POLITECNICO DI TORINO | Compagnia di San Paolo | Georgia Tech

Program for Internalization of Research 2018
Joint Project "Machine Learning to Improve the Reliability of Complex Systems"

Statistical Analysis of the Efficiency of an Integrated Voltage Regulator by means of a Machine Learning Model Coupled with Kriging Regression

R. Trinchero[1], M. Larbi[2], M. Swaminathan[2] and **F.G. Canavero**[1]

[1]*Department of Electronics and Telecommunication, Politecnico di Torino, Torino, Italy*
Email: {riccardo.trinchero,flavio.canavero}@polito.it

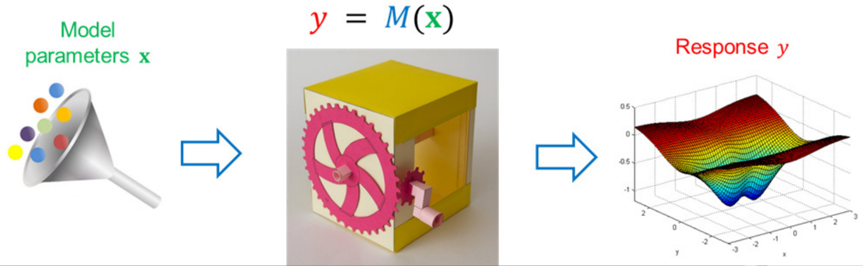
[2]*School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, USA*
Email: mourad.larbi@ece.gatech.com, madhavan.swaminathan@ece.gatech.edu

Statistical Analysis of the Efficiency of an Integrated Voltage Regulator by means of a Machine Learning Model Coupled with Kriging Regression – SPI 2019, 18-21 June, 2019

Intro: design process and needs

- ❑ Robust designs require assessment against uncertainty
- ❑ Optimize design to achieve better product performance

➢ Need for a **computational model**, *i.e. a procedure (e.g., analytical formula, algorithm, …) computing quantities of interest from input parameters*

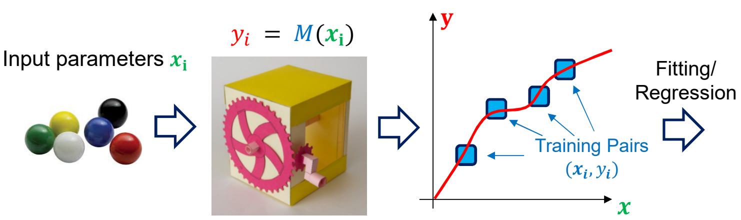


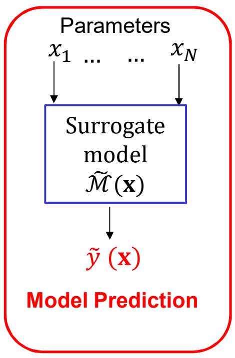
The crux of the matter

- **Computational models** are demanding in terms of memory and CPU time
- **Statistical simulations** are necessary for both
  - the assessment of the design robustness w.r.t. uncertainty and variability
  - the performance optimization

Surrogate Modeling

- A **surrogate model** $\tilde{\mathcal{M}}$ is an approximation of the **full-computational model** $\mathcal{M}$





- Determined from a <u>limited number of runs</u> of the full model
- Simpler **closed-form** relationship
- **Faster** then the full-model $\mathcal{M}$

Several **fitting techniques** are available

Least Squares Support Vector Machines (LS-SVM)

- **Machine learning** technique for both **classification** and **regression**
- **Kernel method** mapping data from the vector space to the feature space
- The # of unknowns is **independent** from **problem dimensionality (non-parametric regression)**

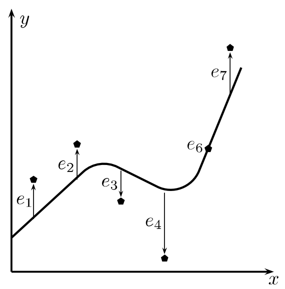
Generic formulation

$$\mathcal{M}(\mathbf{x}) = \sum_{i=1}^{L} \alpha_i K(\mathbf{x}_i, \mathbf{x}) + b$$

The **response** $y = \mathcal{M}_{LS-SVM}(x)$ is given as the **linear combination** of the kernel evaluated at the training samples x_i.

- Linear kernel : $K(\boldsymbol{x}, \boldsymbol{x}_i) = \boldsymbol{x}_i^T \boldsymbol{x}$,
- Polynomial kernel of degree d: $K(\boldsymbol{x}, \boldsymbol{x}_i) = \left(1 + \boldsymbol{x}_i^T \boldsymbol{x}/c\right)^d$,
- Radial basis function RBF kernel : $K(\boldsymbol{x}, \boldsymbol{x}_i) = \exp\left(-\|\boldsymbol{x} - \boldsymbol{x}_i\|^2/\sigma^2\right)$,
- MLP kernel : $K(\boldsymbol{x}, \boldsymbol{x}_i) = \tanh\left(k\, \boldsymbol{x}_i^T \boldsymbol{x} + \theta\right)$,

LS-SVM (cont'd)



$$\mathcal{M}_{LS-SVM}(\mathbf{x}) = \sum_{i=1}^{L} \alpha_i K(\mathbf{x}_i, \mathbf{x}) + b$$

Least squares solution

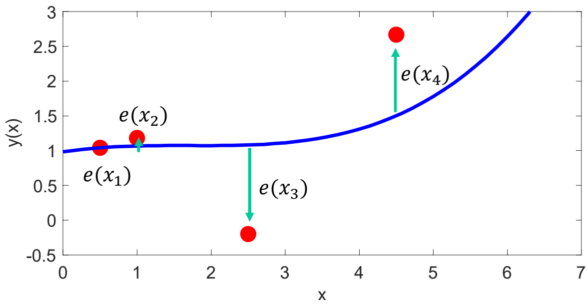
$$\begin{bmatrix} 0 & 1 & \dots & 1 \\ 1 & K(x_1, x_1) + \frac{1}{\gamma} & & K(x_1, x_L) \\ \vdots & & \ddots & \\ 1 & K(x_L, x_1) & & K(x_L, x_L) + \frac{1}{\gamma} \end{bmatrix} \begin{bmatrix} b \\ \alpha_1 \\ \vdots \\ \alpha_L \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ \vdots \\ y_L \end{bmatrix}$$

The coefficients α and b are estimated by solving a **simple linear system**

How accurate is the model prediction ???

Deterministic Regression & Error Function

- Training samples $\{(x_1, y_1), \ldots, (x_4, y_4)\}$



- Build a **surrogate mode** via a **deterministic regression**

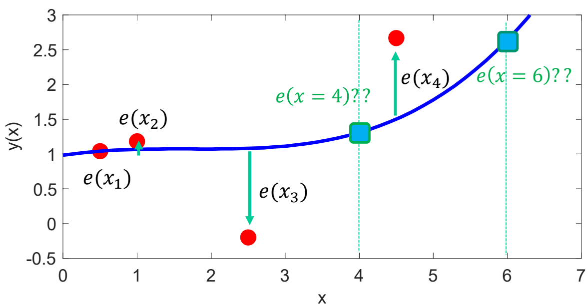
$$y_i = M_{LS-SVM}(x_i) + e(x_i)$$

Surrogate Model — Model Error

Statistical Analysis of the Efficiency of an Integrated Voltage Regulator by means of a Machine Learning Model Coupled with Kriging Regression – SPI 2019, 18-21 June, 2019

Deterministic Regression & Error Function

- Training samples $\{(x_1, y_1), \ldots, (x_4, y_4)\}$



- Build a **surrogate mode** via a **deterministic regression**

$$y_i = M_{LS-SVM}(x_i) + e(x_i)$$

Surrogate Model — Model Error

- Model Prediction

$$y(4) \approx M_{LS-SVM}(4) + ??$$
$$y(6) \approx M_{LS-SVM}(6) + ??$$

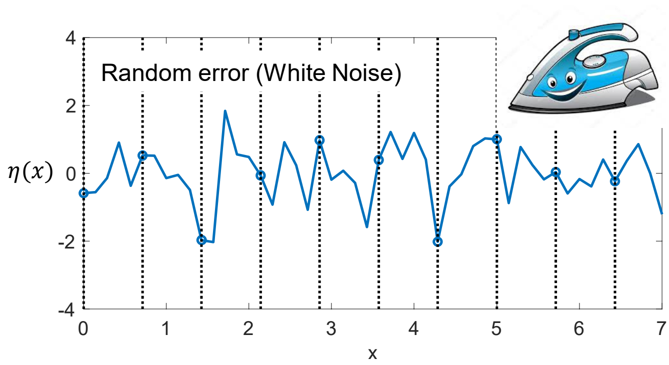
the regression error $e(x)$ is **only known** on the **training samples** x_i

How accurate is the prediction?
Confidence bounds are needed!!!

Probabilistic Interpretation

We can think of the error function $e(x)$ **in a probabilistic sense**

In practice, error is a smooth function

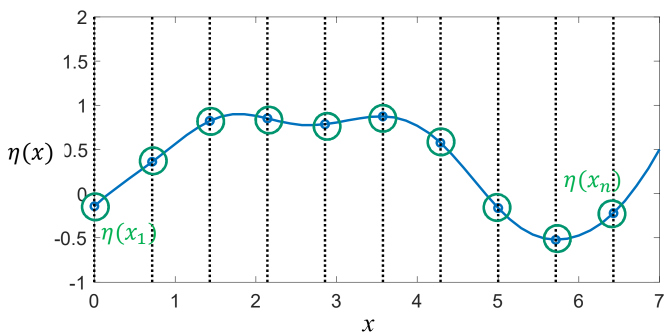


Probabilistic Interpretation

We can think of the error function $e(x)$ **in a probabilistic sense**

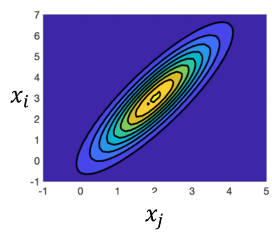
Assumption:

$e(x)$ is **smooth** → The values $\eta(x_1), .., \eta(x_n)$ **are correlated**



Correlation via Covariance function, e.g.,

$$k(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{(x_i - x_j)^2}{2\sigma_l^2}\right)$$

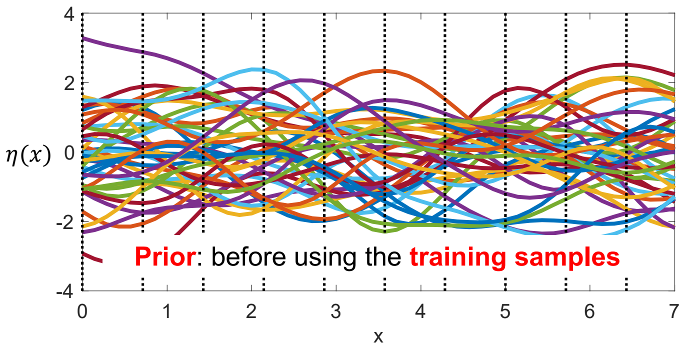


Probabilistic Interpretation

We can think of the error function $e(x)$ **in a probabilistic sense**

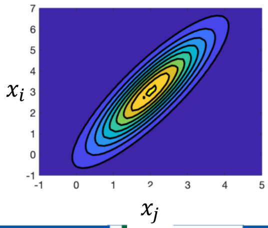
Assumption:

$e(x)$ is **smooth** → The values $\eta(x_1),..,\eta(x_n)$ **are correlated**



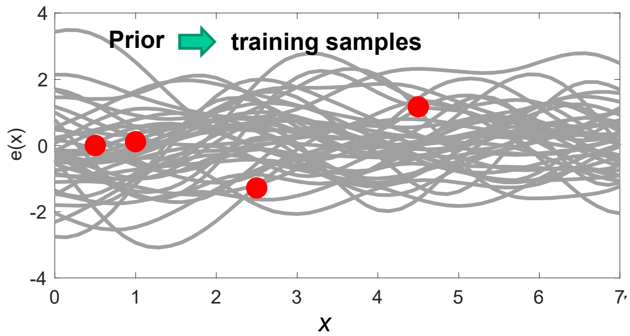
Correlation via Covariance function, e.g.,

$$k(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{(x_i - x_j)^2}{2\sigma_l^2}\right)$$



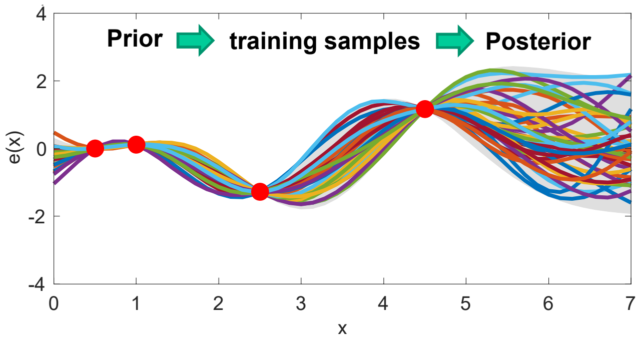
Gaussian Process (GP) Regression: from Prior to Posterior

- The value of the **error** $e(x_i)$ is known on the **training samples** $e(x_i)$



GP Regression: from Prior to Posterior

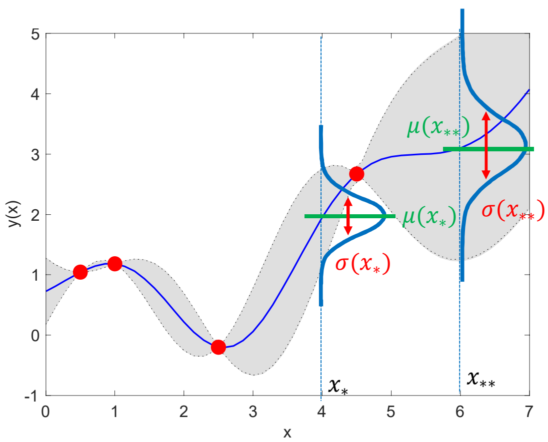
- The value of the **error** $e(x_i)$ is known on the **training samples** $e(x_i)$
- Discarding all the functions not compatible with the training samples



We consider **only all functions fitting our data**
Such functions can be described in terms of a **probability law** →
probabilistic model

GP Regression & Probabilistic Model

- The **resulting model** combines the **deterministic regression** M_{LS-SVM} with a **probabilistic model** of the error function $e(x)$



The model provides for any value x_* a **distribution** [3]

$$M(x_*) \sim N(\mu(x_*), \sigma^2(x_*))$$

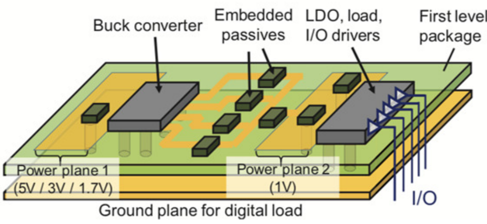
most probable value of the error | **confidence bounds**

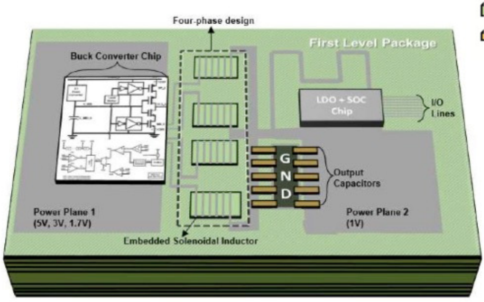
GP allows converting any **regression-based deterministic** model into a **probabilistic** one

[3] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning, MIT Press. Cambridge, Massachusetts, 2006.

Integrated Voltage Regulator (IVR)

- **System-in-package** (SiP) solution with buck converter, low-dropout/load and an **integrated inductor** on an organic package [4]



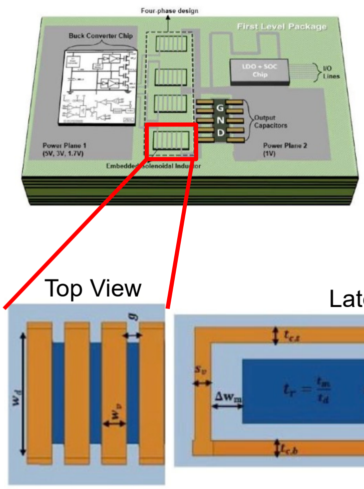


- The **buck converter** is used to drop the voltages of the power plane 1 to the level required by the power plane 2

The converter **efficiency** depends on the **integrated inductor**

[4] H. M. Torun, et al., "A Global Bayesian Optimization Algorithm and Its Application to Integrated System Design," in *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 4, pp. 792–802, April 2018.

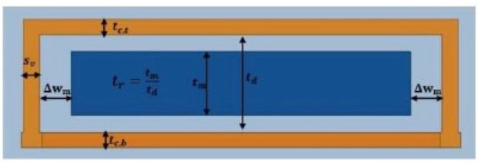
Embedded Solenoidal Inductor [5]



CONTROL PARAMETERS OF SOLENOIDAL INDUCTOR

Uniform random variables		Unit	$\mathcal{U}[Min; Max]$
Gap between windings	g	mil	$\mathcal{U}[4;6]$
Size of via	s_v	μm	$\mathcal{U}[80;120]$
Copper Trace Width	w_v	mil	$\mathcal{U}[9;11]$
Copper Thickness Bottom	$t_{c,b}$	μm	$\mathcal{U}[64;96]$
Copper Thickness Top	$t_{c,t}$	μm	$\mathcal{U}[64;96]$
Dielectric Thickness	t_d	μm	$\mathcal{U}[180;220]$
Dielectric Width	w_d	mil	$\mathcal{U}[59.4;60.6]$
Magnetic Core Width offset	Δw_m	mil	$\mathcal{U}[9;11]$

Lateral View



The IVR efficiency has been investigated by **considering 8 uniformly distributed parameters**

[5] R. Trinchero, at al., "Machine Learning and Uncertainty Quantification for Surrogate Models of Integrated Devices With a Large Number of Parameters," in IEEE Access, vol. 7, pp. 4056-4066, 2019.

IVR Results

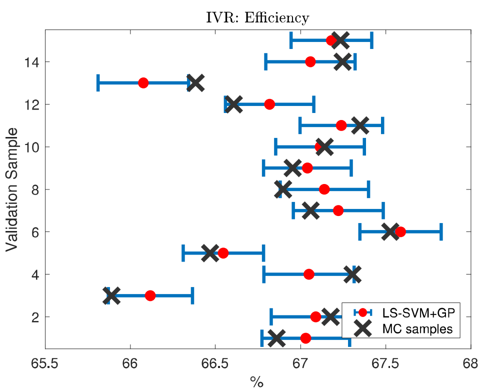
- **A subset of $L = 200$ training samples** is selected via Latin Hypercube Sampling
- The surrogate model predictions are **compared** with the results of a **MC simulation** with 10000 samples

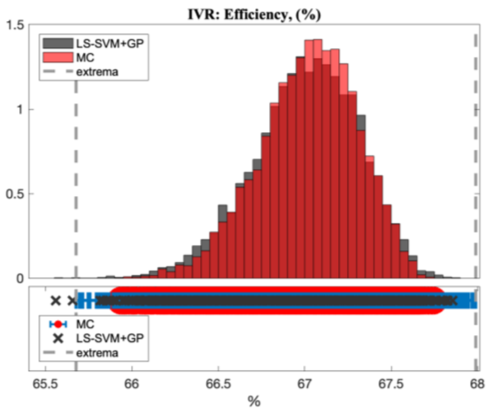
Method	Kernel Regression	RMSE	$\hat{\mu}$	$\hat{\sigma}$	t_{model}	t_{cost}
MC	–	–	67.01	0.31	–	7 days
LS-SVM	Linear	0.158	67.03	0.28	<1s	<1s
	Poly Order 2	0.166	67.04	0.29	<1s	<1s
	Poly Order 3	0.443	67.00	0.54	<1s	<1s
	RBF	0.152	67.02	0.28	1.4s	<1s
GP	–	0.162	67.02	0.28	<1s	2.4s
LS-SVM+GP	Linear	0.162	67.02	0.28	<1s	<1s
	RBF	0.156	67.02	0.28	1.5s	<1s
Sparse PC	Poly Order 9	0.170	67.02	0.28	<1s	<1s

LS-SVM and LS-SVM+GP regression with RBF kernel provides the **most accurate** metamodel

IVR Results (cont'd)

- The prediction and 95% **confidence intervals** estimated by the GP+LS-SVM (RBF) are compared with the results of a **MC simulation** with 10000 samples.





Excellent agreement for both the model prediction and Confidence Intervals

Conclusions

- **Surrogates** based **on Machine Learning regressions** represent effective solution for the UQ in nonlinear problems

- **Surrogates** are built from a **limited set of training samples** provided by the full-model

- **Gaussian Process regression (aka Kriging)** allows building accurate probabilistic models providing an estimation of the **model output** + **confidence bounds**

- FUTURE WORK**:** What is the limit in term of number of parameters?