_k)$$

corr. fun. of quali factor j: $\tau_{j,z_{ij},z_{hj}} = \rho_{ij}\rho_{hj}\mathbb{1}[z_{ij} \neq z_{hj}]$

corr. fun. of quanti. factor: $c(x - x') = \sigma^2 \exp - \left(\frac{(x-x')}{\theta}\right)$

Kernel	6 modalities for soil type:	4 modalities for rainfall:	Total
cov-quali-isotropic	$\begin{pmatrix} 1 & \rho & \rho & \rho \\ & 1 & \rho & \rho \\ & & 1 & \rho \\ & & & \rho \\ & & & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & \mu & \mu & \mu \\ & 1 & \mu & \mu \\ & & 1 & \mu \\ & & & 1 \end{pmatrix}$	2
cov-quali-multiplicative	$\begin{pmatrix} 1 & \rho_1\rho_2 & \rho_1\rho_3 & \rho_1\rho_4 \\ & 1 & \rho_2\rho_3 & \rho_2\rho_4 \\ & & 1 & \rho_3\rho_4 \\ & & & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & \mu_1\mu_2 & \mu_1\mu_3 & \mu_1\mu_4 \\ & 1 & \mu_2\mu_3 & \mu_2\mu_4 \\ & & 1 & \mu_3\mu_4 \\ & & & 1 \end{pmatrix}$	10
mixgp	$\begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ & 1 & \rho_{23} & \rho_{24} \\ & & 1 & \rho_{34} \\ & & & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & \mu_{12} & \mu_{13} & \mu_{14} \\ & 1 & \mu_{23} & \mu_{24} \\ & & 1 & \mu_{34} \\ & & & 1 \end{pmatrix}$	21

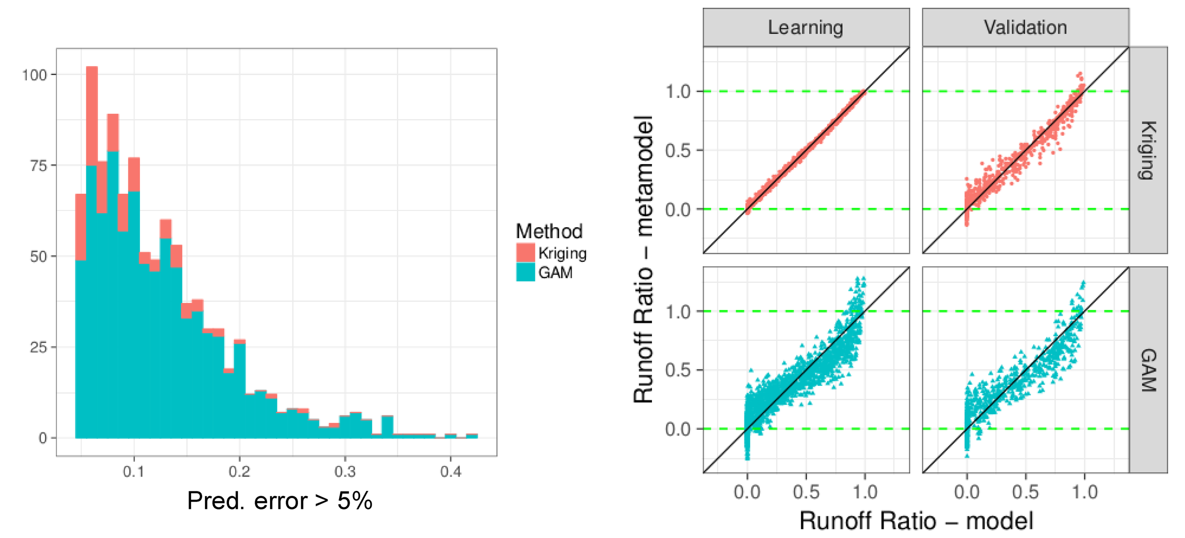
Sampling of the most influent and the most accessible input parameters

Key variable	Sampling range
Curve Number	[63 , 99]
Slope	[2% , 20%]
Contributive area length(m	[25 , 300]
VFS Water table depth (m)	[0.5 , 4]
Rainfall typical event	Summer/Winter, short/long
VFS Soil type	clo, scl, SIL, CLO , SCL, SAL

Results

- Comparison of two methods : Kriging & Generalized Additive Model
- Kriging was adapted to qualitative variables with adapted covariance model

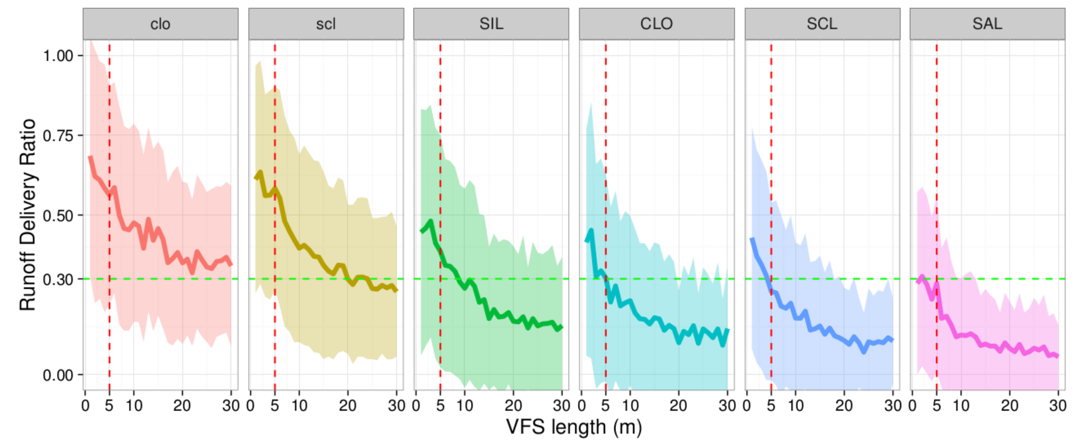
Output variable is the Runoff Delivery Ratio = Runout/Runin :
RDR = 0 ⇔ efficiency = 100%; RDR = 1 ⇔ efficiency = 0%



GAM R^2 = 0.84
Kriging R^2 = 0.96

- ➢ Kriging is very efficient with qualitative variables
- ➢ cov_quali_multiplicative and mixgp give the best prediction and are similar
- ➢ Still some prediction error due to the plateau
- ➢ High sensitivity of RDR to the Curve Number

- Uncertainty analysis : Average efficiency and uncertainty on large sample simulations (24 000)



- ➢ a very large variability of optimal length / type of soil
- ➢ possibility to quantify uncertainty of the results

Conclusion

- Qualitative variables were properly taken into account by the GP adaptation
- Good quality of prediction (94 % of variance) but some bad predictions
- The optimal VFS length is very sensitive to the Curve Number
- MM is a promising tool to test the toolkit to perform UA and GSA at low cost

Limitations and perspectives:

- On larger ranges (Curve Number in p.), the current MM is not satisfying (due to a large plateau of null values of the output variable)
- Continue testing and compare methods on :
  - prediction uncertainty
  - sensitivity of prediction quality to the sampling size
  - global sensitivity analysis (Sobol)

Ref. Carluer, N et al. 2017. *Defining context-specific scenarios to design vegetated buffer zones that limit pesticide transfer via surface runoff.* Sci. Of Total Env. , 575. // Munoz-Carpena, R.; Lauvernet, C. & Carluer, N., 2018. *Shallow water table effects on water, sediment, and pesticide transport in vegetative filter strips -- Part 1: nonuniform infiltration and soil water redistribution.* Hydrol. Earth System Sci., 22. // Lauvernet, C. & Munoz-Carpena, R. 2018. *Part 2: model coupling, application, factor importance, and uncertainty.* Hydrol. Earth System Sci., 22. // Chen, Wang & Yang, 2013, *Stochastic Kriging with Qualitative factors.* WSC '13 Proceedings. // Panodou & Roustant, 2017 *mixgp R package: Kriging models for mixed data.*

Acknowledgments: Many thanks to Robert Faivre, Nicolas Forquet, Esperan Padonou, Oivier Roustant for our fruitful discussions on metamodeling, and to Nadia Carluer, Guy Le Hénaff, Etienne Leblois, Rafael Munoz-Carpena, for BUVARD development. Thanks to LEFE-MANU who supported this study.