The Machine Learning Approach to Dynamic Security Assessment

Simon Tindemans, TU Delft
Jochen Cremer, Imperial College London

TU Delft
Imperial College London

IEEE BDA Tutorial Series: Big Data & Analytics for Power Systems
10:00 am-12:00 pm, PST, Monday, Dec. 2, 2019

Presenters

Simon Tindemans

- Assistant professor at Delft University of Technology (NL)
- Visiting researcher at the Alan Turing Institute (UK)
- **Research interest:** machine learning for risk assessment, efficient computation and apportioning of risks

Jochen Cremer

- PhD student (final year) at Imperial College London (UK)
- **Research interest**: intersection of machine learning and mathematical optimization applied to the operation of the power system

TUDelft

Imperial College London

Disclaimer: the materials presented in this tutorial will be somewhat biased towards our own research.

Opening credits

Collaborators and funders of our work in this area

- **Imperial College London**: Ioannis Konstantelos, Mingyang Sun, Federica Bellizio, Goran Strbac
- **RTE**: Jean Maeght, Nicolas Omont, Samir Issad, Patrick Panciatici, Antoine Marot, Benjamin Donnot

iTesla
Innovative Tools for Electrical System Security within Large Areas

iPST project

Rte
Réseau de transport d'électricité

TU Delft
Imperial College London

Disclaimer: the materials presented in this tutorial will be somewhat biased towards our own research.

Outline

~ 20 mins
- Dynamic Security Assessment (DSA): what and why?
- The Machine Learning approach to DSA

~ 30 mins
- The 'offline' process
  - Classifier training
  - Data generation

~ 40 mins
- The 'online' process
  - Targeting simulations
  - DSA for online control
- Questions

TUDelft
Imperial College London

TU Delft
Imperial College London

Introduction

Problem statement

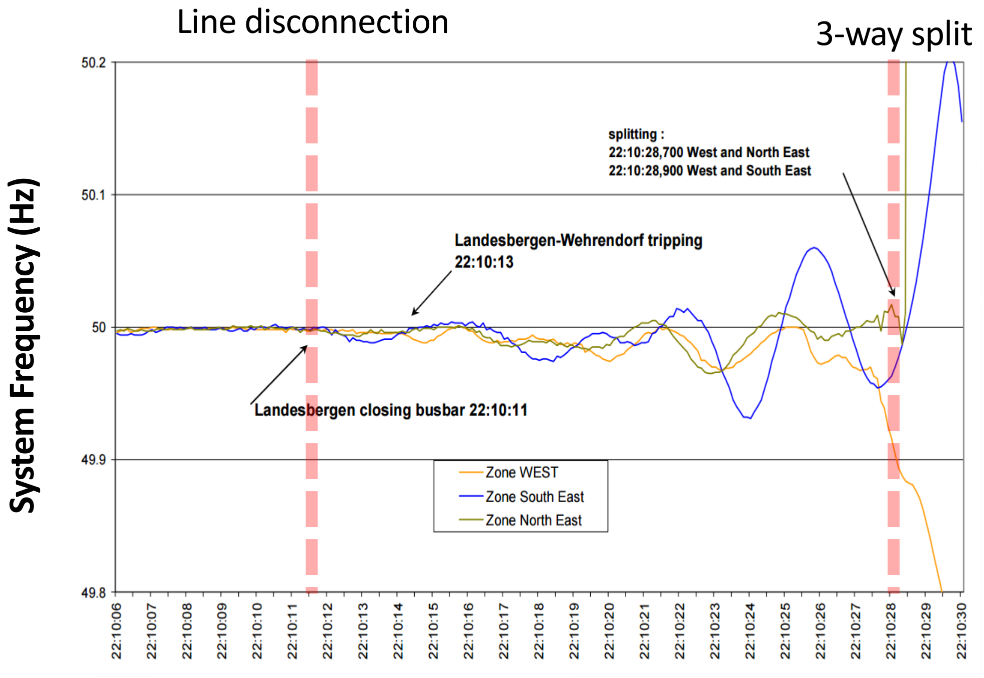
Powerline crossing at Ems river

double-circuit 380kV

TU Delft

Imperial College London

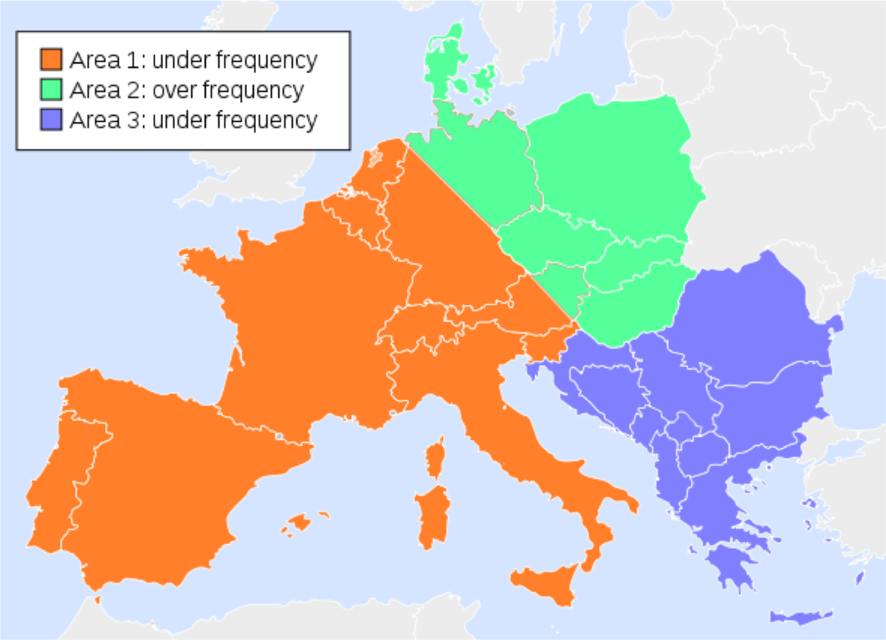
4 November 2006 – cascading faults



Source: UCTE Final Report – System Disturbance on 4 November 2006

TU Delft
Imperial College London

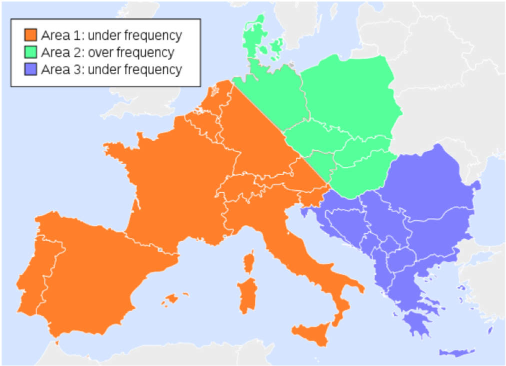
Pan-European disturbance

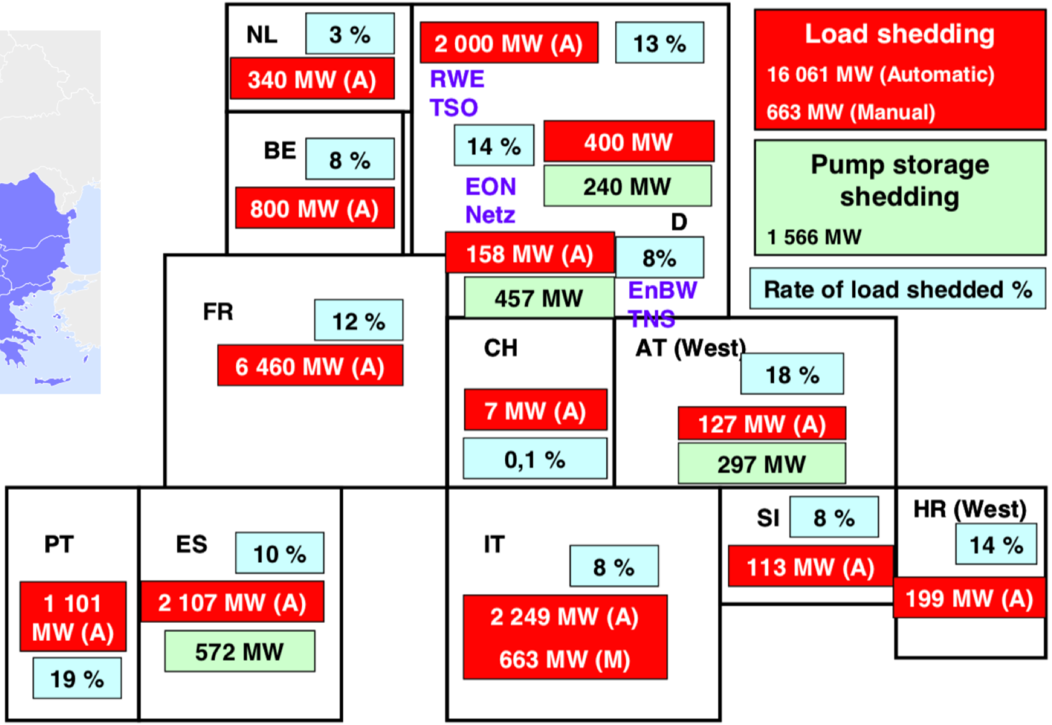


Source: UCTE Final Report – System Disturbance on 4 November 2006

TU Delft
Imperial College London

Load shedding

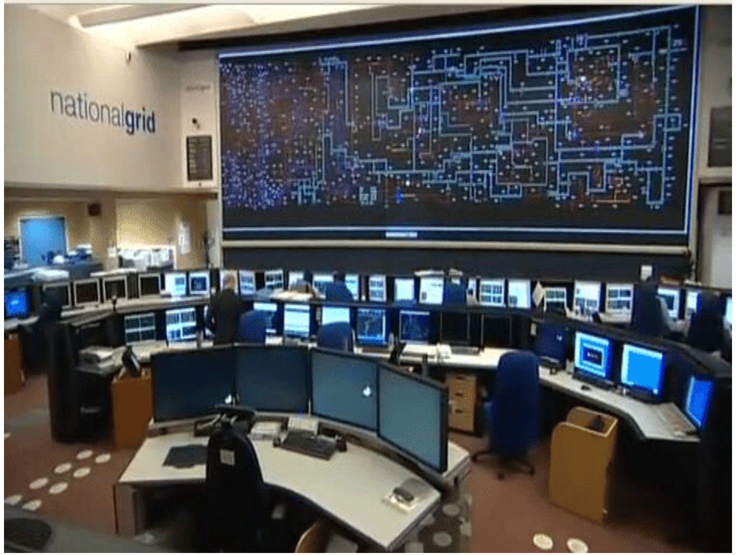




TU Delft
Imperial College London

Source: UCTE Final Report – System Disturbance on 4 November 2006

The system operator's challenge



System operators are responsible for the reliable supply of power to end users

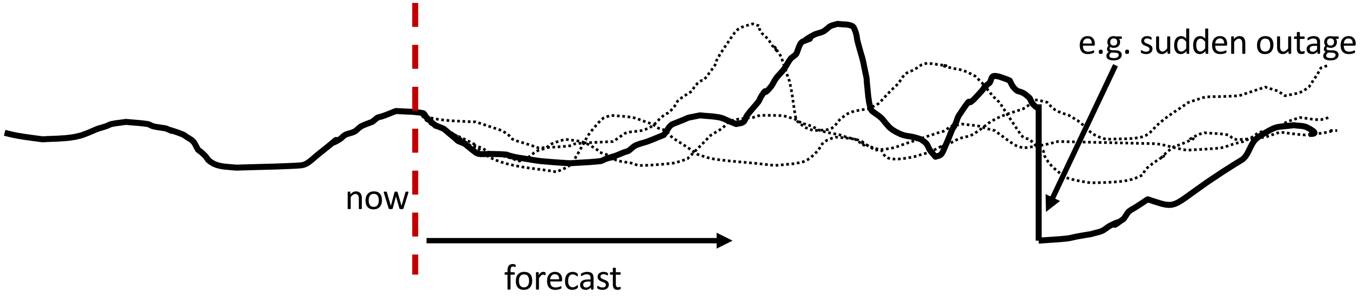
1. Electricity markets 'propose' a solution
2. System operators check this solution and prepare for uncertainties
3. Operators can override market outcomes, but this is **expensive and/or carbon-intensive.**
4. When multiple TSOs are involved, things get harder.

TUDelft
Imperial College London

Uncertainties and decisions

Uncertainties

- continuous (e.g. wind forecasts)
- discrete (sudden failures, aka **contingencies**)



When to act?

day ahead | hour ahead | 'corrective' actions

- most uncertainty
- most options

- limited uncertainty
- restricted options

- 'no' uncertainty
- very few actions and limited time

TU Delft
Imperial College London

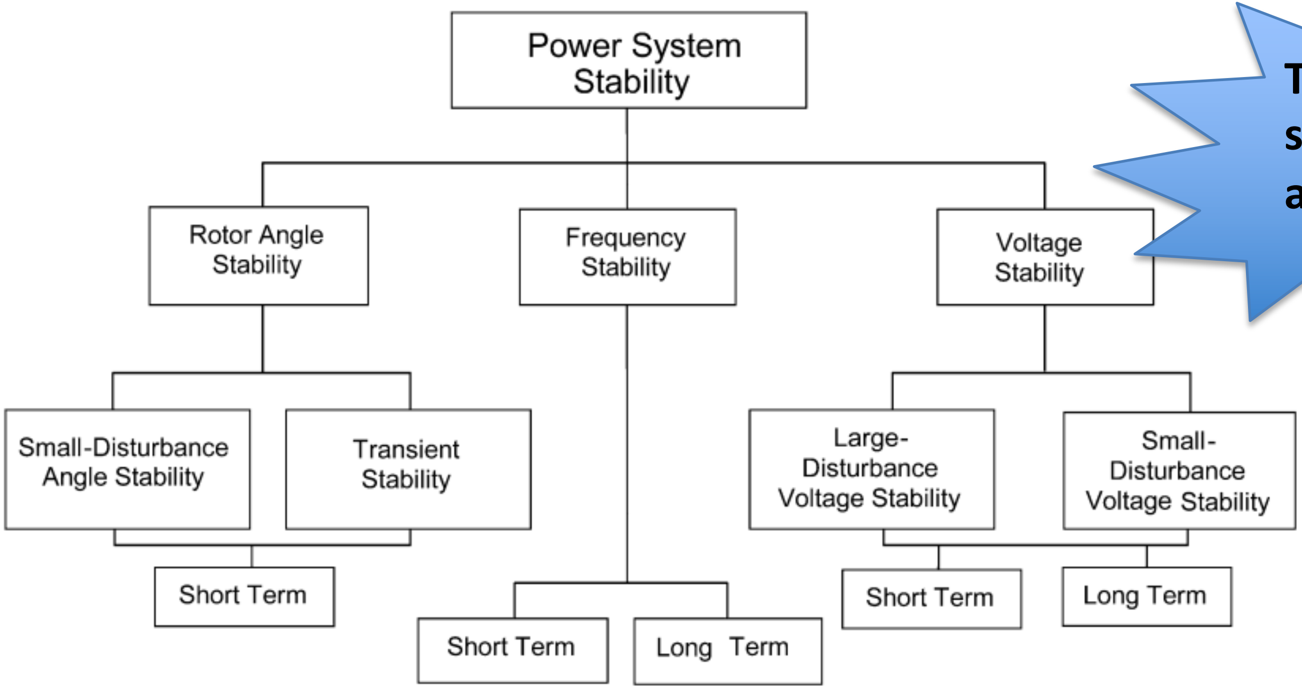
Reliability, Security and Stability

- "**Reliability** of a power system refers to the probability of its satisfactory operation over the long run."
- "**Security** of a power system refers to the degree of risk in its ability to survive imminent disturbances (contingencies) without interruption of customer service."
- "(..) **Stability** is the ability of an electric power system, for a given initial operating condition, to regain a state of operating equilibrium after being subjected to a physical disturbance, with most system variables bounded so that practically the entire system remains intact."

implies

P. Kundur *et al.*, "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," in *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387-1401, Aug. 2004.

Classification of stability



Time-domain simulations are required

P. Kundur *et al.*, "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," in *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387-1401, Aug. 2004.

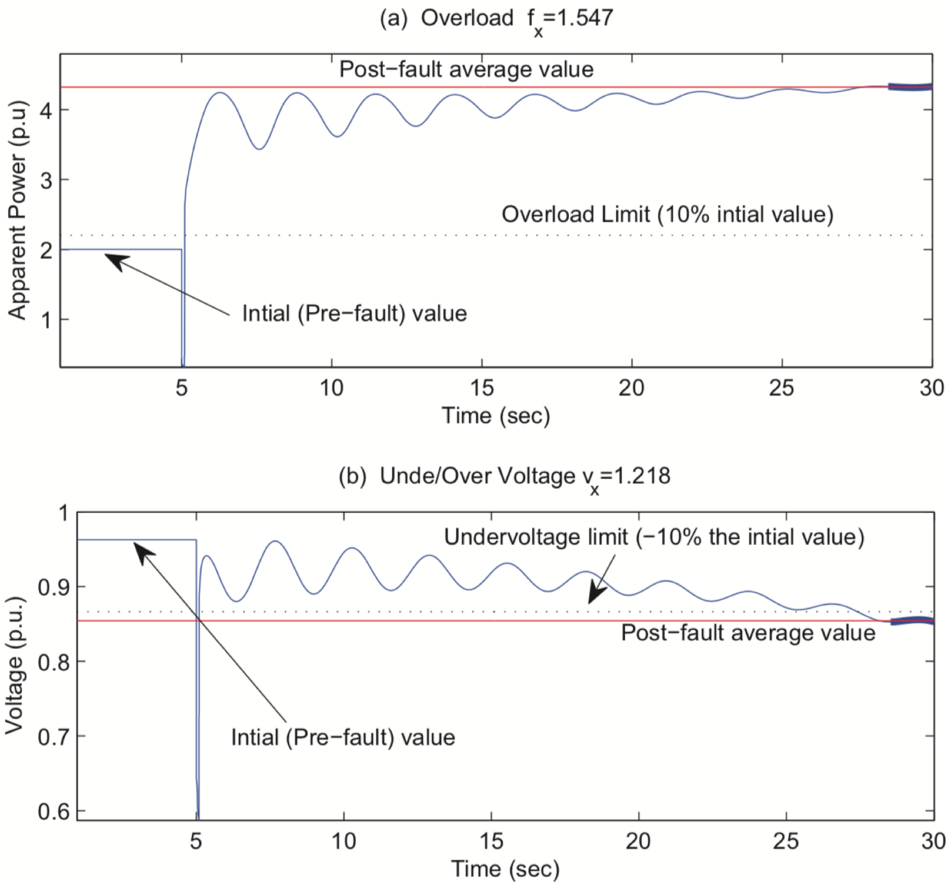
Foresight through simulations

High-fidelity simulation

- Time-domain
- All components

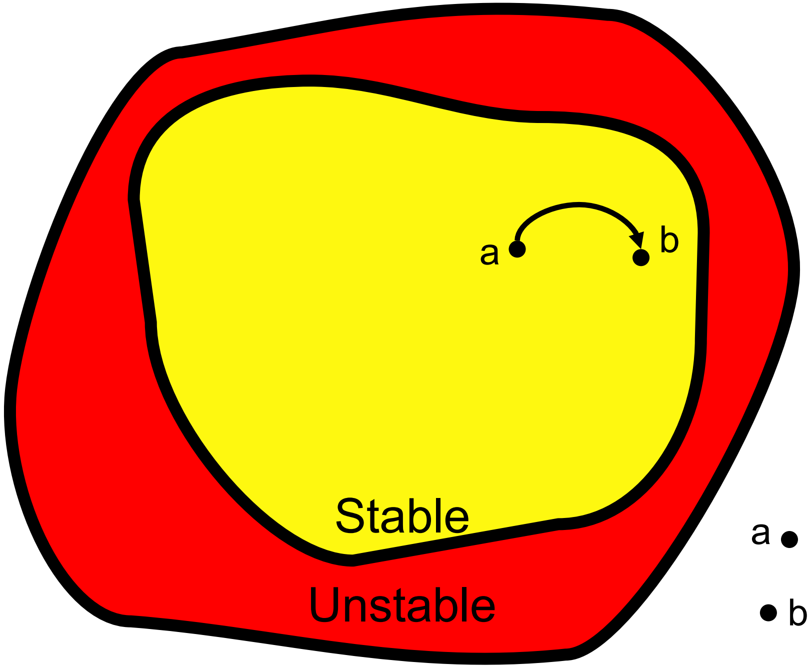
Case study

- Three-phase fault
- Initiates line trip



F. R. S. Sevilla and L. Vanfretti, "Static stability indexes for classification of power system time-domain simulations," *2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Washington, DC, 2015.

System response to a disturbance

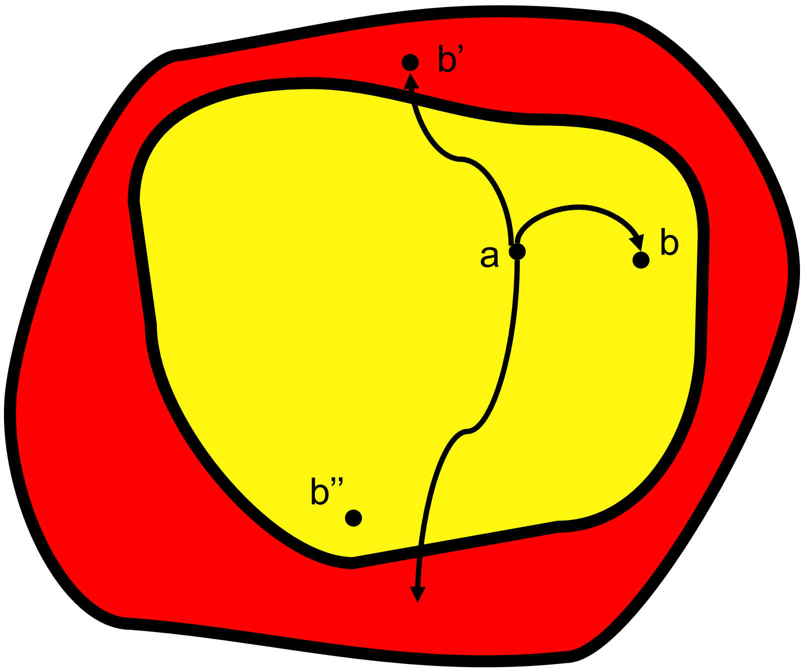


a ● pre-fault operation condition

● b post fault steady-state operation

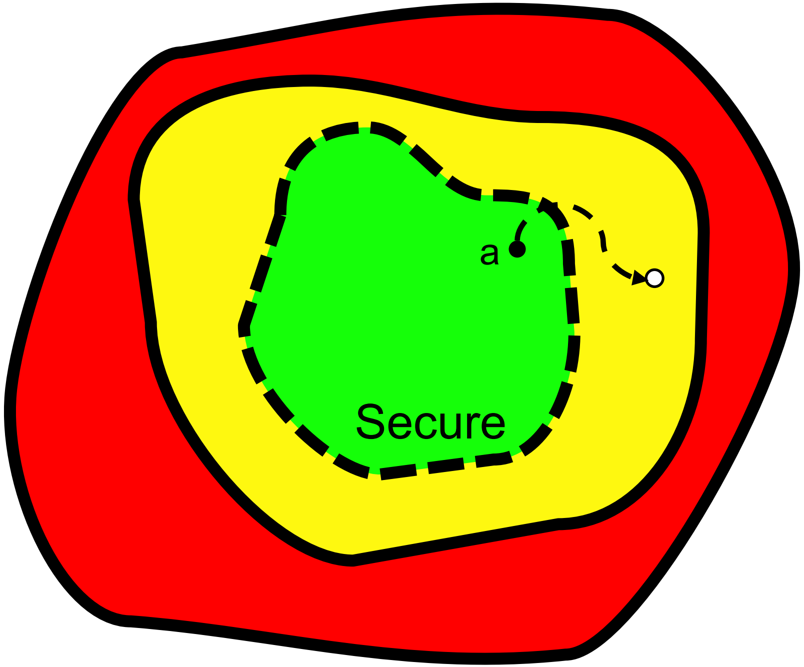
time-domain trajectory

Transient stability



1. steady-state stable and trajectory stable in b
2. steady-state unstable in b’
3. steady-state stable in b’’, however the trajectory is unstable

Security



Security is operator-defined:

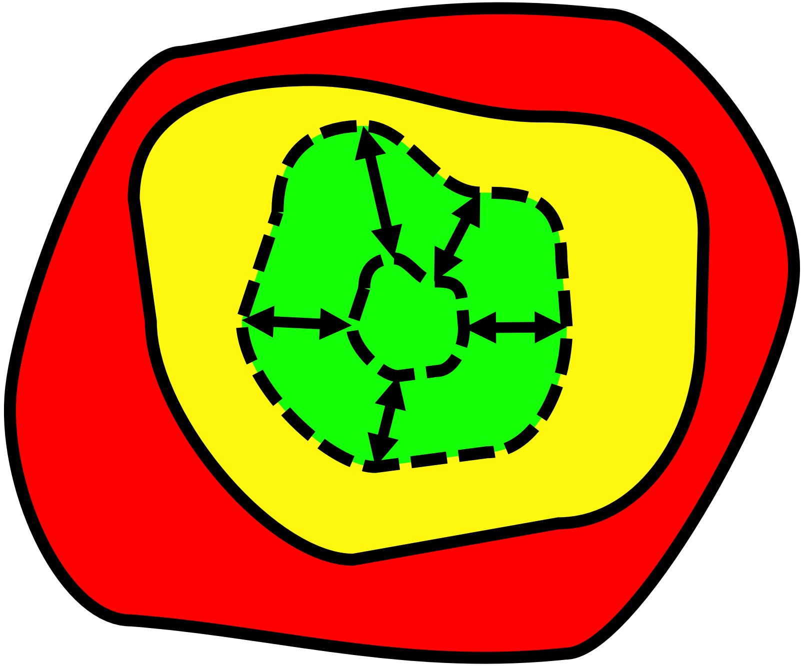
Example: N-1 security

- Post-fault stability for any *single* equipment fault.
- All operational constraints are satisfied at all times

In the remainder, we will assume a **secure contingency list.**

TU Delft
Imperial College London

Security vs cost



"survive imminent disturbances (contingencies) "

There is a trade-off:

- 'the smaller' the region the more secure
- 'the larger' the region the cheaper to operate

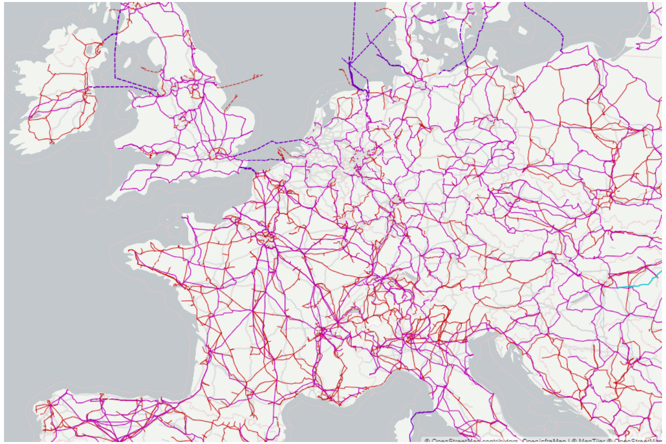
TUDelft
Imperial College London

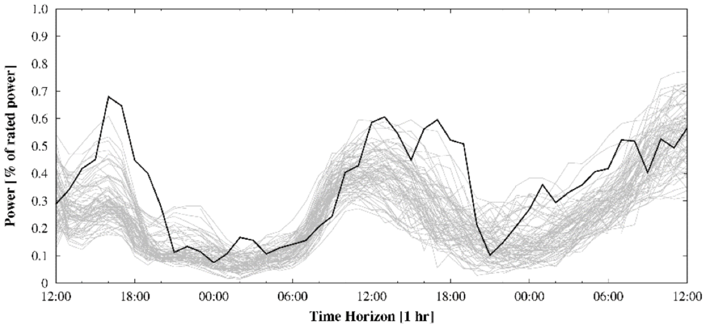
Introduction

Machine learning for dynamic security assessment

TU Delft
Imperial College London

Computational burden





- Dynamic simulations are carried out with a detailed model that accounts for each asset.
- These simulations must be considered for each outage and each operation condition
- **Result:** Too many cases to be simulated in real time

Security assessment: analysis tools

1. Actual experiments	✖	
2. Time-domain simulations	10k+ coupled ODEs ~minutes	**tool of choice** **... but slow**
3. Quasi-steady state simulations with AC, linearised AC or DC power flows	10k+ nonlinear equations ~seconds	**quick ...** **but biased**
4. Proxy (aka emulator) for time-domain simulator	e.g. decision tree or NN ~ms	based on data + mathematics, not physics (watch this space)

TU Delft
Imperial College London

Not a new idea. See e.g.:
Wehenkel, L., (1998). *Automatic Learning Decision trees and Techniques in Power Systems*, ISBN 978-0-7923-8068-9.

The general idea

Offline analysis

Security Rules

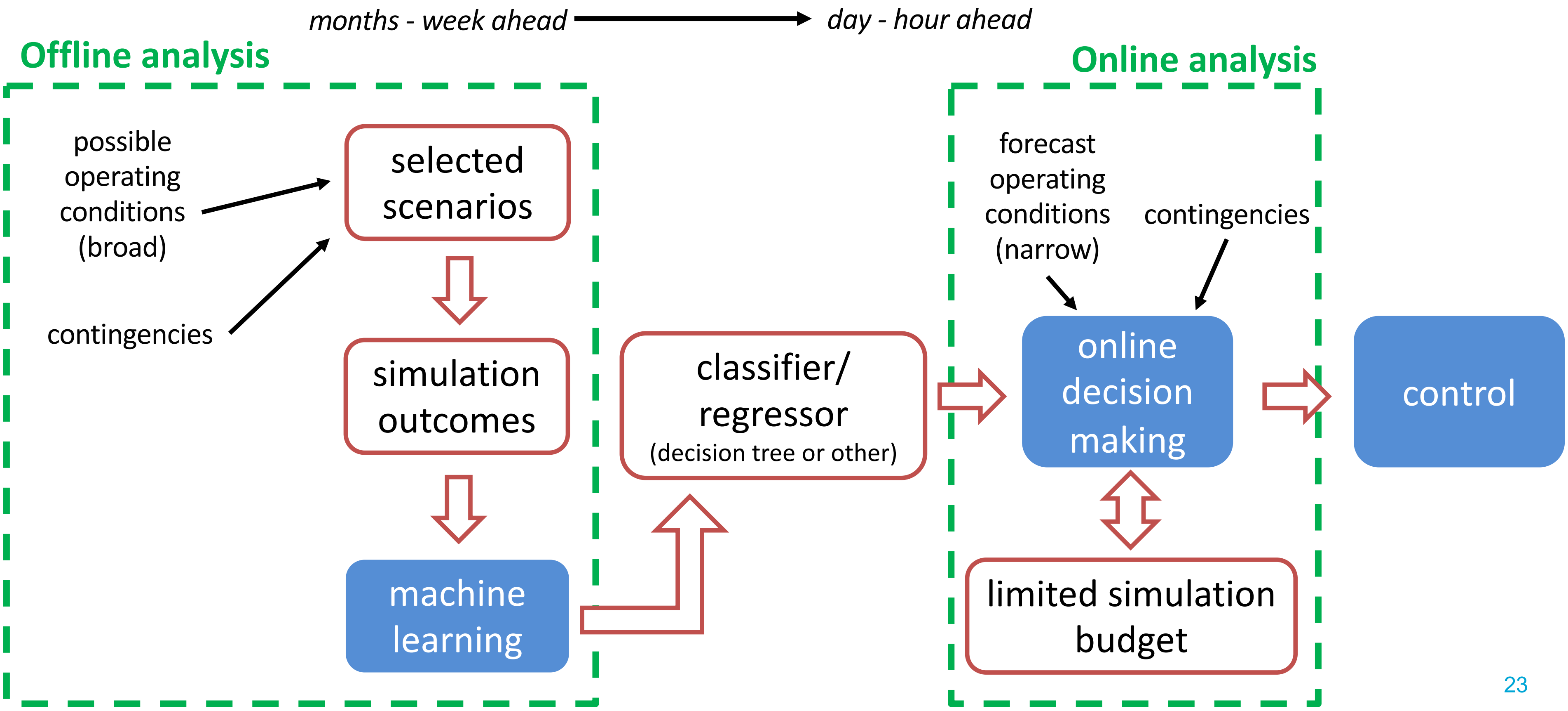
Online analysis

every few days

proxies of dynamic stability

improved decision-making

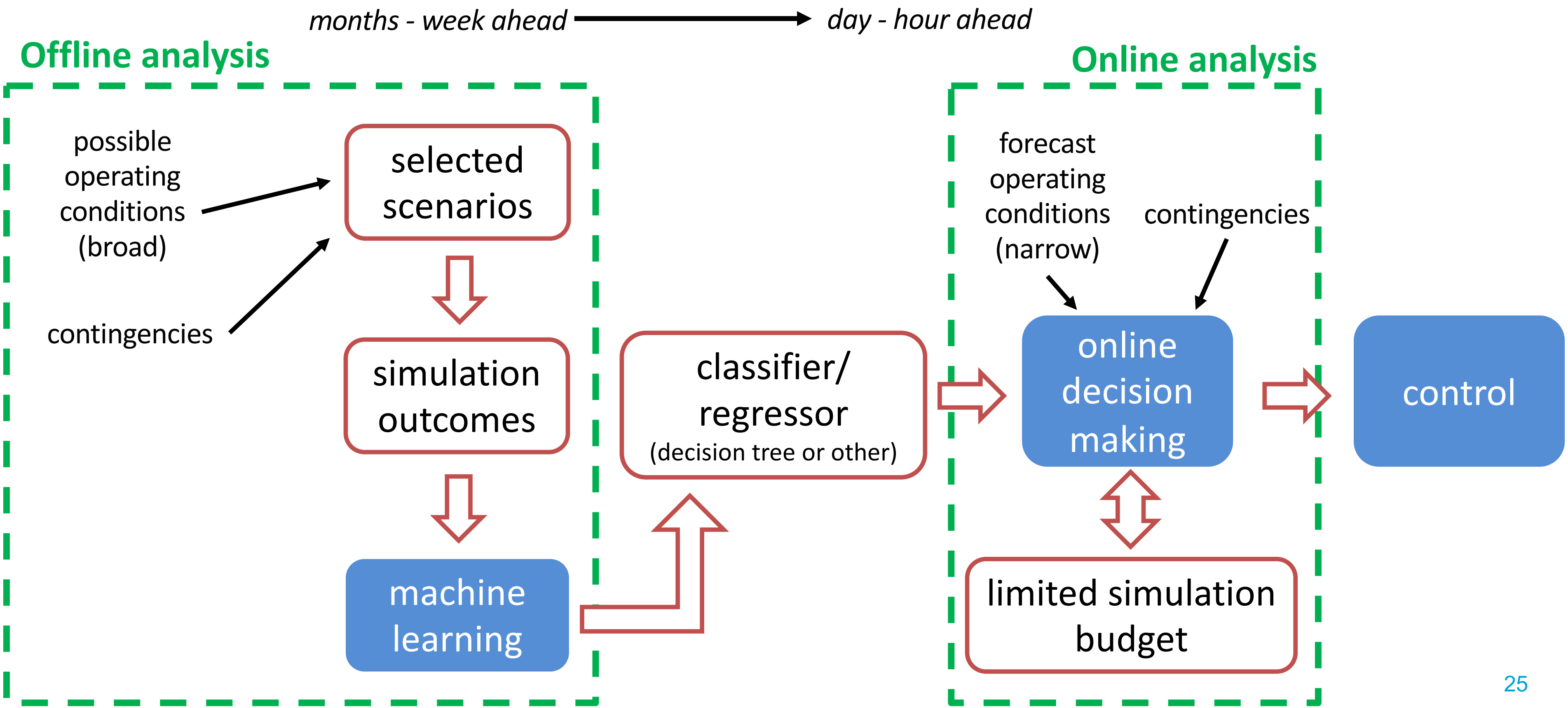
Using machine learning for DSA



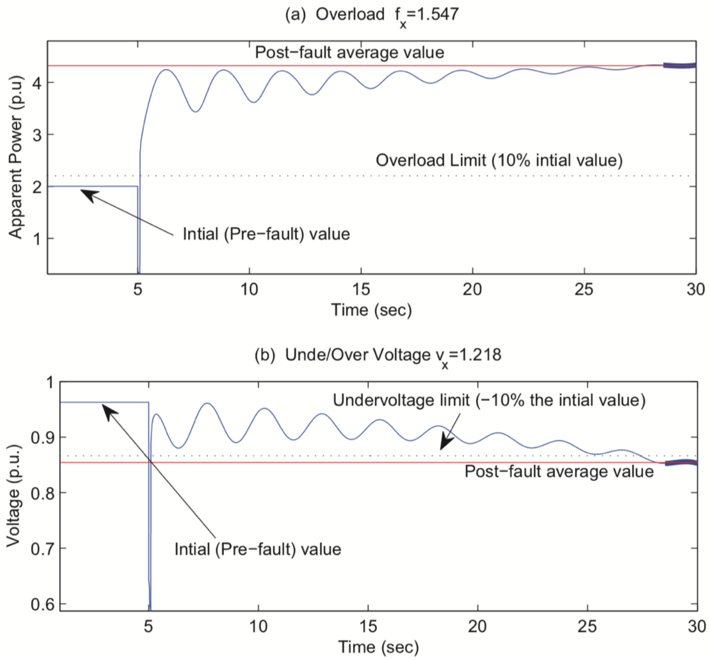
Offline workflow

Classifier training and evaluation

Using machine learning for DSA



Extracting indicators from simulations



Examples

- Small stability [S&V 2014]
  - Mode identification
- Static [S&V 2015]
  - Overload
  - Voltage
- Dynamic [S&V 2015]
  - Integrated square generator angle

F. R. S. Sevilla and L. Vanfretti, "A small-signal stability index for power system dynamic impact assessment using time-domain simulations," *2014 IEEE PES General Meeting*

F. R. S. Sevilla and L. Vanfretti, "Static stability indexes for classification of power system time-domain simulations," *2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*

Mapping onto classification problem

Features

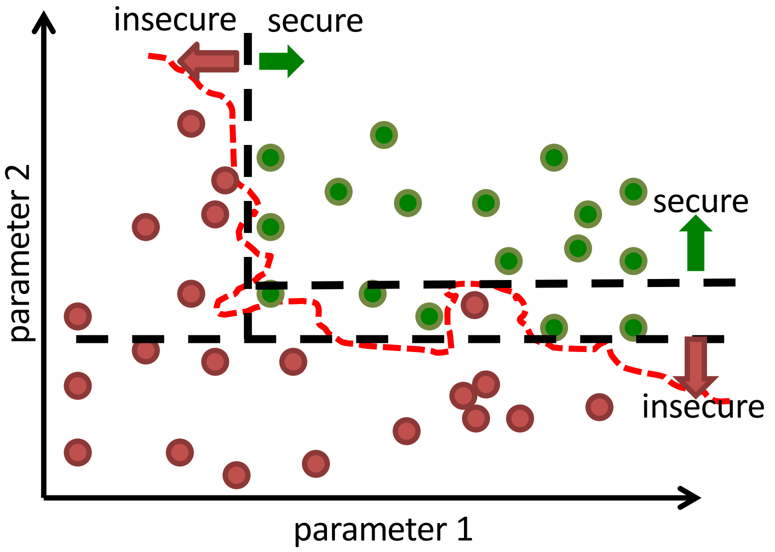
- Nodal voltages angles, active/reactive injections
- Line active/reactive power
- Topological features

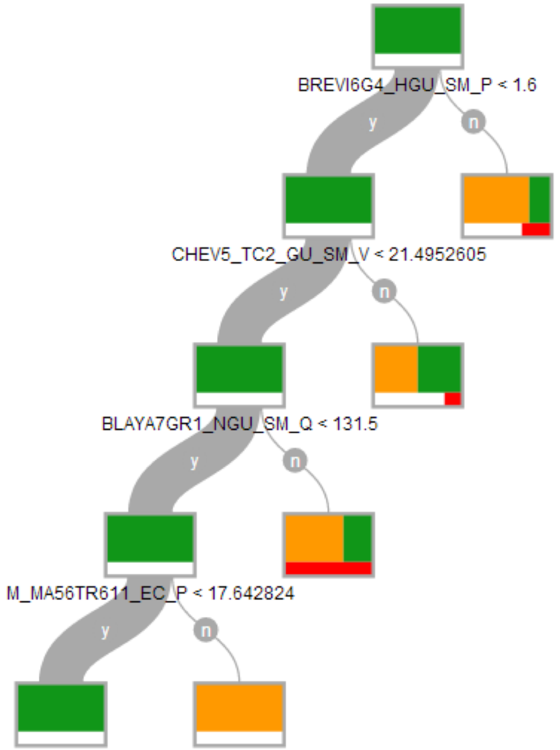
Outcomes

- Secure/insecure, **with respect to each contingency** [classification]
- Distance to security boundary [regression]

Decision trees for power system security

Two-dimensional example



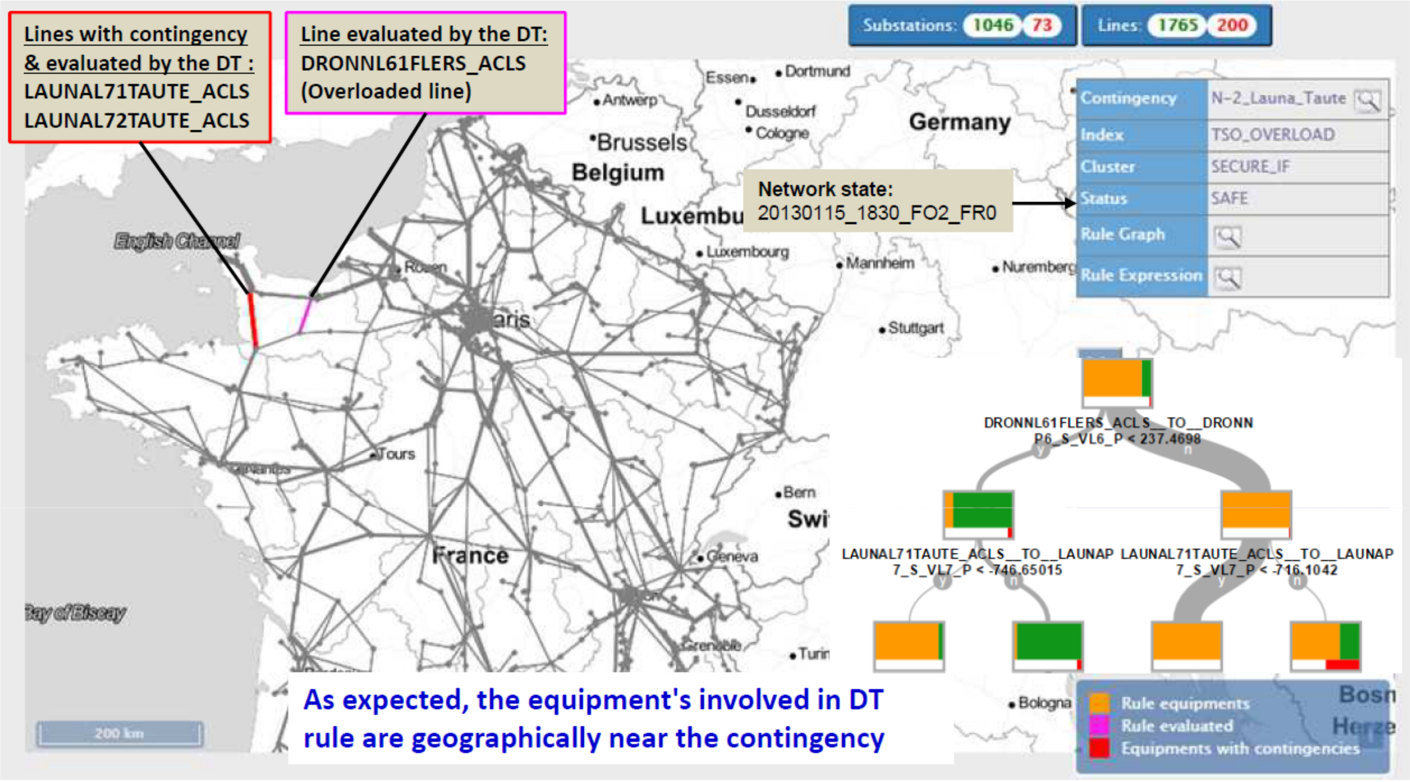


Decision trees:
- Limited **expressive** power
- Fantastic **interpretability**

TUDelft
Imperial College London

Example rule

Security rule for N-2 outage for Launay - Taute line in Normandy



Alternative: random forest (Breiman, 2001)

- Good **expressive power**
- Limited **interpretability**
- Very few **hyperparameters**

(a personal favourite)

Use an *ensemble* of decision trees. For condition x:

$$\mathrm{s_c}(x) \cong \frac{1}{K}\sum_{k=1}^{K} \hat{f}_c^k\,(x)$$

$$\hat{f}_c^k(x) \in \{0,1\}$$

Classifier:

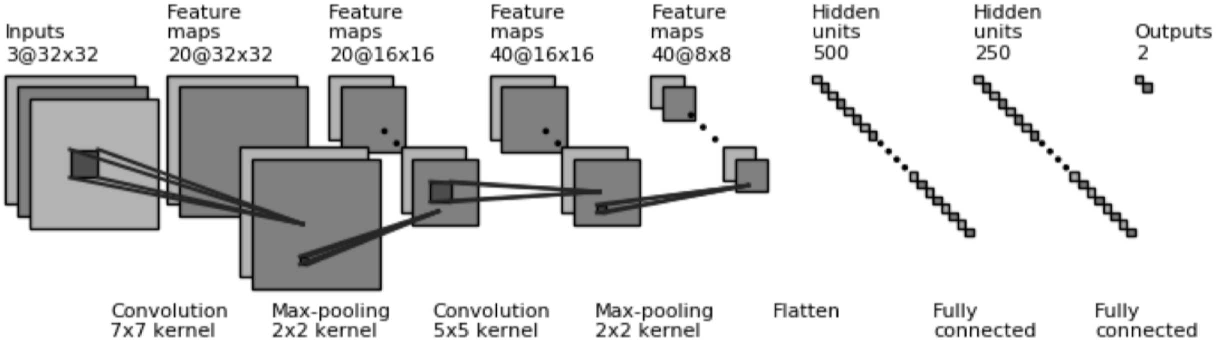
$$\hat{f}_c(x) = \begin{cases} 0\ if\ s_c(x) \leq 0.5 \\ 1\ if\ s_c(x) > 0.5 \end{cases}$$

Individual decision trees $\hat{f}_c^k(x)$ are randomized in two ways:

- Random subset of features for training
- Bootstrap aggregation (‘*bagging*’), i.e. random resampling of training data *with replacement*.

TU Delft
Imperial College London

Alternative: Neural networks



- Great **expressive power**
- Limited **interpretability**
- Many **hyperparameters**

TABLE I
REPRESENTATION OF CNN ARCHITECTURE.

Layer	Shape	Weights size	Bias size	# param
Input	[b,162,162,3]			
Conv1	[b,162,162,20]	[9,9,3,20]	[20]	4.880
Max-pool1	[b,81,81,20]			
Conv2	[b,81,81,40]	[7,7,20,40]	[40]	39.240
Max-pool2	[b,40,40,40]			
Conv3	[b,40,40,80]	[5,5,40,80]	[80]	80.080
Max-pool3	[b,20,20,80]			
Flatten	[b,32000]			
FC1	[b,250]	[32000,250]	[250]	8.000.250
FC2 (output)	[b,2]	[250,2]	[2]	502
Total				**8.124.952**

Arteaga, J-M. H., Hancharou, F., Thams, F., & Chatzivasileiadis, S. *Deep Learning for Power System Security Assessment*. In *Proceedings of IEEE Powertech 2019* IEEE.

TUDelft
Imperial College London

Prediction = making mistakes

		Predicted Class: Secure	Predicted Class: Insecure
True Class	Secure	**True positive**	**False negative**
True Class	Insecure	**False positive**	**True negative**

$$\text{(Simple) classification error} = \frac{N_{FP}+N_{FN}}{N_{FP}+N_{TP}+N_{FN}+N_{TN}}$$

HPC implementation

10,000 operating points

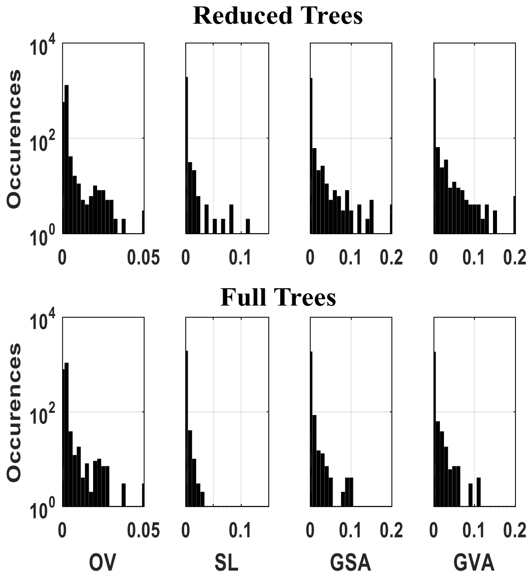
1980 contingencies

11 stability indicators

<0.5% average classification error

CURIE Supercomputer (10,000 cores)

Histogram of test errors for four indices



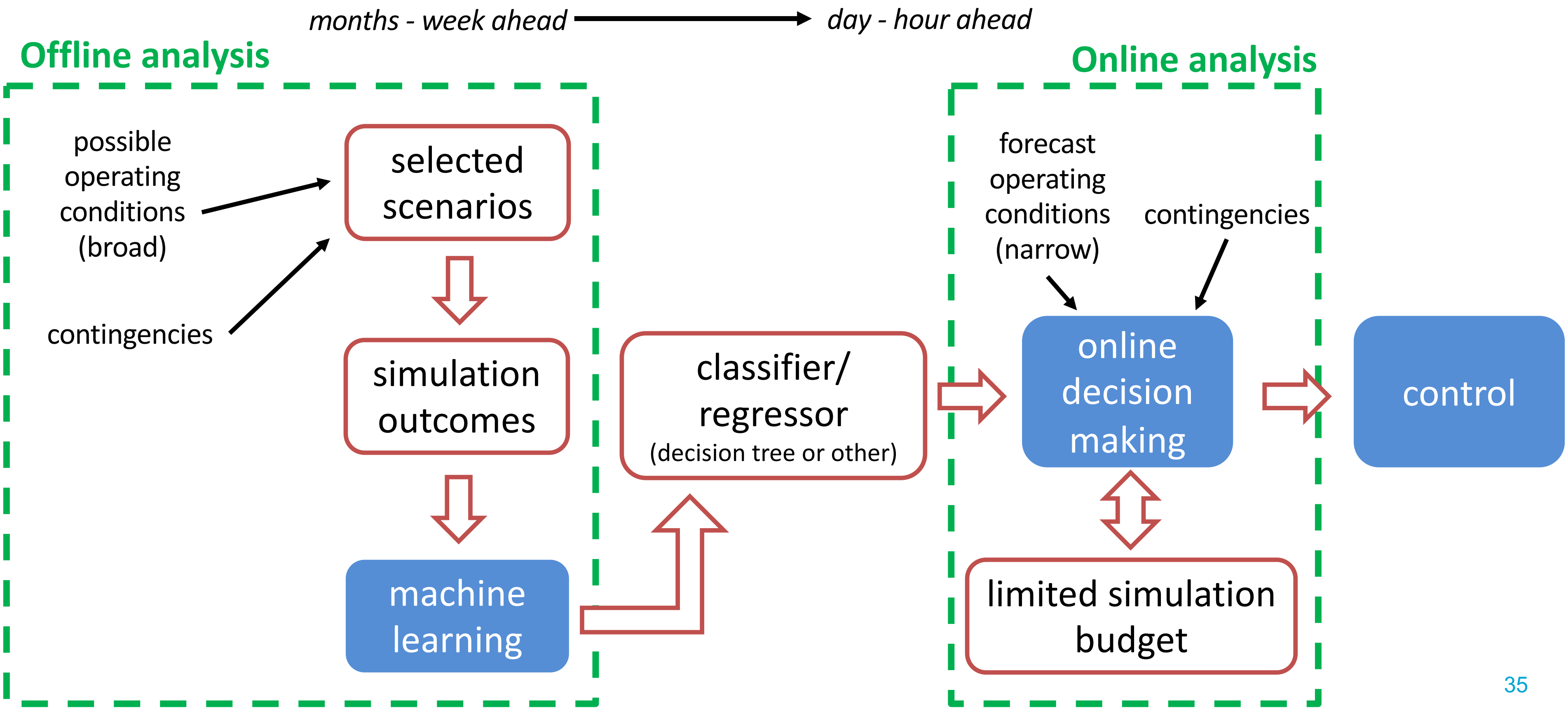
I. Konstantelos, G. Jamgotchian, S. Tindemans, P. Duchesne, S. Cole, C. Merckx, G. Strbac and P. Panciatici, “Implementation of a Massively Parallel Dynamic Security Assessment Platform for Large-Scale Grids”, IEEE Transactions on Smart Grid, May 2017.

Offline workflow

Selection and generation of training data

TU Delft
Imperial College London

Using machine learning for DSA

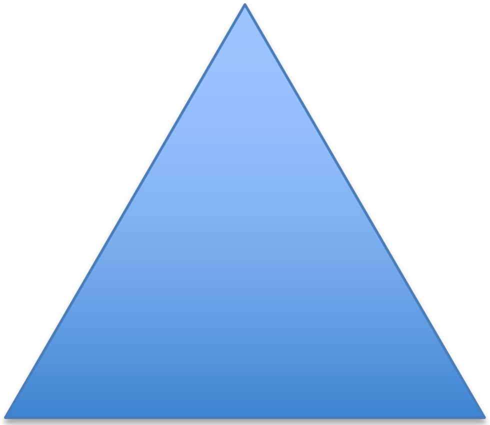


Challenge

Generate a set of training scenarios, such that:

- the classifier has a low error rate, …
- for known and unknown scenarios, …
- with finite computational resources.

Three approaches in the literature



Approach 1:
Use all imaginable states

- E.g. upper/lower bounds for injections
- Generates *many* infeasible or unlikely states
- Need clever methods to ‘zoom in’ (see table). *But is it enough?*

TABLE I
UNCLASSIFIED INPUT VOLUMES FOR PGLIB-OPF NETWORKS

case	$\lvert x \rvert$	V^{BT}	$\lvert HP \rvert$	V^{HP}	$\frac{-\log_{10}(V)}{\lvert x \rvert}$
AC-OPF without N-1 security and without uncertainty					
case3_lmbd	4	6.3e-02	28	3.3e-02	37.0%
case5_pjm	7	1.0e+00	99	6.9e-03	30.9%
case14_ieee	6	2.4e-01	54	6.9e-04	52.7%
case24_ieee_rts	20	9.2e-01	184	2.3e-06	28.2%
case30_ieee	7	6.2e-03	61	8.8e-06	72.2%
case39_epri	19	9.9e-02	203	7.0e-08	37.7%
case57_ieee	10	3.8e-02	231	4.9e-06	53.1%
case73_ieee_rts	62	1.0e+00	608	6.1e-16	24.5%
case118_ieee	72	1.7e-02	1000	1.6e-17	23.3%
case162_ieee_dtc	23	6.1e-04	371	1.5e-11	47.1%
case200_tamu	69	9.3e-01	1000	6.0e-11	14.8%
case300_ieee	125	1.0e-12	1000	3.4e-40	31.6%
case500_tamu	111	8.6e-02	1000	5.4e-26	22.8%
AC-OPF considering N-1 security and uncertainty					
case39_epri	25	2.6e-01	271	2.0e-05	18.8%
case162_ieee_dtc	29	2.2e-04	394	6.0e-10	31.8%
Median all cases	23	8.6e-02	271	7.0e-08	31.6%

A. Venzke, D.K. Molzahn, S. Chatzivasileiadis,(2019). Efficient Creation of Datasets for Data-Driven Power System Applications, arXiv:1910.01794

TUDelft
Imperial College London

Approach 2: Use historical data

- Historical states are (almost) guaranteed to be relevant
- They may not be *sufficient*
  - in volume
  - in variety
- We need to ‘enrich’ (interpolate/extrapolate) the historical data

Challenge: complex dependency patterns

I. Konstantelos, M. Sun, S. H. Tindemans, S. Issad, P. Panciatici and G. Strbac, "Using Vine Copulas to Generate Representative System States for Machine Learning," in *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 225-235, Jan. 2019.

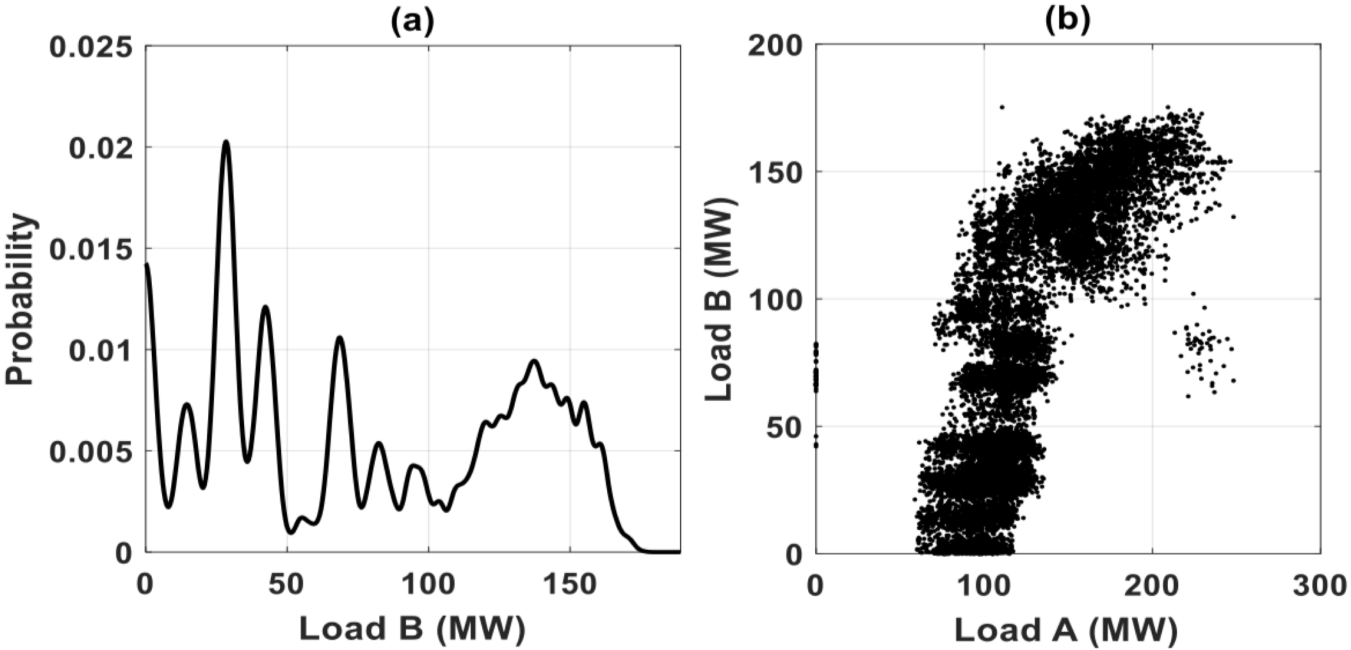
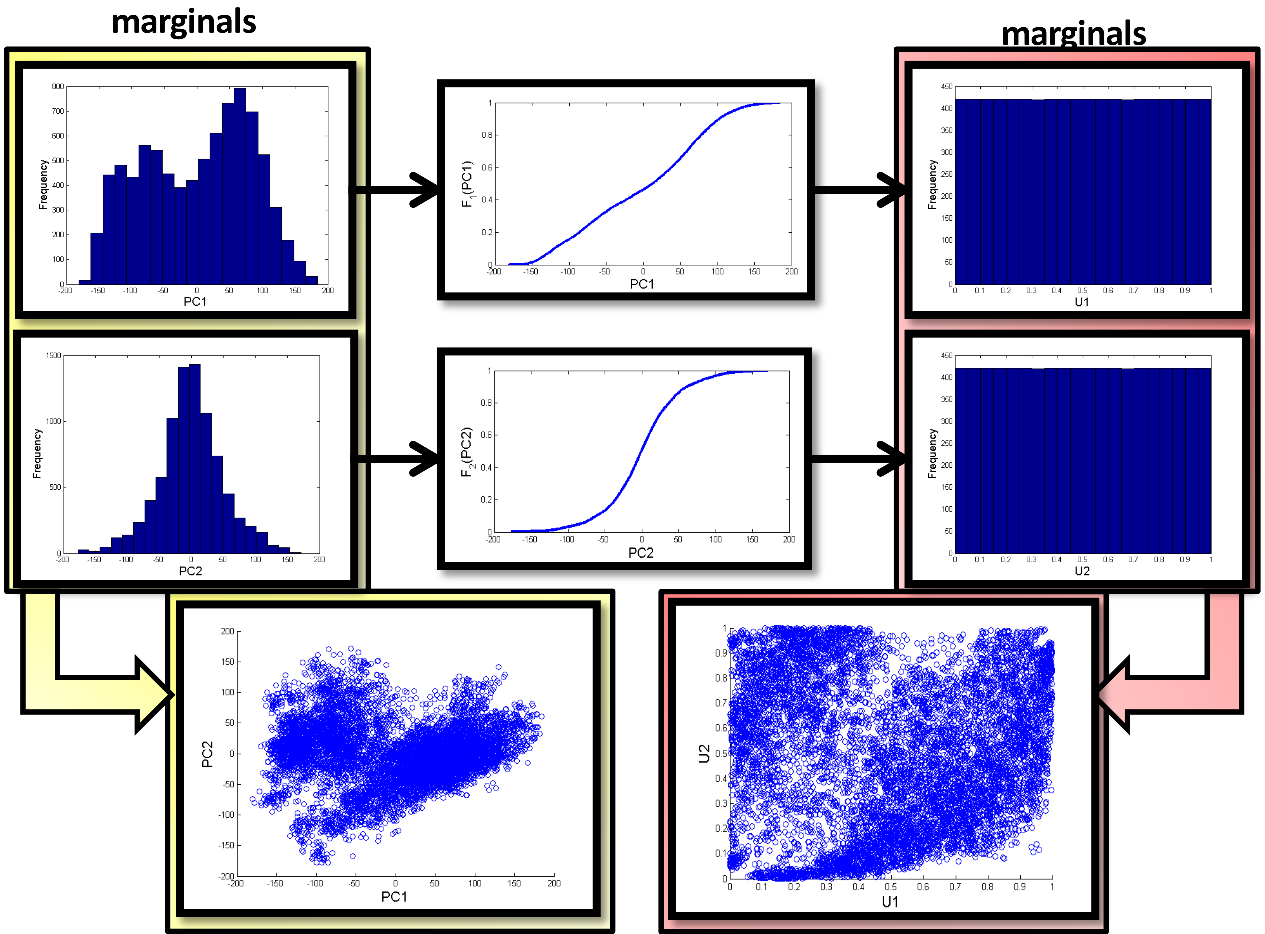


Fig. 1. (a) Marginal probability distributions of 5-minute load measurements over 3 months from a bus in the region of Nancy, France (March 2012). (b) Non-linear dependence between load measurements of two other buses in the same region.

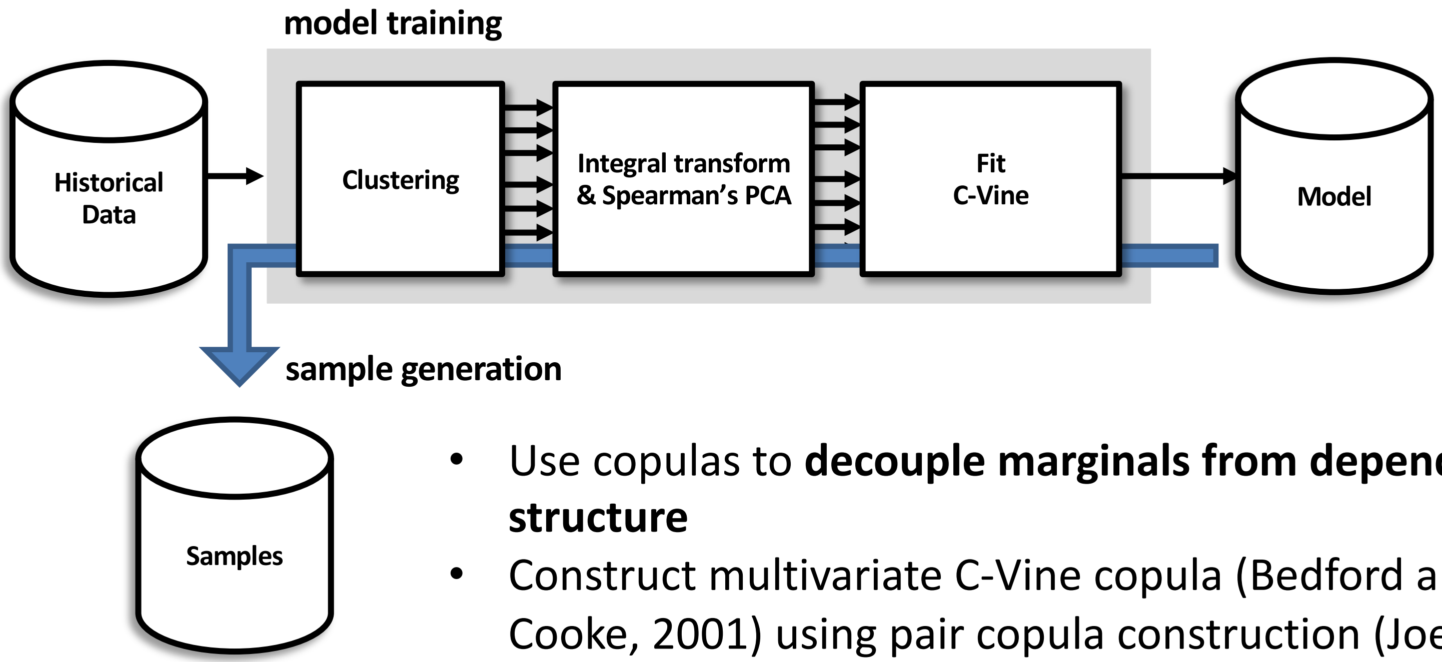
TUDelft
Imperial College London

Decoupling marginal distributions from dependence



41

The modelling pipeline



- Use copulas to **decouple marginals from dependence structure**
- Construct multivariate C-Vine copula (Bedford and Cooke, 2001) using pair copula construction (Joe, 1996)
- We truncate the C-Vine to limit impact of dimension

TUDelft
Imperial College London

Does it work?

- 128 variables (118 loads, 10 wind generators)
- 3 months data, 5-minute intervals
- 10 clusters; 97.5% variance used to select truncation; C-vine parametrised using Clayton, Frank, Gaussian, Gumbel, Student-t copulas (and rotations)

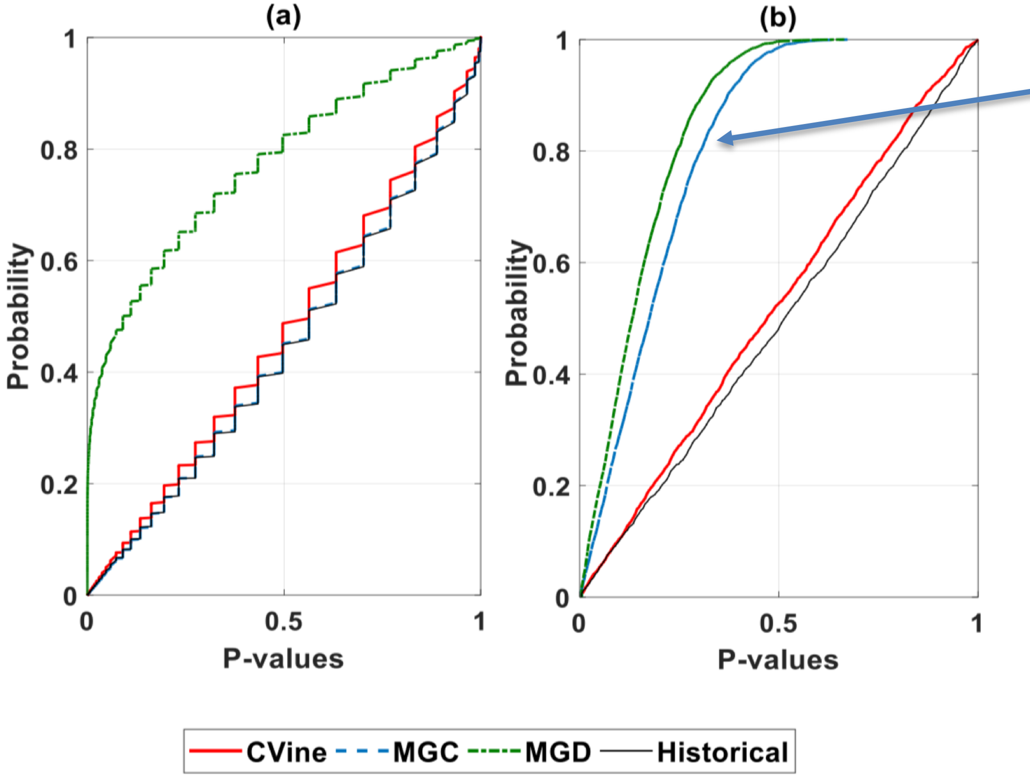
- Generated 40,000 samples
- Test on 1000 random subsets of 200 observations of sample and historical data
- Test metrics
  - Kolmogorov-Smirnov for marginals
  - Energy test (Aslan & Zech, 2005)

TU Delft
Imperial College London

Statistical match to historical data

Marginals (Kolmogorov Smirnov)

Full distribution (energy test)



multivariate
Gaussian
copula

TU Delft
Imperial College London

Application test

Model

- IEEE 118 bus system
- Set of four line outages, analyses independently
- Dispatch determined using OPF
- Post-fault generation redispatch of +- 10% permitted

Training and testing

- For a given state, verify whether post fault solution exists
- Classification using decision tree (Matlab 2017a default parameters)
- Training on bus angles and post-fault violations (true/false)
- 10-fold cross-validation

Impact on machine learning

four different contingencies

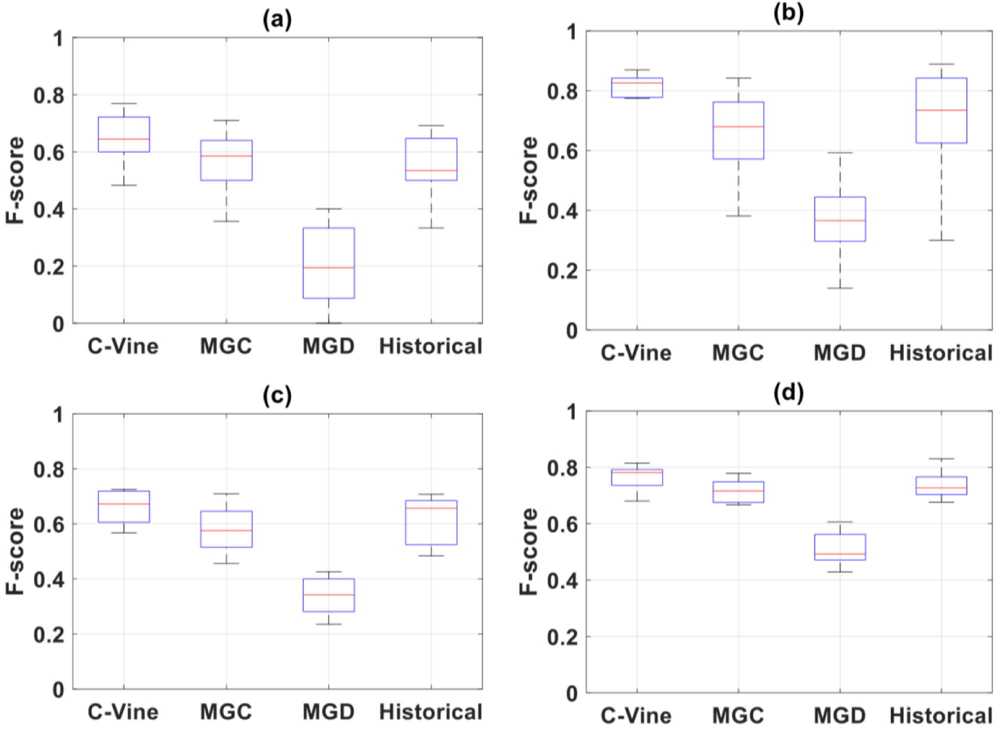
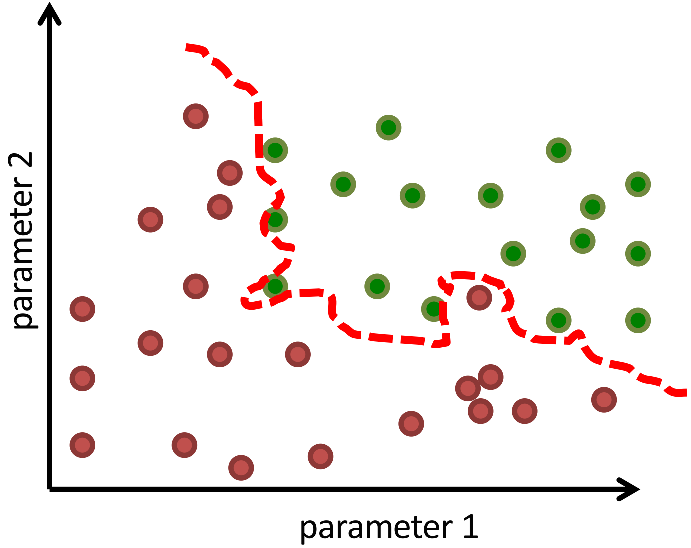


Fig. 7: F-score boxplots for contingencies (a) L148 (b) L139 (c) L54 (d) L71.

MGC = multivariate Gaussian copula ; MGD = multivariate Gaussian distribution

Approach 3: Improving sample efficiency

Classifiers are computed based on point samples



Obvious bright ideas

- Don't analyse situations that would never occur
- Don't analyse situations for which the outcome is obvious

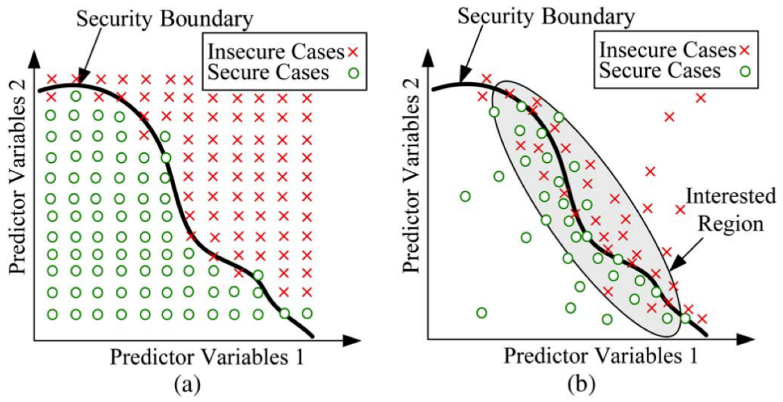
TU Delft
Imperial College London

Parametric approaches

Approaches typically rely on three assumptions:

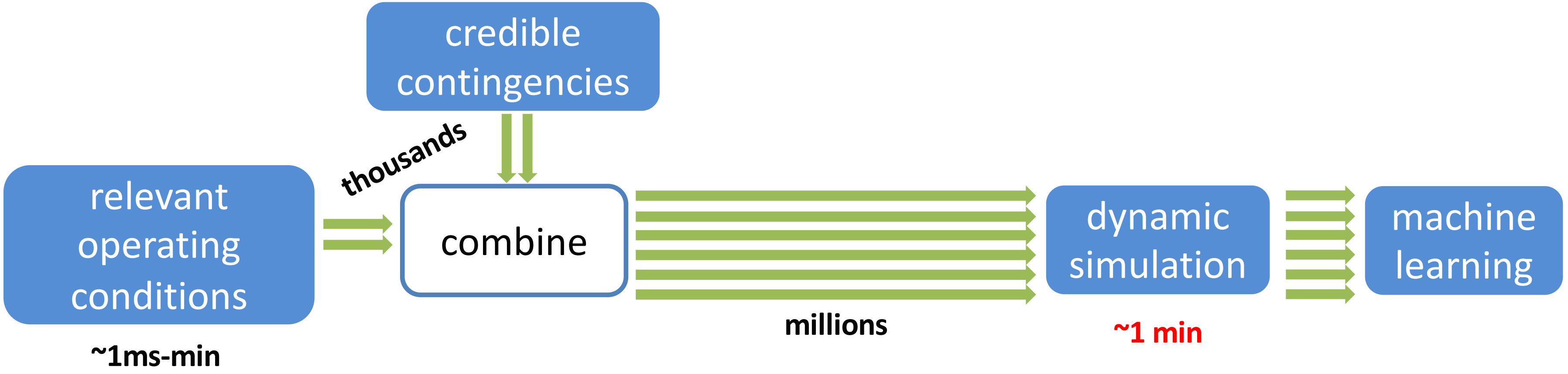
- Meaningful definition of 'distance' from the security boundary.
- 'Easy' sampling distributions.
- 'Nice' properties of the security boundary.

Chengxi Liu et al. (2014), IEEE Transactions on Power Systems



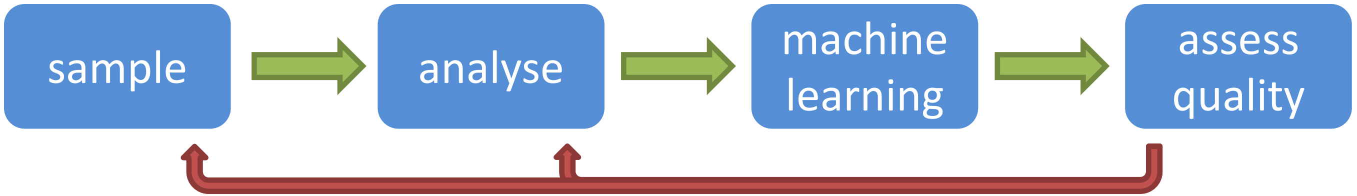
TU Delft
Imperial College London

The basic offline learning process



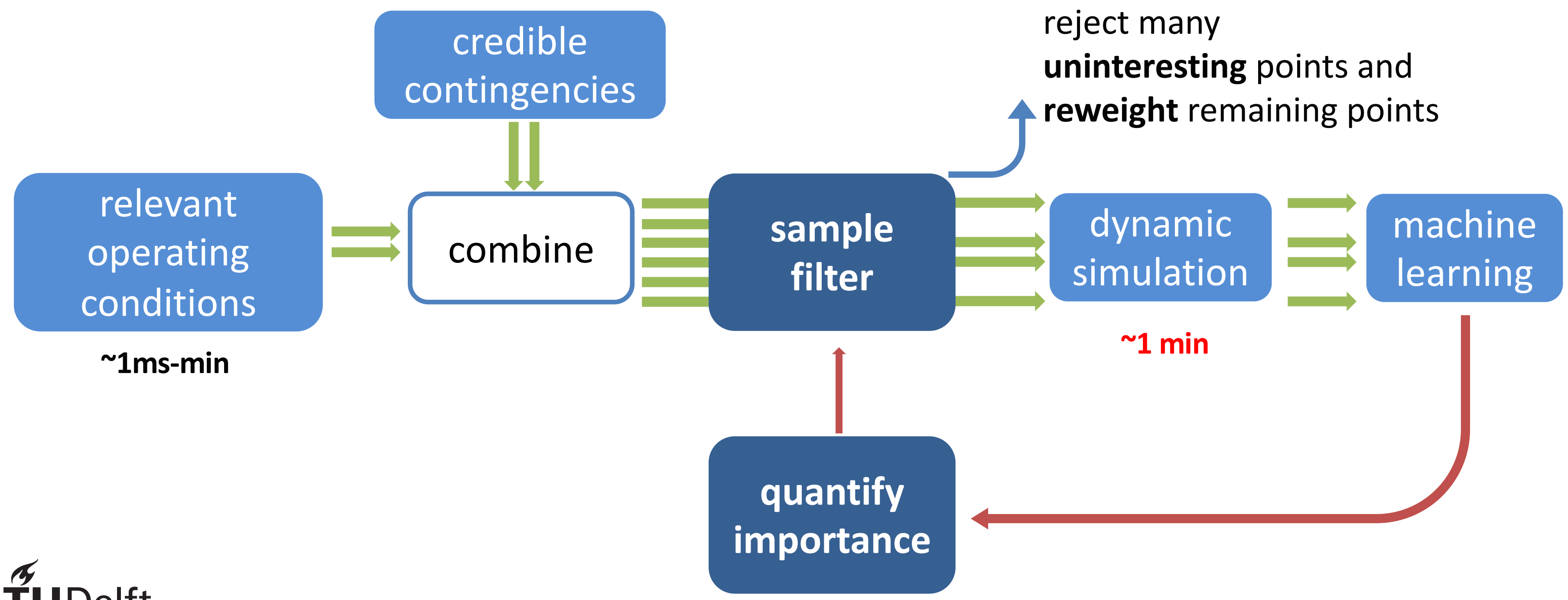
Research: the *active learning* approach

- Passive learning
  - Learner does not influence data gathering process
  - Inherently linear



- Active learning (e.g. Settles, *Active Learning,* 2012)
  - Generate data that is useful to the Learner
  - Machine learning equivalent of ‘optimal design of experiments’
  - Applied when sample generation **and/or** analysis is expensive

Active learning process



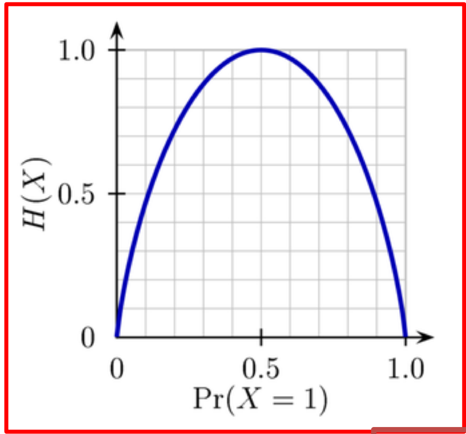
TUDelft
Imperial College London

Defining importance

1. For condition x and contingency c, predict score using ensemble method (e.g. random forest)

$$s_c(x) \cong \frac{1}{K}\sum_{k=1}^{K} \hat{f}_c^k(x)$$

$$\hat{f}_c^k(x) \in \{0,1\}$$



2. Define importance $I(x,c)$ by measuring *ensemble disagreement* (entropy)

$$I(x;c) = -\frac{1}{\log_2 2}\left[s_c(x)\log_2 s_c(x) + (1 - s_c(x))\log_2(1 - s_c(x))\right]$$

Sample filter

Accept/reject algorithm

$$\Pr(accept|x, c) = b + (1 - b)I(x, c)$$

Results in:

$$f_{bias}(x, c) = e\, f_0(x) + (1 - e) f_{int}(x, c)$$

where

$f_0(x) = \text{unbiased distribution}$

$f_{int}(x, c) \propto I(x, c)$

$e = \text{effective exploration}$

TU Delft
Imperial College London

Small ‘proof of concept’– IEEE 118 bus model

Procedure

1. Check DC load flow feasibility
2. Map actual measurements from French transmission grid onto network
3. Generate 500,000 random load samples
4. For each sample, analyse the impact of every single line outage (186)

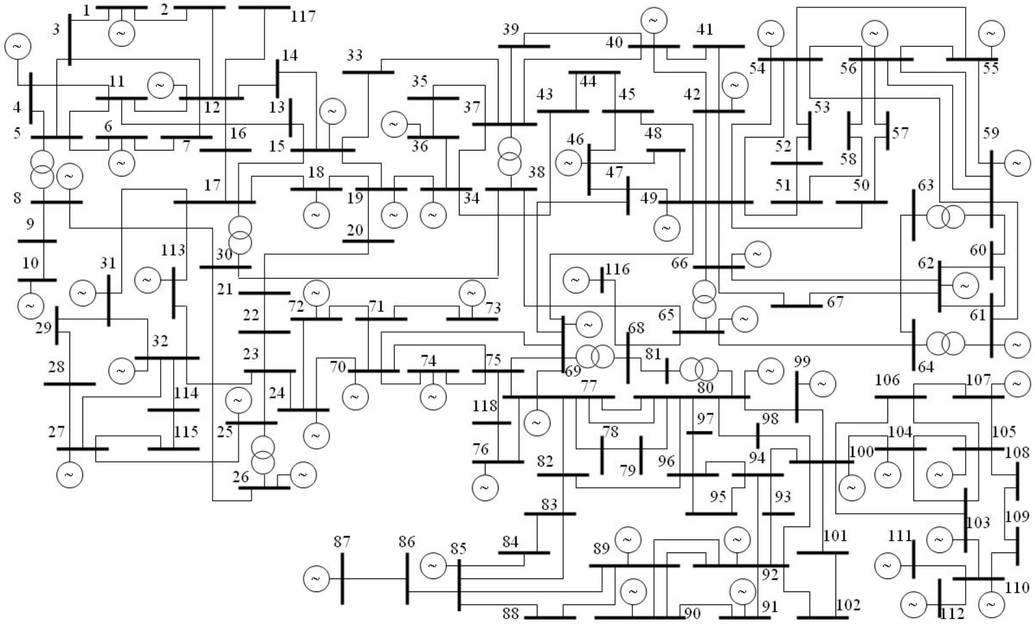
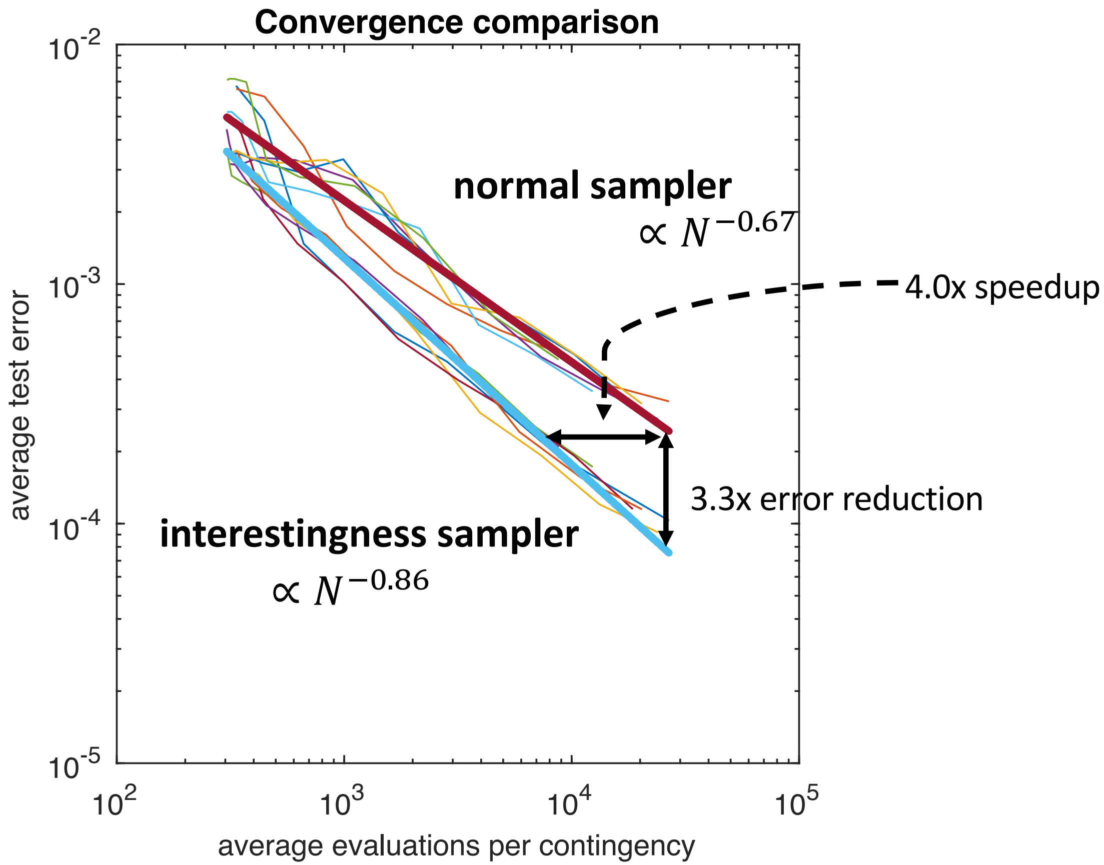


Рис.1. IEEE тестовая схема, состоящая из 118 узлов

Study details

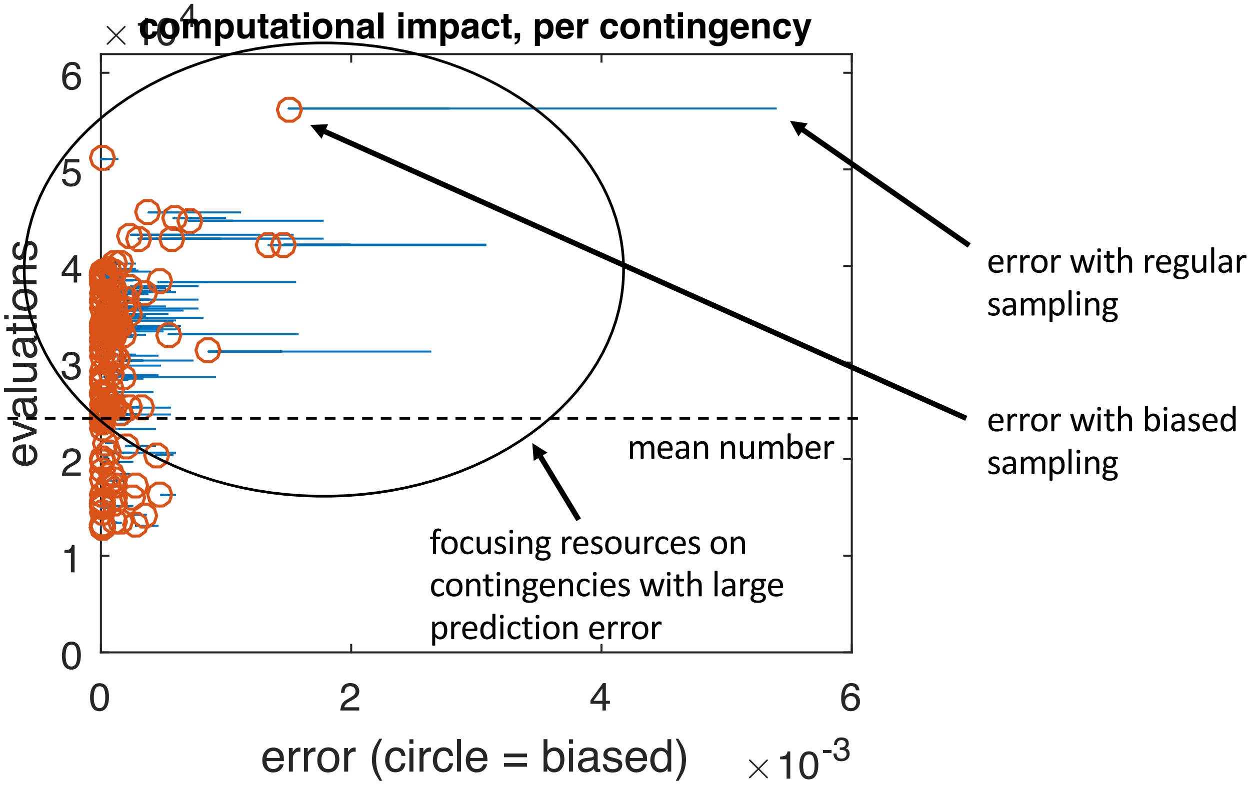
- Offline simulation of importance sampling
  - 450,000 states are classified as ‘good’ or ‘bad’ in combination with 186 contingencies
  - Use results as an unbiased ‘pool’ of samples
- Importance sampling parameters
  - Exploration fraction $e = 0.5$ (conservative choice: max 2x slowdown)
  - Minimum acceptance rate 1%
  - Update importance functions in batches
- Validation using 50,000 states x 186 contingencies
  - Compare classification errors against unbiased samples using identical computational budget

TU Delft
Imperial College London



Results

- Measure performance across *all* decision trees for given number of evaluations
- Robust improvement, for various importance functions
- Active learning advantage improves with learning



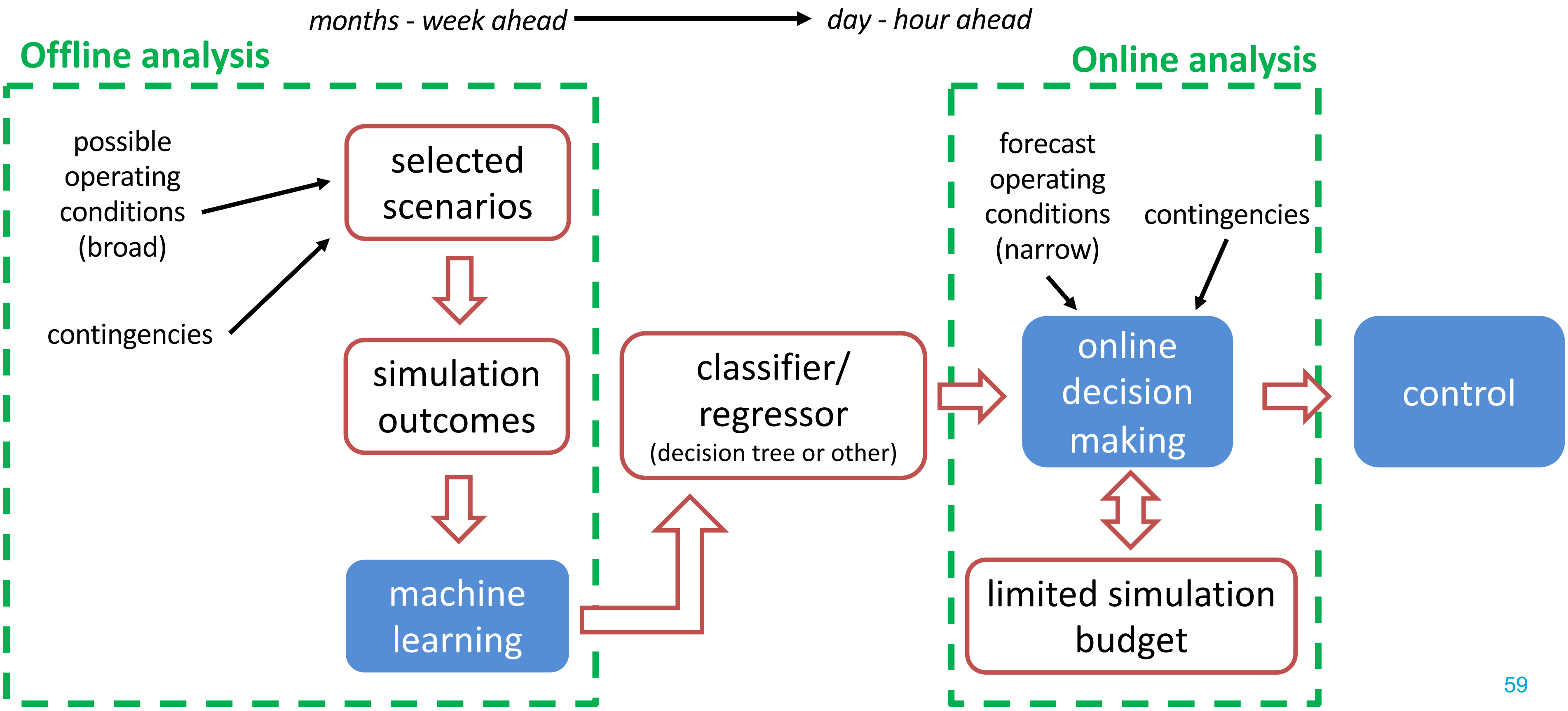
TUDelft
Imperial College London

Online workflow

Using machine learning to effectively operate simulations

TU Delft
Imperial College London

Using machine learning for DSA



Challenge

- Quantifying the risk of relying on a Machine Learning approach is not trivial
- When moving these Machine Learning approaches toward practical tools it is important to understand and manage the risk involved
- In comparison: Physics-based methods typically offer insights analytically or numerically for assessing the confidence of the output

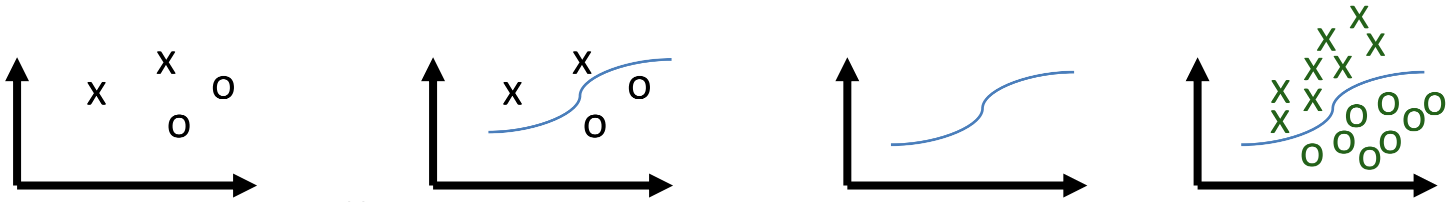
The machine learning approach

Close to online

- Study more potential operating conditions instead of only a few
- Replace dynamic simulations (slow) with a machine-learning-based estimator (fast)

Offline

- Prepare (train) an estimator on actual dynamic simulations using similar operating conditions

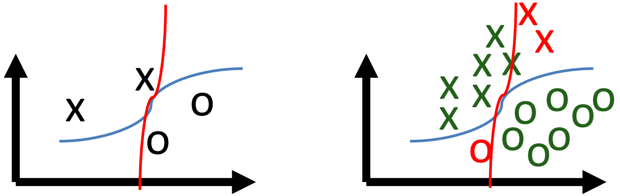


Offline

Close to online

Jochen Cremer, Ioannis Konstantelos, Goran Strbac, "From Optimization-based Machine Learning to Interpretable Security Rules for Operation", *IEEE Transactions on Power Systems, 2019*

The estimation can be inaccurate



Close to online

True Class		Predicted Class: Positive	Predicted Class: Negative
	Positive	**TP**	**FN**
	Negative	**FP**	**TN**

Estimated and actual stability limits are different!

<u>Two types of inaccurate predictions</u>:

FP: Is stable but we think it is unstable (**BAD**)

FN: Is unstable but we think it is stable (**VERY BAD!**)

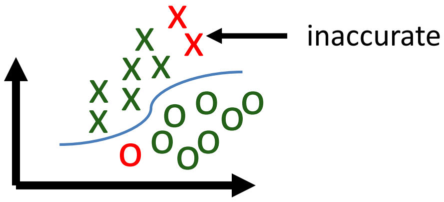
This can have a severe effect!

TU Delft

Imperial College London

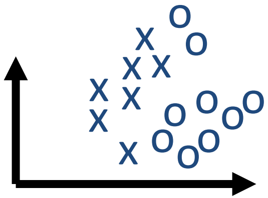
Combined approach

Machine-learning



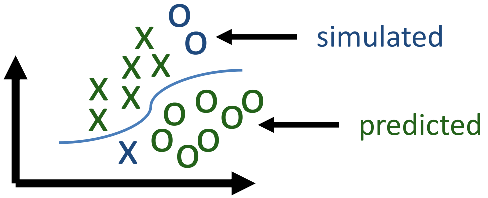
- Fast
- Sometimes inaccurate

Physics-based



- Slow
- Always accurate

Combined approach



TU Delft
Imperial College London

Understanding the costs of inaccurate predictions

		Predicted Class	
		Positive	Negative
True Class	Positive	**0**	C_{FN}
	Negative	C_{FP}	**0**

Cost skew: $C_{FN} \gg C_{FP}$

TU Delft
Imperial College London

What a classifier can do

Classify points

- is x positive?

Rank points

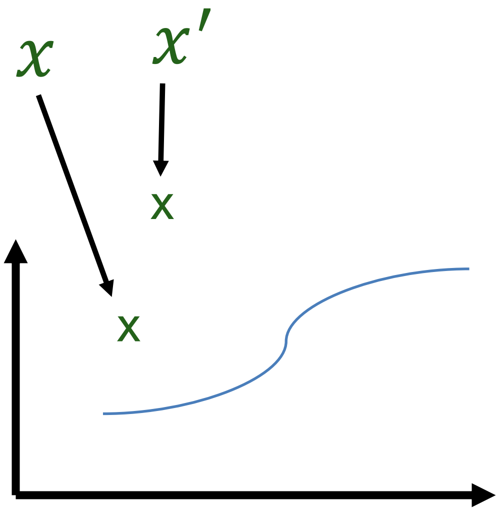
- Is x ‘more positive’ than x’?

Output a score $\boldsymbol{s(x)}$

- ‘How positive’ is x?

Output a probability estimate $\boldsymbol{\widehat{p}(x)}$

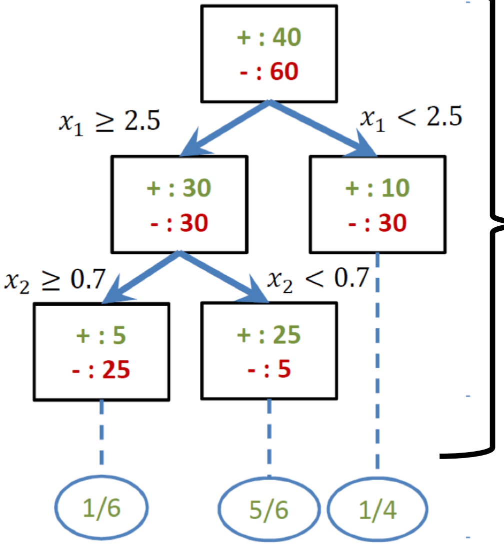
- What is the (estimated) probability that x is positive?



Nikos Nikolaou, *“Cost-sensitive learning with AdaBoost”*, 2017

Probability estimation is not trivial

- Typically, classifiers don't output probability estimates. **Scores** can be used
- E.g. decision tree



Based on training set

Scores s(x) as probability estimates $\widehat{\boldsymbol{p}}(\boldsymbol{x})$?

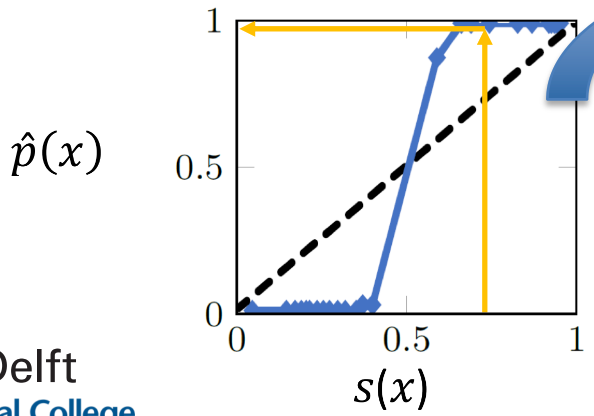
No!

Nikos Nikolaou, *"Cost-sensitive learning with AdaBoost"*, 2017

TU Delft
Imperial College London

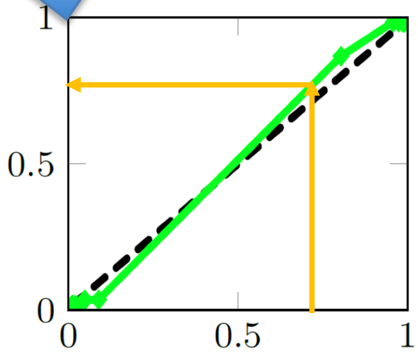
Calibration

- $s(x) \in [0,1]$ score for operating condition x
- A classifier is **calibrated** if $\hat{\boldsymbol{p}}(\boldsymbol{x}) \longrightarrow s(x)$, as $N \longrightarrow \infty$
- An intuition for a calibrated classifier: E.g., 70% of the operating points with $s(x)$ = 0.7 should belong to the 'stable' class



Platt scaling:

Find $\hat{p}(x) = \frac{1}{1+e^{A\,s(x)+B}}$



The risk of relying on machine learning

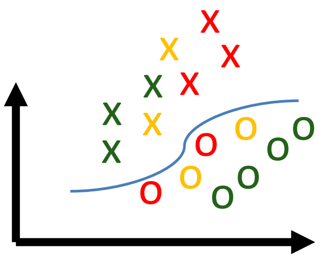
Machine-learning

$$R_{stable} = p_o p_c \widehat{\boldsymbol{p}}(\boldsymbol{x})\boldsymbol{C_{FN}}$$

$$R_{unstable} = p_o(1 - p_c)(1 - \widehat{\boldsymbol{p}}(\boldsymbol{x}))\boldsymbol{C_{FP}}$$

Predict class with lower risk

Residual risk: $R_{stable} \vee R_{unstable}$



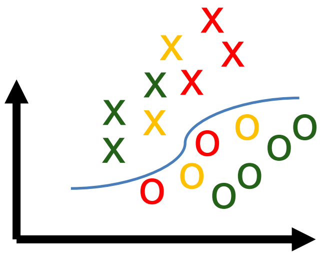
High risk

Medium risk

Low risk

TU Delft
Imperial College London

Combined approach

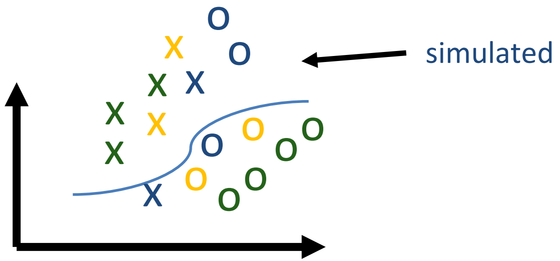


High risk

Medium risk

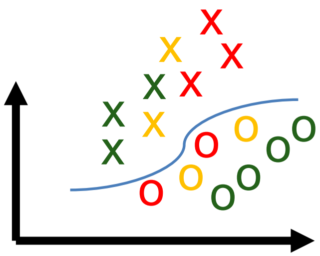
Low risk

Perform simulations on the operating conditions with high risk



Multiple contingencies

Contingency 1

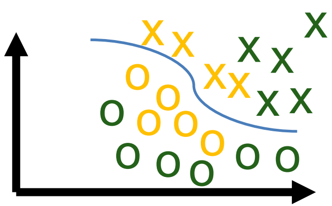


High risk

Medium risk

Low risk

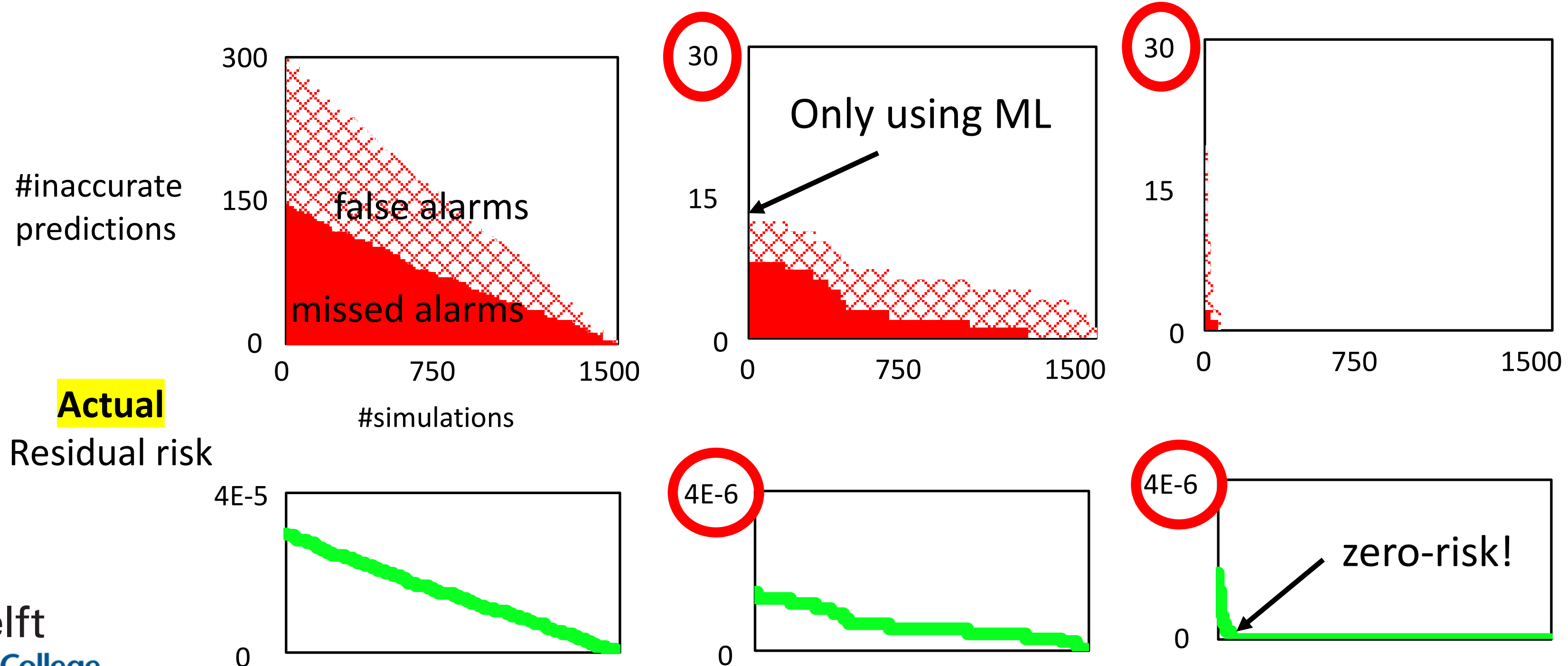
Contingency 2



Perform simulations according to risks

A case study

System: IEEE 6 bus system, 1500 conditions (stable 1322, unstable 178)
Machine-Learning: AdaBoost ensemble, cross-validation, train/test 70/30



No Machine-learning **Standard classifier** **Combined approach**

TUDelft
Imperial College London

Summary

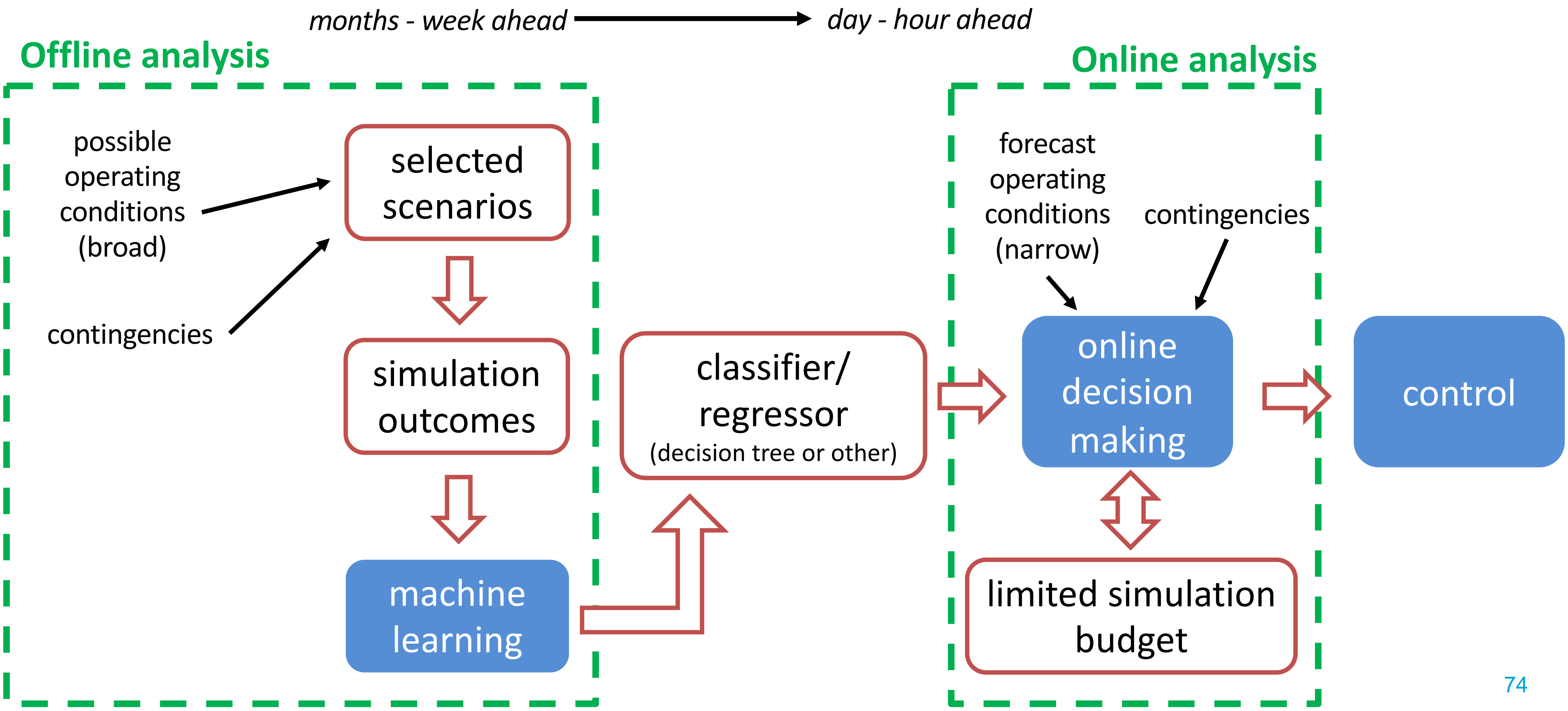
- Combining Machine Learning and current DSA methods results in fast and accurate security assessment
- Zero-risk can be achieved – allowing for no disadvantage when using Machine Learning
- Parameter estimations are uncertain

Online workflow

Using machine learning to ensure stability in operations

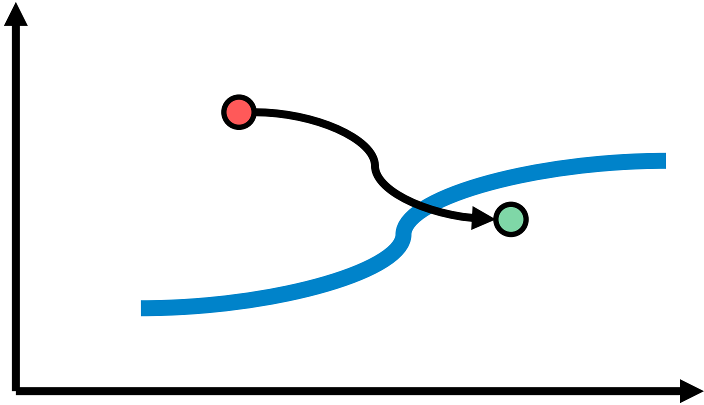
TU Delft
Imperial College London

Using machine learning for DSA



Challenge

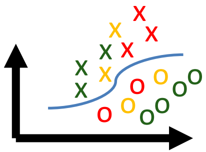
- Compute appropriate measures to re-establish stability once an unstable operating condition is detected
- How conservative do we need to be? Cost versus risk
- Stability is a complex system-level attribute
- Finding cost-effective measures to ensure stability is even more complex



TU Delft
Imperial College London

Our approach

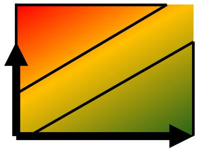
- Using Machine Learning to train a proxy implementable in a control approach



High risk

Medium risk

Low risk

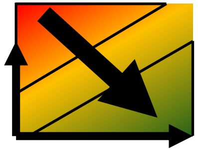


High risk of instability

Medium risk

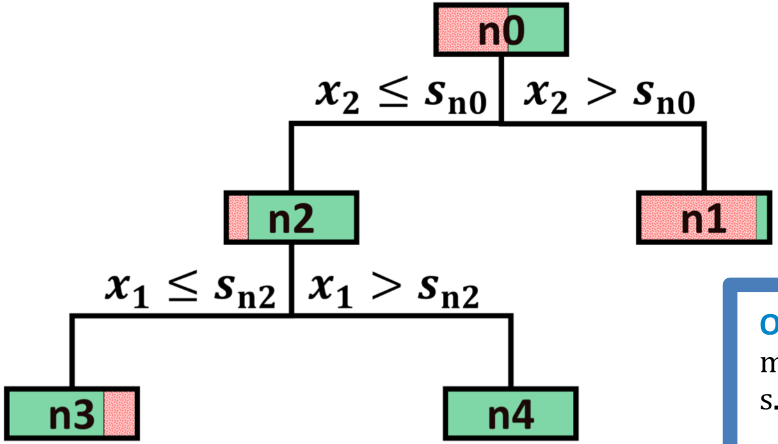
Low risk

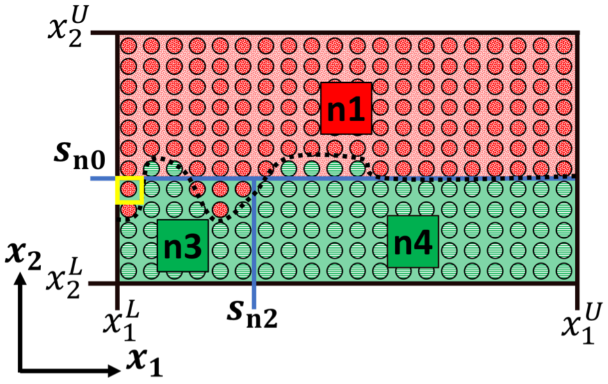
Find low-risk regions



Move operating condition to 'low risk' region

Classifier in control





OPF formulation: MILP

$$\min \, operating \; cost(x)$$

s.t.

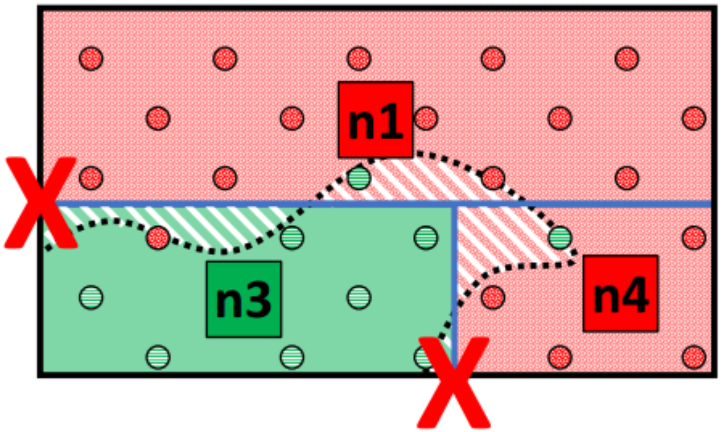
$$g(x) \leq 0$$
$$h(x) = 0$$
$$x \leq a_j + m_j(1 - y_j) \quad \forall j \in J$$
$$x > \tilde{a}_j + \tilde{m}_j(1 - y_j) \quad \forall j \in J$$
$$\sum_{j \in J} y_j = 1$$
$$m_j^i = \max\{a_{j'}^i : j' \in J\} - a_j^i$$
$$\widetilde{m_j^i} = \min\{\widetilde{a_{j'}^i} : j' \in J\} - \widetilde{a_j^i}$$

TU Delft
Imperial College London

Jochen Cremer, Ioannis Konstantelos, Simon Tindemans, Goran Strbac, "Sample-Derived Disjunctive Rules for Secure Power System Operation," 2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), Boise, USA 2018

Using a classifier in control is not trivial

- DT tries to approximate the true stability boundary
- Typical measure of DT's quality: classification error



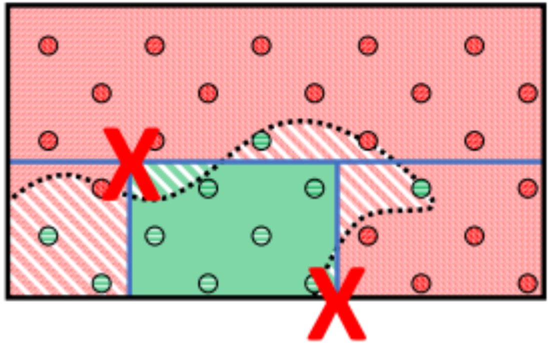
Class. error < 10%
Control error > 90%

Classification error cannot predict a rule's performance when used for inferring suitable mitigation control actions.

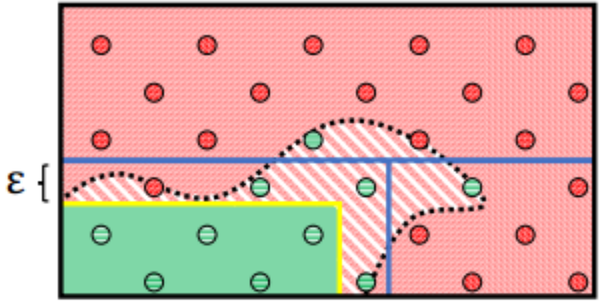
Jochen Cremer, Ioannis Konstantelos, Simon Tindemans, Goran Strbac, "Data-driven Power System Operation: Exploring the Balance between Cost and Risk", IEEE Transactions on Power Systems, 2018

Four different approaches

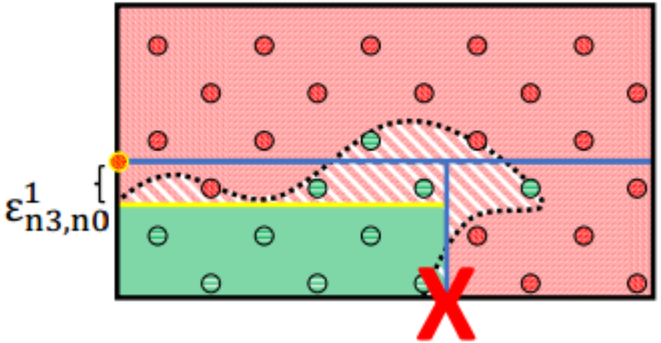
Asymmetric weighting



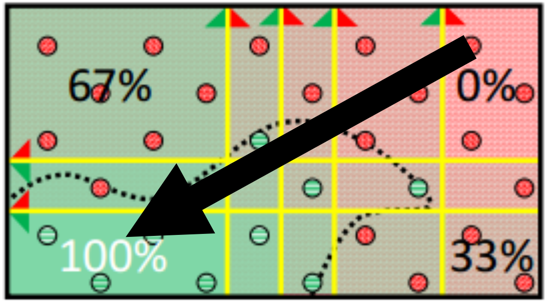
Single-ε



Condition-specific-ε



Probabilistic



Normalized operating cost

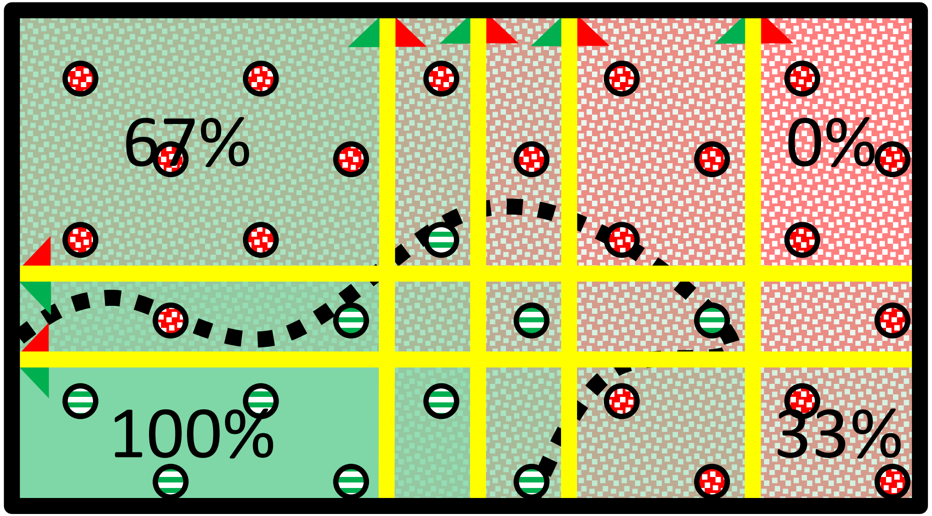
Estimated risk

$$\min\,(1-\alpha)f'(x) + \alpha\, R(x)$$

User-defined parameter

Jochen Cremer, Ioannis Konstantelos, Simon Tindemans, Goran Strbac, "Data-driven Power System Operation: Exploring the Balance between Cost and Risk", IEEE Transactions on Power Systems, 2018

Combination approach



Optimal Power Flow

$$\begin{aligned} \min \quad & (1-\alpha)\,\hat{\mathbb{E}}(\text{cost}) - \alpha\,\hat{\mathbb{E}}(\text{stability}) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \leq 0 \\ & x^L < x \leq x^U , \end{aligned}$$

Extension by index l for each learner

$$\begin{aligned} & x \leq a_{l,d} + m_d(1 - y_{l,d}) && \forall d \in D_l, \forall l \in L \\ & x > \tilde{a}_{l,d} + \tilde{m}_d(1 - y_{l,d}) && \forall d \in D_l, \forall l \in L \\ & \sum_{d \in D_l} y_{l,d} = 1 && \forall l \in L \\ & y_{l,d} = \{0, 1\} && \forall d \in D_l, \forall l \in L \\ & m_d = \max\{a_{d'} : d' \in D_l\} - a_{l,d} && \forall d \in D_l, \forall l \in L \\ & \tilde{m}_d = \min\{\tilde{a}_{d'} : d' \in D_l\} - \tilde{a}_{l,d} && \forall d \in D_l . \forall l \in L, \end{aligned}$$

Probability estimate for each learner

$$h_l = \sum_{d \in D_l} p_{l,d}\, y_{l,d} \quad \forall l \in L,$$

Soft voting

$$\hat{\mathbb{E}}(\text{stability}) = F = \sum_{l \in L} w_l h_l(x)$$

TU Delft
Imperial College London

The balance of cost and risk

- This approach balances economic cost of dispatching with the risk of relying on Machine Learning
- Each operating point has a unique solution
- This approach is a MILP in DC or MINLP in AC

TU Delft
Imperial College London

Case study

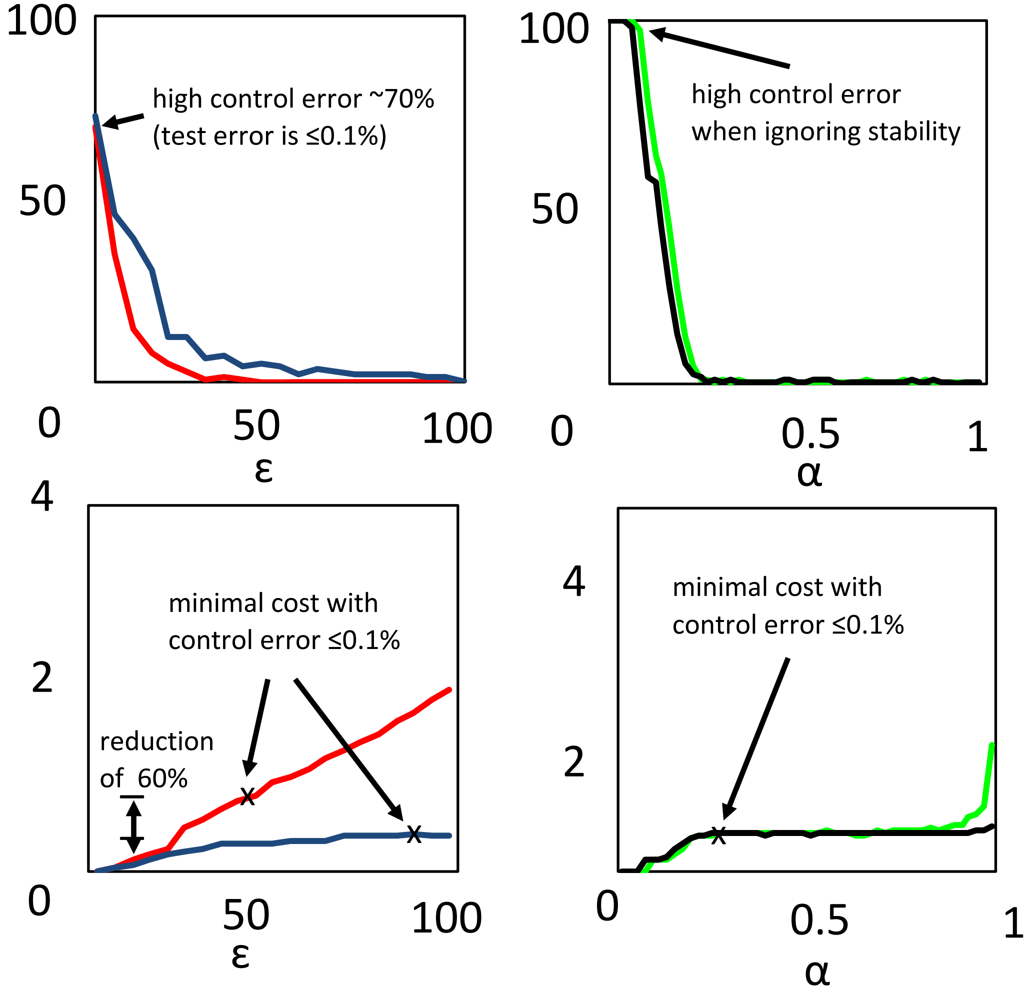
- IEEE 39 bus system
- Considering steady state security allows for SCOPF comparison
- Kumaraswamy distribution in loads
- Implemented in Python 3.5.2, scikit-learn
- Optimization environment: Pyomo, Solver: Gurobi 7.02

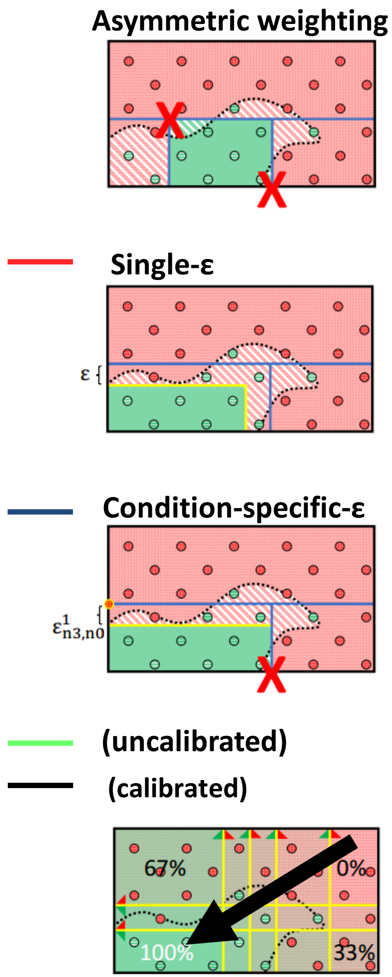
TU Delft
Imperial College London

Cost-risk

Control error *k* [%]

Cost increase *f* [%]





TU Delft

Imperial College London

Computational results

- MILP condition-specific approach: 82 binaries
- MILP probabilistic approach: 21 binaries

- Solver time for all approaches: <0.1sec

Summary

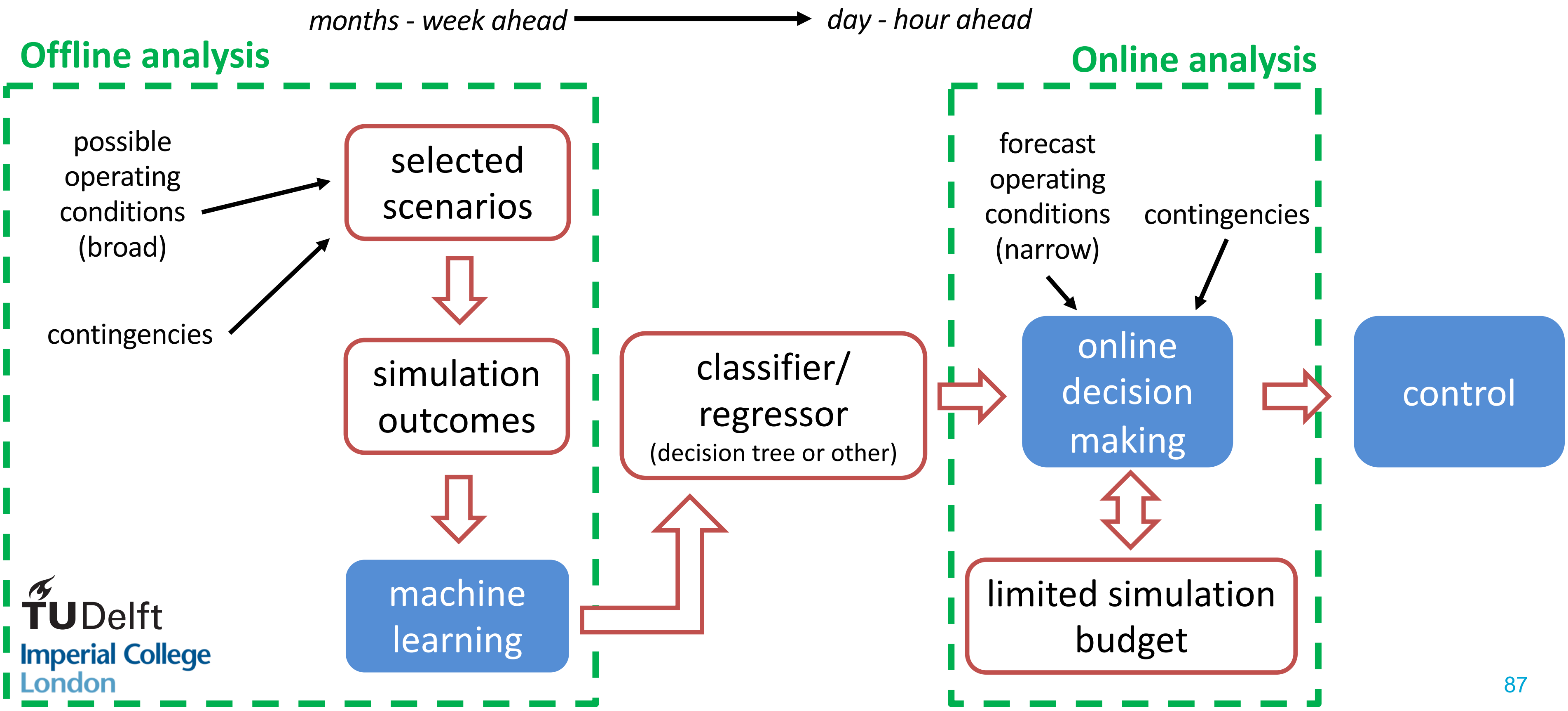
- All approaches ensure security (zero risk)
- **Cost-risk balance:** Probabilistic approach & condition-specific approach are best
- Condition-specific approach requires iterations in training
- Calibrated probabilistic approach outperforms all other approaches in the combination of time, robustness and cost-risk balance

TU Delft
Imperial College London

Summary and outlook

TU Delft
Imperial College London

Using machine learning for DSA



Major open challenges

- Optimal balance between online and offline
  - Offline sample distribution
  - Dealing with topological changes
- Combination of physical models and data-driven approaches
- Coordination with complex control actions

Thank you! We are happy to answer questions

TU Delft
Imperial College London

References

- P. Kundur *et al.*, "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," in *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387-1401, Aug. 2004.
- Wehenkel, L., (1998). *Automatic Learning Decision trees and Techniques in Power Systems*, ISBN 978-0-7923-8068-9.
- F. R. S. Sevilla and L. Vanfretti, "A small-signal stability index for power system dynamic impact assessment using time-domain simulations," *2014 IEEE PES General Meeting*
- F. R. S. Sevilla and L. Vanfretti, "Static stability indexes for classification of power system time-domain simulations," *2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Washington, DC, 2015.
- Arteaga, J-M. H., Hancharou, F., Thams, F., & Chatzivasileiadis, S. *Deep Learning for Power System Security Assessment*. In *Proceedings of IEEE Powertech 2019* IEEE.
- I. Konstantelos, G. Jamgotchian, S. Tindemans, P. Duchesne, S. Cole, C. Merckx, G. Strbac and P. Panciatici, “Implementation of a Massively Parallel Dynamic Security Assessment Platform for Large-Scale Grids”, IEEE Transactions on Smart Grid, May 2017.
- A. Venzke, D.K. Molzahn, S. Chatzivasileiadis,(2019). Efficient Creation of Datasets for Data-Driven Power System Applications, arXiv:1910.01794
- I. Konstantelos, M. Sun, S. H. Tindemans, S. Issad, P. Panciatici and G. Strbac, "Using Vine Copulas to Generate Representative System States for Machine Learning," in *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 225-235, Jan. 2019.
- C. Liu et al., "A Systematic Approach for Dynamic Security Assessment and the Corresponding Preventive Control Scheme Based on Decision Trees," in IEEE Transactions on Power Systems, vol. 29, no. 2, pp. 717-730, March 2014
- Krishnan, V., McCalley, J. D., Henry, S., & Issad, S. (2011). Efficient Database Generation for Decision Tree Based Power System Security Assessment. *IEEE Transactions on Power Systems*, *26*(4), 2319–2327.
- Settles, B. (2012). *Active Learning*. https://doi.org/10.2200/S00429ED1V01Y201207AIM018
- Jochen Cremer, Ioannis Konstantelos, Goran Strbac, “From Optimization-based Machine Learning to Interpretable Security Rules for Operation”, *IEEE Transactions on Power Systems, 2019*
- Nikos Nikolaou, *“Cost-sensitive learning with AdaBoost”,* 2017
- Jochen Cremer, Ioannis Konstantelos, Simon Tindemans, Goran Strbac, “Sample-Derived Disjunctive Rules for Secure Power System Operation,” 2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), Boise, USA 2018
- Jochen Cremer, Ioannis Konstantelos, Simon Tindemans, Goran Strbac, “Data-driven Power System Operation: Exploring the Balance between Cost and Risk”, IEEE Transactions on Power Systems, 2018

TU Delft

Imperial College London