An intrinsically probabilistic approach to analyzing stochasticity and uncertainty in fusion plasmas using information geometry

G. Verdoolaege

Department of Applied Physics, Ghent University, Ghent, Belgium

7th Asia-Pacific Conference on Plasma Physics (AAPPS-DPP)
Nagoya, Japan
November 14, 2023

GHENT UNIVERSITY

infusion

EUROfusion

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 — EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Overview

1. Motivation
2. Information geometry
3. Classification of edge-localized modes (ELMs)
4. Regression analysis for scaling laws
   - Conventional techniques
   - Geodesic least squares regression (GLS)
   - ELM energies and waiting times
   - Global confinement scaling in tokamaks
5. Conclusions

Overview

1. Motivation
2. Information geometry
3. Classification of edge-localized modes (ELMs)
4. Regression analysis for scaling laws
   - Conventional techniques
   - Geodesic least squares regression (GLS)
   - ELM energies and waiting times
   - Global confinement scaling in tokamaks
5. Conclusions

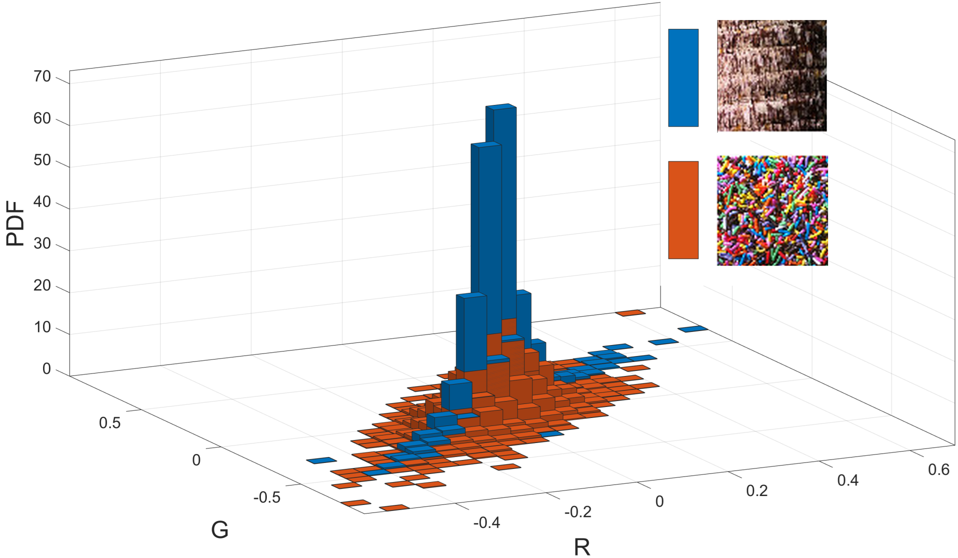
The power of simplicity

- Adoption in data-intensive communities
- Minimal parameter tuning
- Solid foundations

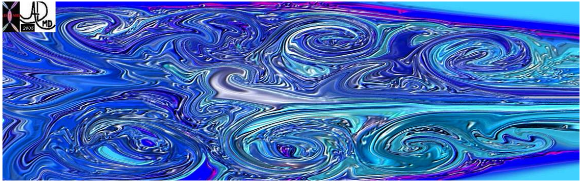
vs. 

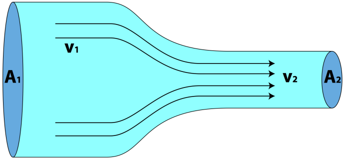
Classification of image textures

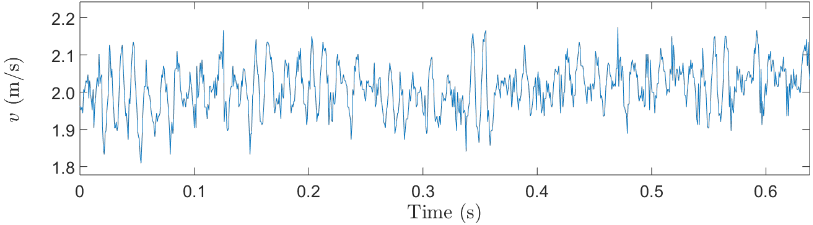
Color texture distributions



Velocity fluctuations in fluid flow







Intrinsically probabilistic data

The probabilistic nature of data

The fundamental object resulting from a measurement is a probability distribution. Any further processing of the data (statistical inference, pattern recognition) should respect this inherent probabilistic nature.

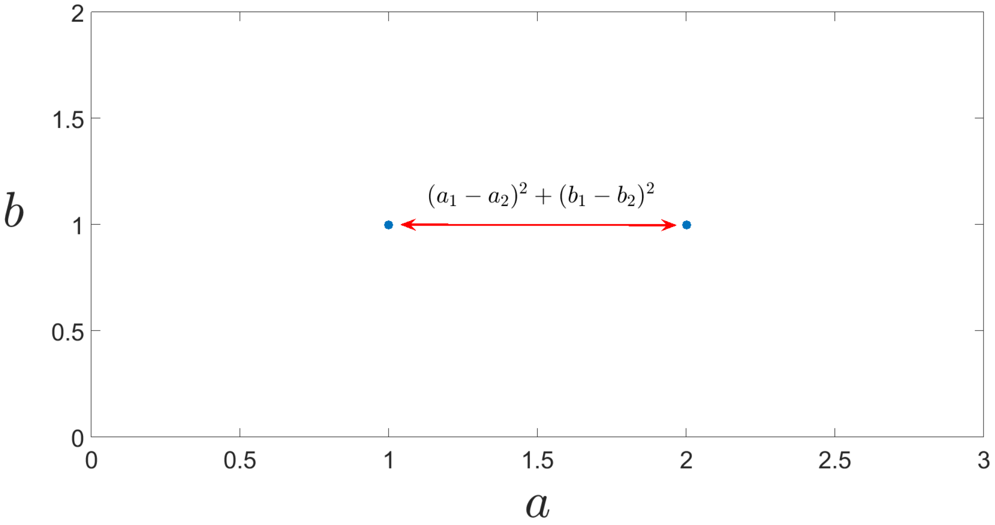
- PDF $>$ (mean, error bar)
- More flexibility, more information, but concise description
- Need to compare, classify, regress PDFs $\rightarrow$ pattern recognition in PDF spaces
- Need similarity measure or *distance* between distributions

Overview

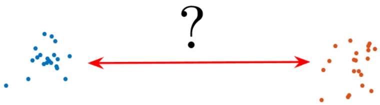
1. Motivation
2. Information geometry
3. Classification of edge-localized modes (ELMs)
4. Regression analysis for scaling laws
   - Conventional techniques
   - Geodesic least squares regression (GLS)
   - ELM energies and waiting times
   - Global confinement scaling in tokamaks
5. Conclusions

Difference/distance between measurements

Euclidean distance

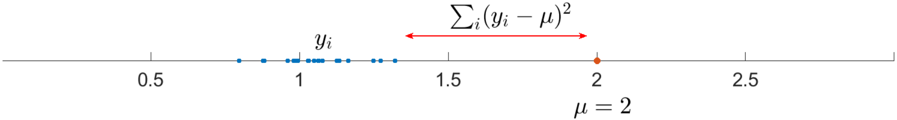


Which distance?

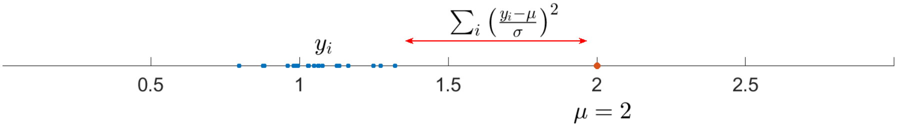


A point and a distribution

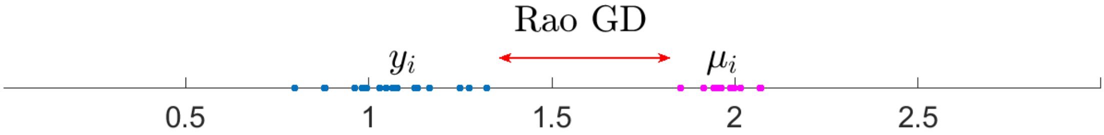
Sum of squares



Mahalanobis distance



Rao geodesic distance



Principles of information geometry

- Family of probability distributions $\rightarrow$ differentiable manifold
- Parameters $=$ coordinates
- Unique metric tensor: Fisher information matrix

$$\text{Parametric probability model: } p\left(\boldsymbol{x}|\boldsymbol{\theta}\right) \Longrightarrow$$

$$g_{\mu\nu}\left(\boldsymbol{\theta}\right) = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^\mu\partial\theta^\nu}\ln p\left(\boldsymbol{x}|\boldsymbol{\theta}\right)\right], \quad \mu,\nu = 1,\ldots,m$$

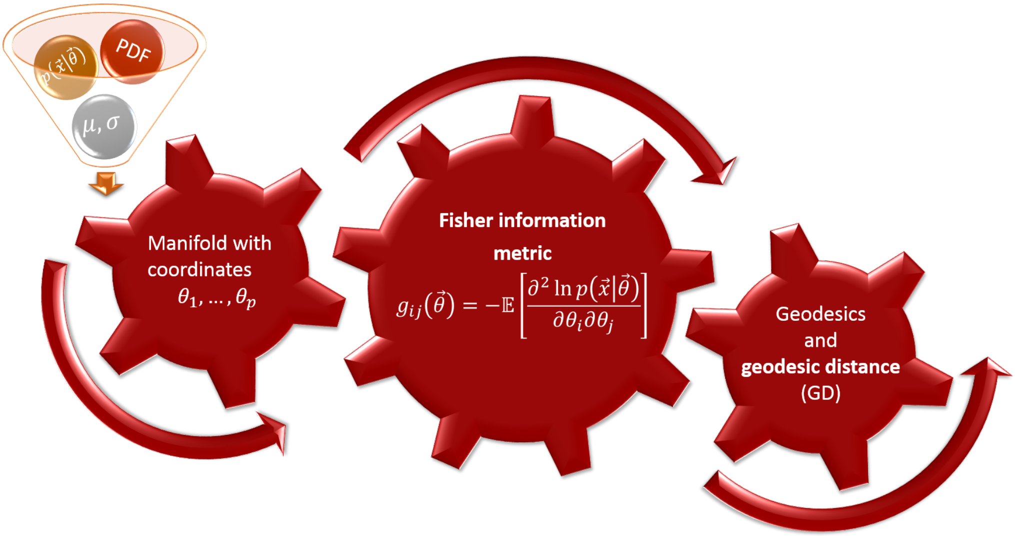
$\boldsymbol{\theta}$ = m-dimensional parameter vector

- Parallel transport via Levi-Civitta connection
- Line element:

$$ds^2 = g_{\mu\nu}d\theta^\mu d\theta^\nu$$

- Minimum-length curve: *geodesic*
- *Rao geodesic distance* (GD)

Information geometry scheme



The Gaussian manifold

- Probability density function (PDF):

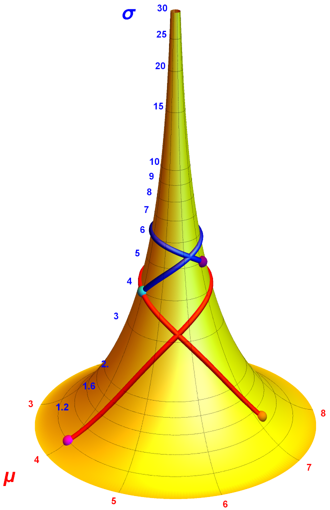
$$p(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

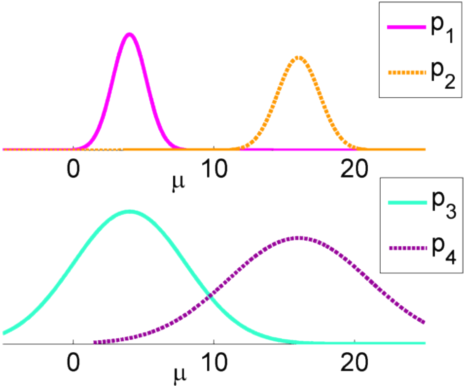
- Line element:

$$\mathrm{d}s^2 = \frac{\mathrm{d}\mu^2}{\sigma^2} + 2\frac{\mathrm{d}\sigma^2}{\sigma^2}$$

- Hyperbolic geometry: Poincaré half-plane model, pseudosphere, …
- Analytic geodesic distance

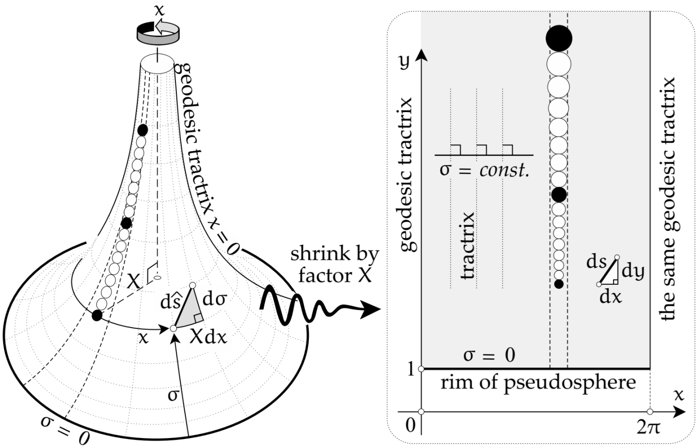
Geodesic intuition: the pseudosphere

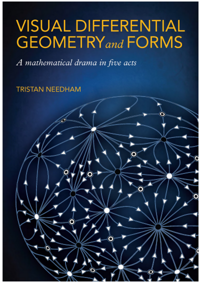




Poincaré half-plane

Conformal mapping pseudosphere $\leftrightarrow$ ***Poincaré half-plane***





T. Needham, Visual differential geometry and forms, PUP, 2021

Overview

1. Motivation

2. Information geometry

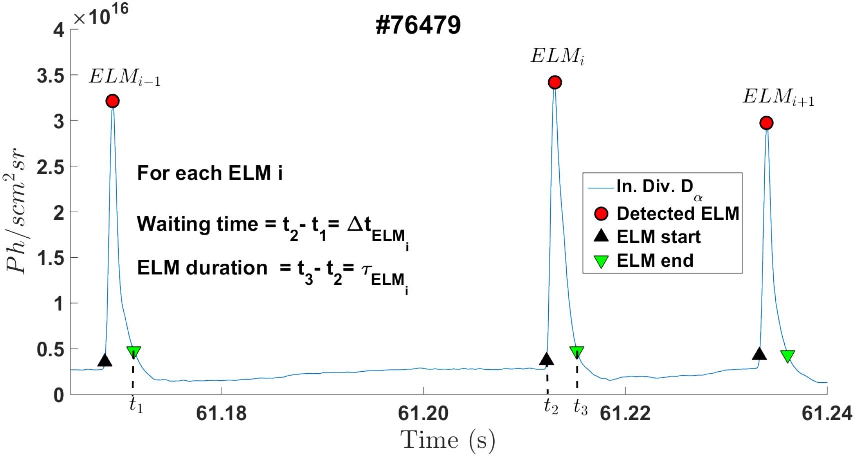
3. Classification of edge-localized modes (ELMs)

4. Regression analysis for scaling laws
   - Conventional techniques
   - Geodesic least squares regression (GLS)
   - ELM energies and waiting times
   - Global confinement scaling in tokamaks

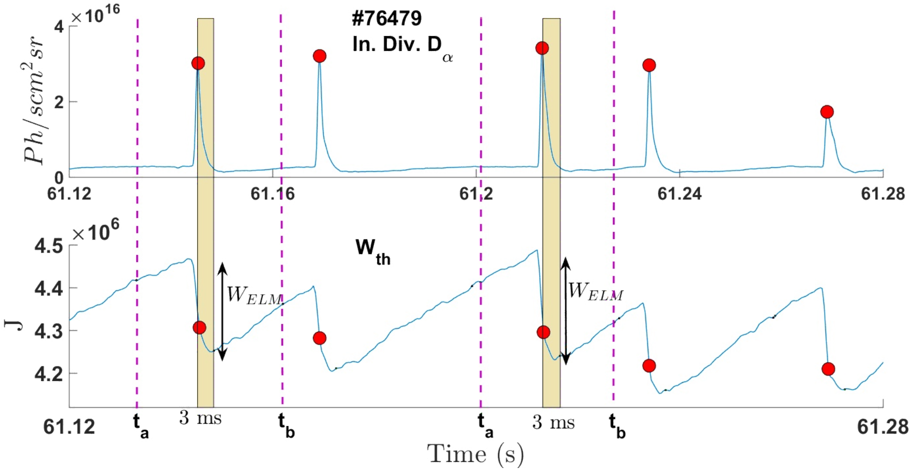
5. Conclusions

ELM waiting times

JET data



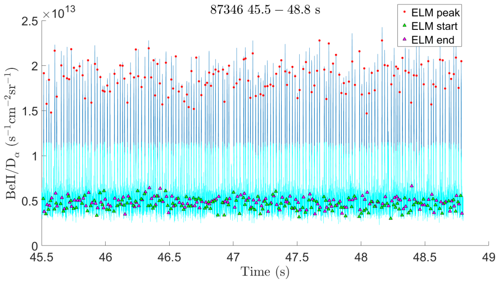
ELM energy losses



ELM types

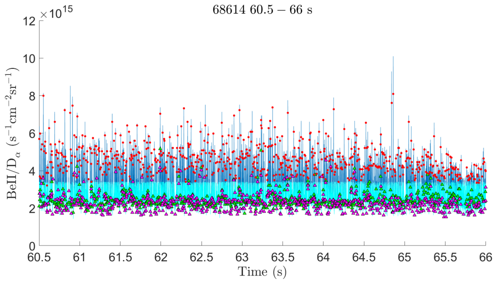
Type I:

- Large and slow ($\lesssim$ 150 Hz)
- Frequency increases with heating power



Type III:

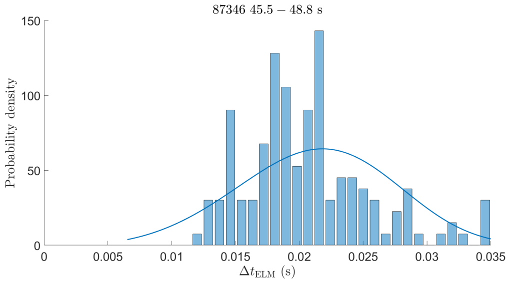
- Small and fast ($\gtrsim$ 100 Hz)
- Frequency decreases with heating power



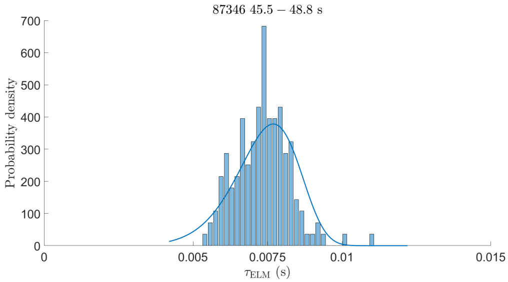
G. Verdoolaege *et al.*, 38th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (MaxEnt2018), London, UK, 2018

ELM timing distributions

Inter-ELM time Δt_{ELM}



ELM duration τ_{ELM}



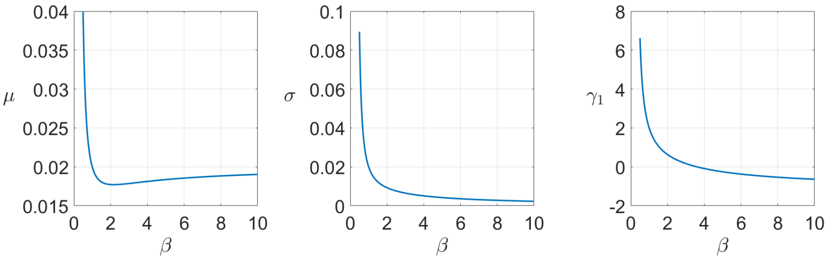
Weibull distribution

- Minimum recovery time t_m
- Prob(ELM) $\sim \left(\frac{t-t_\mathrm{m}}{\alpha}\right)^{\beta-1}$
- Waiting time distribution:

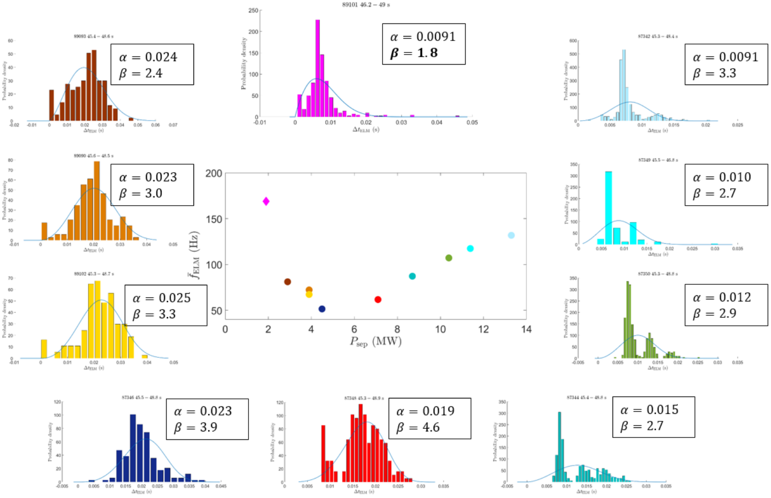
$$p(t) = \frac{\beta}{\alpha}\left(\frac{t-t_\mathrm{m}}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{t-t_\mathrm{m}}{\alpha}\right)^{\beta}\right], \qquad t > t_\mathrm{m}$$

A.J. Webster *et al.*, Phys. Rev. Lett. **110**, 155004, 2013

- Here: $t_\mathrm{m} = 0$

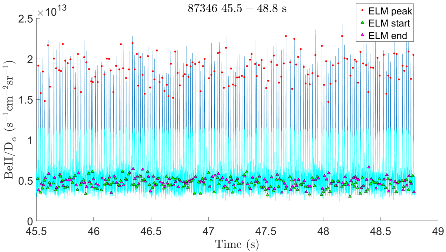


ELM types in power scan

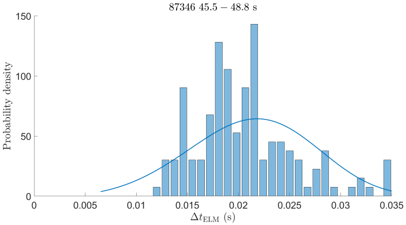


G. Verdoolaege *et al.*, MaxEnt 2018

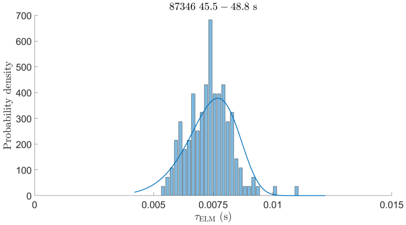
Type I ELM distributions



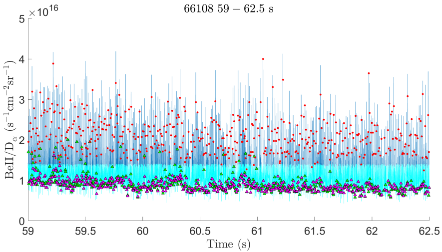
Δt_{ELM}



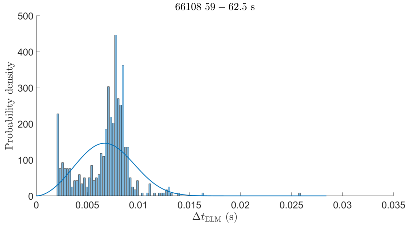
τ_{ELM}



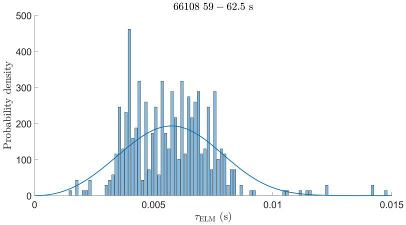
Type I high-frequency ELM distributions



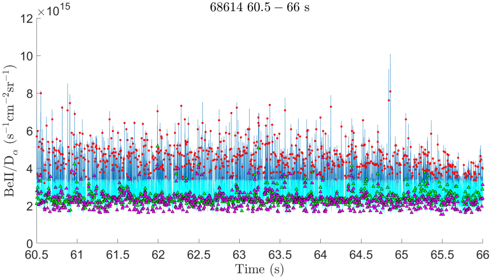
Δt_{ELM}



τ_{ELM}

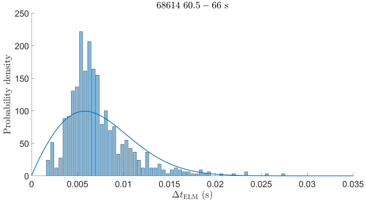


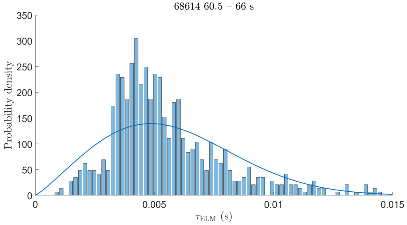
Type III ELM distributions



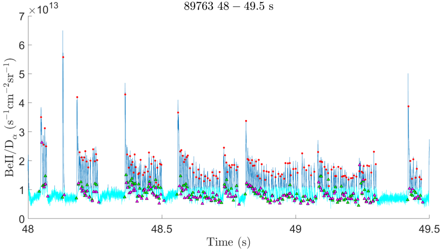
Δt_{ELM}

τ_{ELM}

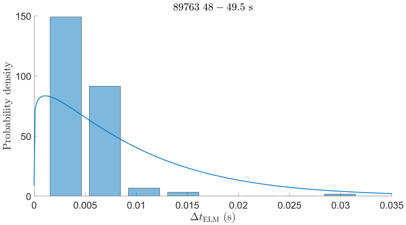




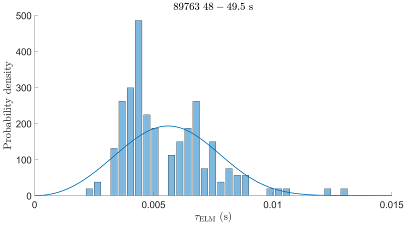
Compound ELM distributions



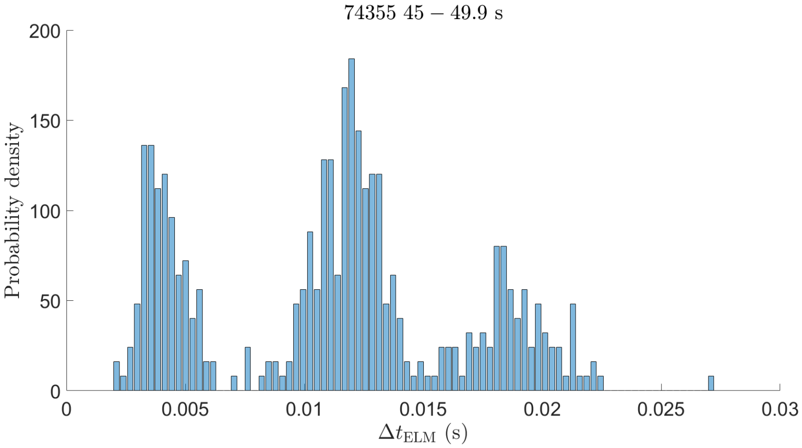
$\Delta t_{\rm ELM}$



$\tau_{\rm ELM}$



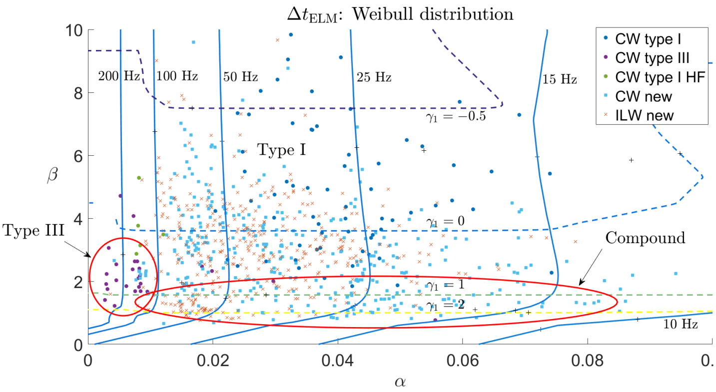
Multimodal distributions



C. Bowman *et al.*, Nucl. Fusion **58**, 016021, 2018

Distribution map Δt_{ELM}

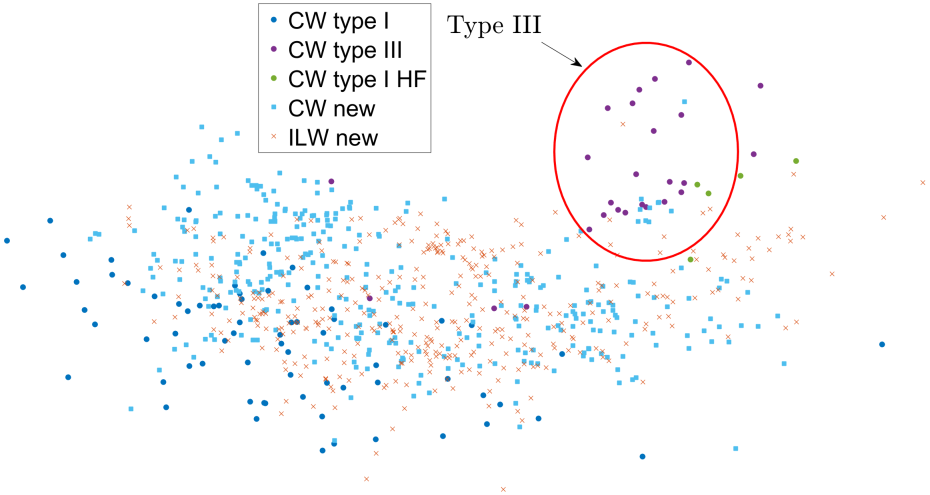
- 453 plasmas from JET carbon wall
- 379 plasmas from JET ITER-like wall
- 100 reference plasmas (A. Shabbir *et al.*, Rev. Sci. Instrum. **87**, 11D404, 2016)



G. Verdoolaege et al., 45th EPS Conference on Plasma Physics, ECA **42A**, P.1078, 2018

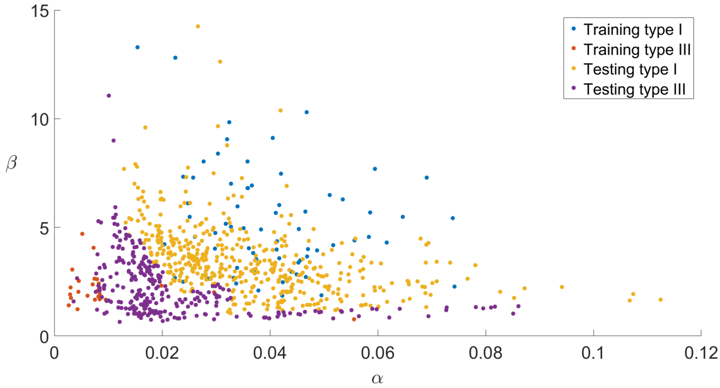
Projected map Δt_{ELM}

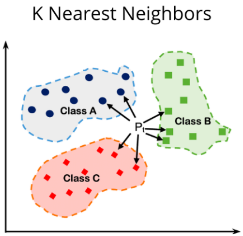
Approximately isometric projection with multidimensional scaling



ELM classification with GD

Based on Δt_{ELM} distributions, nearest neighbor



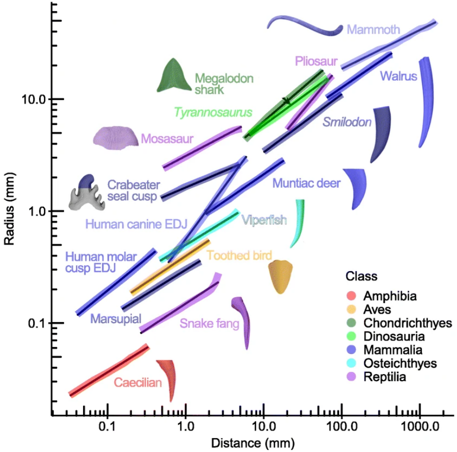


A. Shabbir *et al.*, Rev. Sci. Instrum. **87**, 11D404, 2016

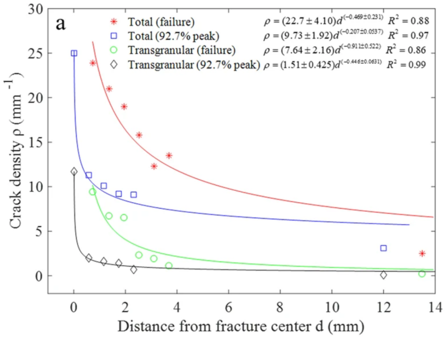
Overview

Overview

Ubiquity of power law models



A.R. Evans *et al.*, BMC Biology, **19**, 58, 2021



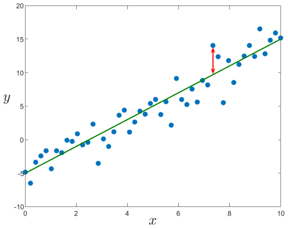
F. Meng *et al.*, Scientific Reports, **9**, 10705, 2019

Least squares and maximum a posteriori estimation

- Multilinear regression model on logarithmic scale:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_p x_p + \epsilon$$
$$\epsilon \sim \mathcal{N}(0, \sigma^2),\ \sigma \text{ known}$$

- Parameter estimation $\rightarrow$ distance minimization: expected $\leftrightarrow$ measured

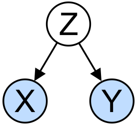


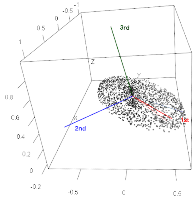
- Workhorse: ordinary least squares (OLS)

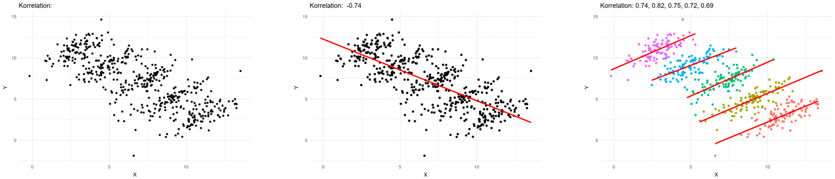
- Maximum likelihood (ML) / maximum *a posteriori* (MAP)

$$\frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{1}{2} \frac{\left[y - f(x, \theta) \right]^2}{\sigma^2} \right\}$$

- Measurement uncertainty
- Model uncertainty:
  - Linear, power law, . . . ?
  - Missing variables
  - Confounding variables
- Multicollinearity: e.g. $I_p \propto B_t$
- Heterogeneity: multi-machine database
- Simpson's paradox:







Robust Bayesian regression

- Errors in all variables (loglinear):

$$y = \eta + \epsilon_y, \quad x_1 = \xi_1 + \epsilon_{x_1}, \quad \ldots \qquad \eta = \alpha_0 + \sum_{j=1}^{p} \alpha_j \xi_j$$

$$\epsilon_y \sim \mathcal{N}\left(0, \sigma_y^2\right), \quad \epsilon_{x_1} \sim \mathcal{N}\left(0, \sigma_{x_1}^2\right), \quad \ldots \qquad \sigma_{\text{mod}}^2 = \sigma_y^2 + \sum_{j=1}^{p} \alpha_j^2 \sigma_{x_j}^2$$

- Robust likelihood:

$$p\left(\{y_{i_k,k}\}, \{x_{i_k,j,k}\} \middle| \{\alpha_0, \alpha_j\}, \{\gamma_k\}\right) = \prod_k \prod_{i_k} \frac{1}{\sqrt{2\pi\gamma_k^2 \sigma_{\text{mod},i_k,k}^2}} \exp\left[-\frac{1}{2} \frac{\left(y_{i_k,k} - \eta_{i_k,k}\right)^2}{\gamma_k{}^2 \sigma_{\text{mod},i_k,k}^2}\right]$$

1 for each device (→ γ_k)

Overview

The minimum distance approach

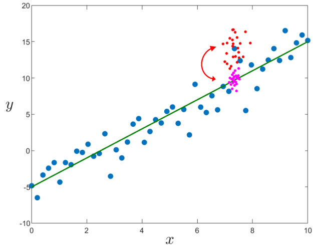
- ***Minimum distance estimation*** (Wolfowitz, 1952):

  Which distribution does the model predict?

  vs.

  Which distribution do you observe?

- Gaussian case: different means *and* standard deviations
- Kullback-Leibler divergence, Hellinger divergence (Beran, 1977), ...
- Observed distribution: kernel density estimate



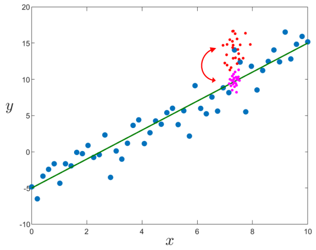
Geodesic least squares regression (GLS)

- Geodesic least squares: ***GLS***

$$\prod_k \prod_{i_k} \frac{1}{\sqrt{2\pi\sigma^2_{\text{tot},i_k,k}}} \exp\left[-\frac{1}{2}\frac{\left(y_{i_k,k} - \eta_{i_k,k}\right)^2}{\sigma^2_{\text{mod},i_k,k}}\right]$$

↕ Rao geodesic distance (GD)

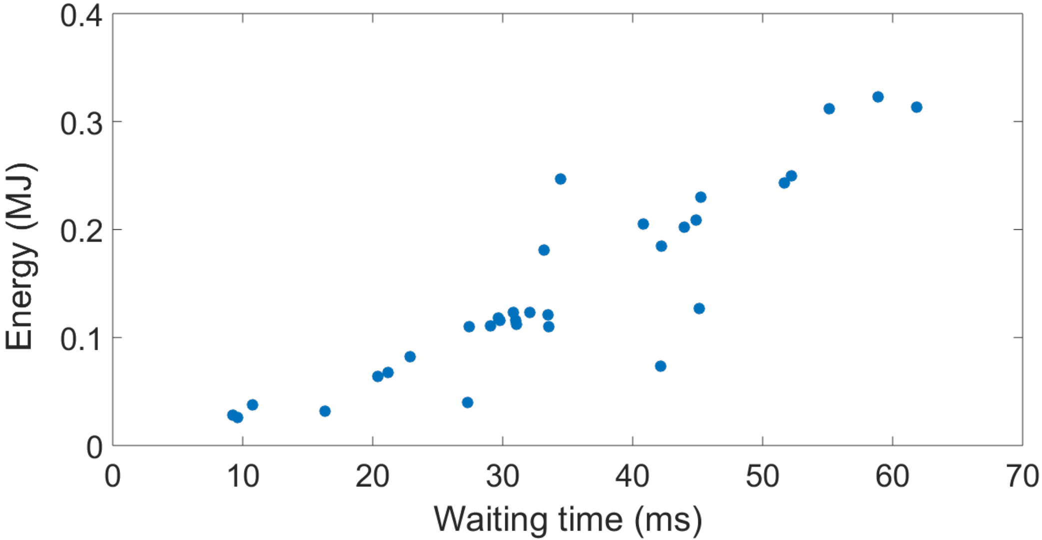
$$\frac{1}{\sqrt{2\pi}\,\sigma_{\text{obs}}} \exp\left[-\frac{1}{2}\frac{(y - y_i)^2}{\sigma_{\text{obs}}{}^2}\right]$$



G. Verdoolaege *et al.*, Nucl. Fusion, **55**, 113019, 2015
G. Verdoolaege *et al.*, Entropy, **17**, 4602–4626, 2015

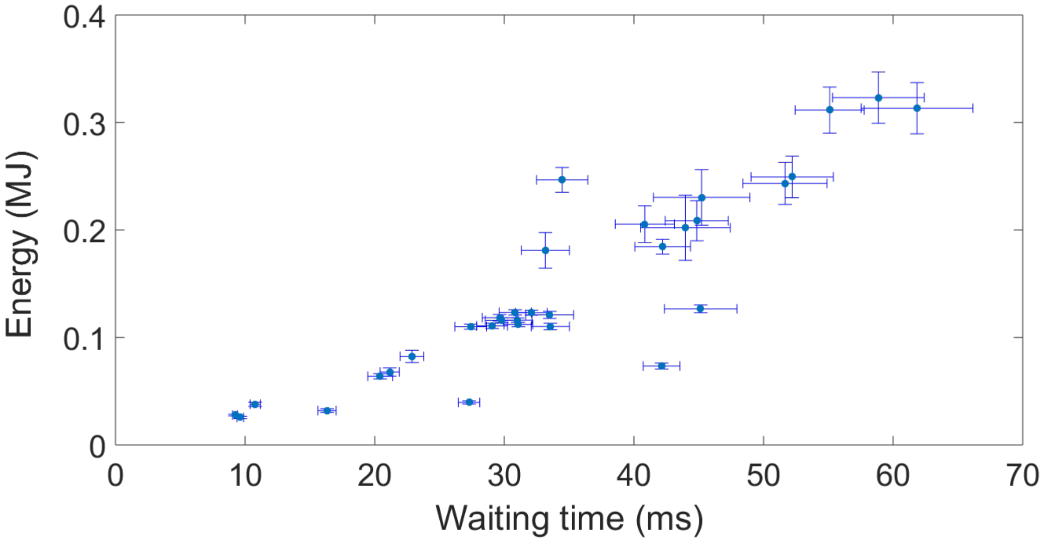
Overview

Average waiting times and energies



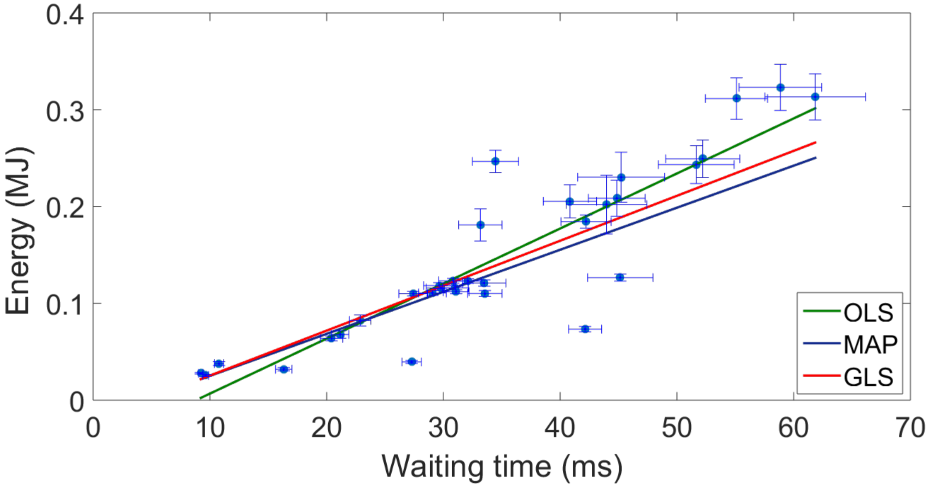
Error bars on averages

- Standard deviation / $\sqrt{n} \rightarrow$ error bars

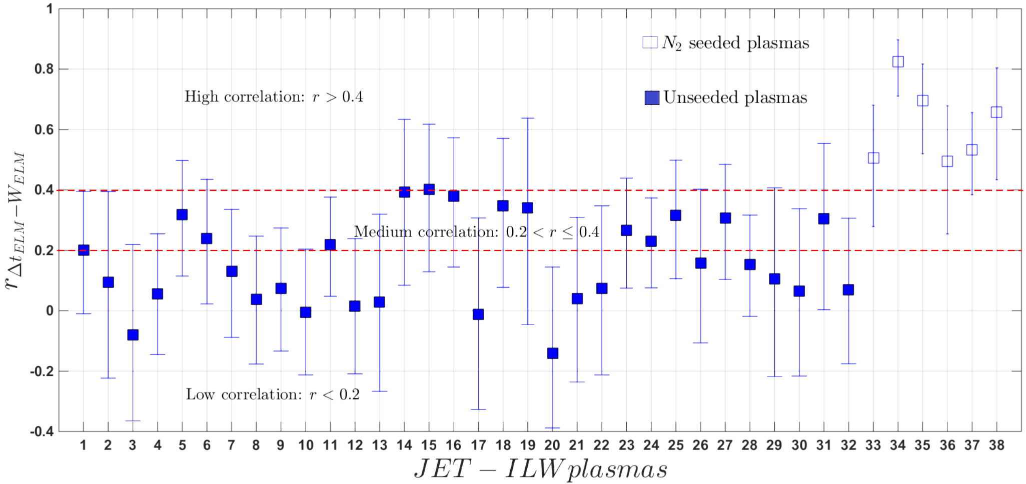


Regression on averages

$$E_{\text{ELM}} = \alpha_0 + \alpha_1 \Delta t_{\text{ELM}}, \qquad \sigma_{E,\text{obs}} \propto \mu_{E,\text{obs}}$$

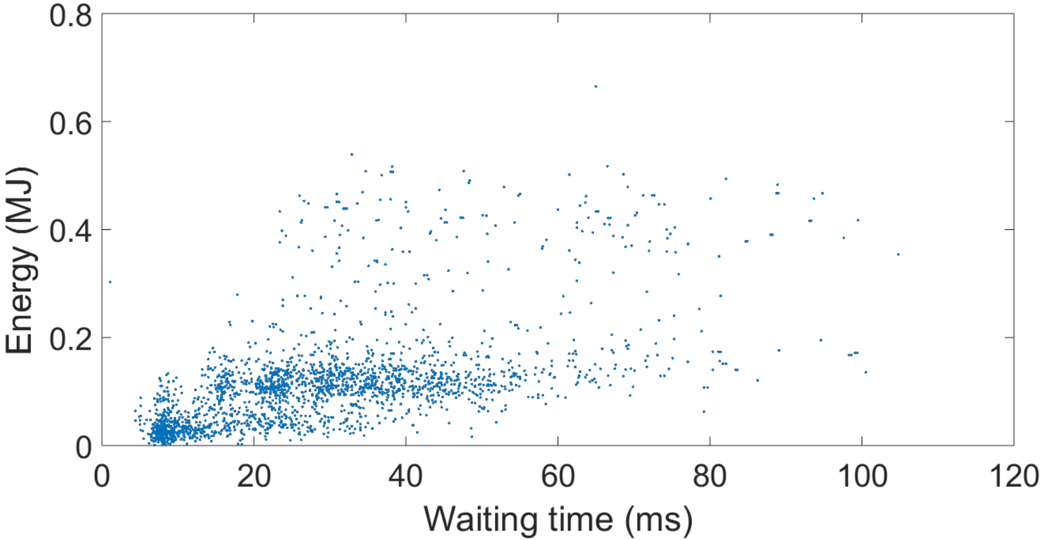


Correlations in individual plasmas



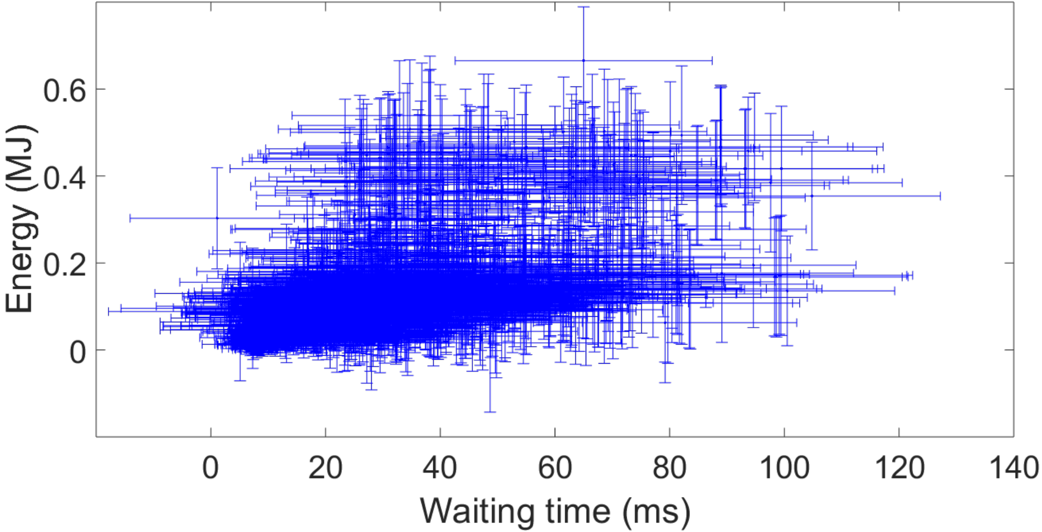
A. Shabbir *et al.*, Nucl. Fusion **57**, 036026, 2017

Individual waiting times and energies



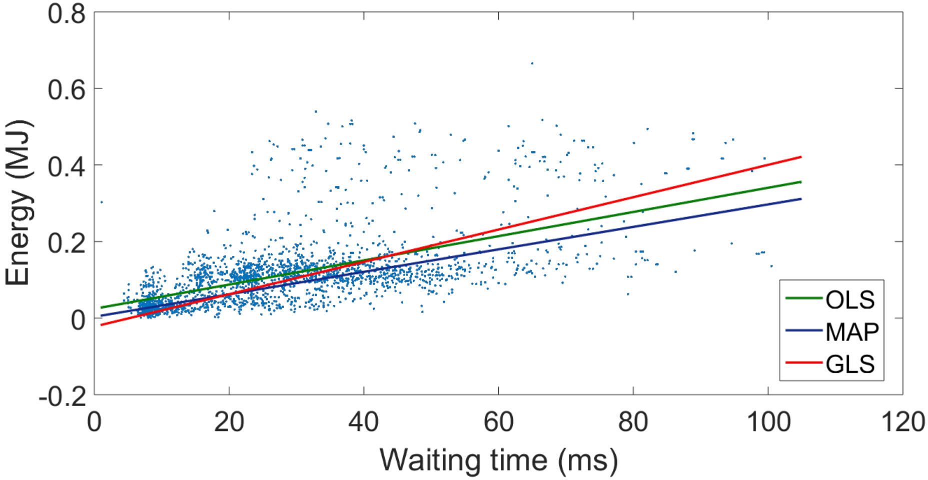
Error bars for individual ELMs

- Standard deviation → error bars



Regression on individual measurements

$$E_{\mathrm{ELM}} = \alpha_0 + \alpha_1 \Delta t_{\mathrm{ELM}}$$



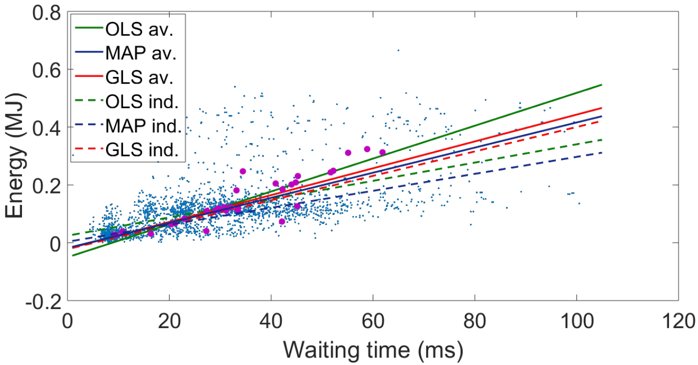
Average vs. collective trend

Average

Method	α_0 (MJ)	α_1 (MJ/s)
OLS	-0.050	5.7
GLS	-0.021	4.6

Individual

Method	α_0 (MJ)	α_1 (MJ/s)
OLS	0.024	3.2
GLS	-0.022	4.2



Overview

Global confinement database update

- Motivation and use of (confinement) scaling laws:
  - Based purely on experimental data (or almost)
  - Benchmark for experimental performance
  - Experimental design and modeling
- Multi-machine ***Global H-mode Confinement Database*** (*1989 $\rightarrow$ ITPA 2001)
- IPB98(y,2) ELMy H-mode global energy confinement scaling (1998)
- ITPA TC-28: Revision of ITER confinement database (2015–2020):
  - Add data closer to ITER conditions and expand parameter ranges
  - Add data from devices with fully metallic walls
  - Reconcile with single-machine scans ($\bar{n}_\mathrm{e}$, $P_\mathrm{l,th}$, β, $\ldots$)
  - Explore new predictor variables (e.g. δ, $n_\mathrm{e,sep}$, torque, $\ldots$)
  - Robust regression analysis

Issues with IPB98

Density scaling — Power degradation — No δ scaling

$$\tau_{\mathrm{E,th}} = 0.0562\, I_{\mathrm{p}}^{0.93}\, B_{\mathrm{t}}^{0.15}\, \bar{n}_{\mathrm{e}}^{0.41}\, P_{\mathrm{l,th}}^{-0.69}\, R_{\mathrm{geo}}^{1.97}\, (1+\delta)^{\alpha_\delta}\, \kappa_{\mathrm{a}}^{0.78}\, \epsilon^{0.58}\, M_{\mathrm{eff}}^{0.19}$$

$$\Omega_{\mathrm{i}}\tau_{\mathrm{E,th}} = 4.24 \times 10^{-7}\, \rho_*^{-2.69}\, \beta_{\mathrm{t}}^{-0.90}\, \nu_*^{-0.0081}\, q_{\mathrm{cyl}}^{-2.99}\, (1+\delta)^{\alpha_\delta}\, \kappa_{\mathrm{a}}^{3.29}\, \epsilon^{0.71}\, M_{\mathrm{eff}}^{0.96}$$

β degradation — No collisionality scaling — No δ scaling

Updates since DB4

- DB2.8 (IPB98) → DB3v13F → DB4.5 → ***DB5.2.3***
- Enhanced data validation and W_{fast} in ASDEX Upgrade (AUG)
- New data:

Metallic wall ***High-Z***

- AUG full W (AUG-W): 825 new points [1]
- JET ITER-like wall (ILW): 866 new points [2]

Carbon-based wall ***Low-Z***

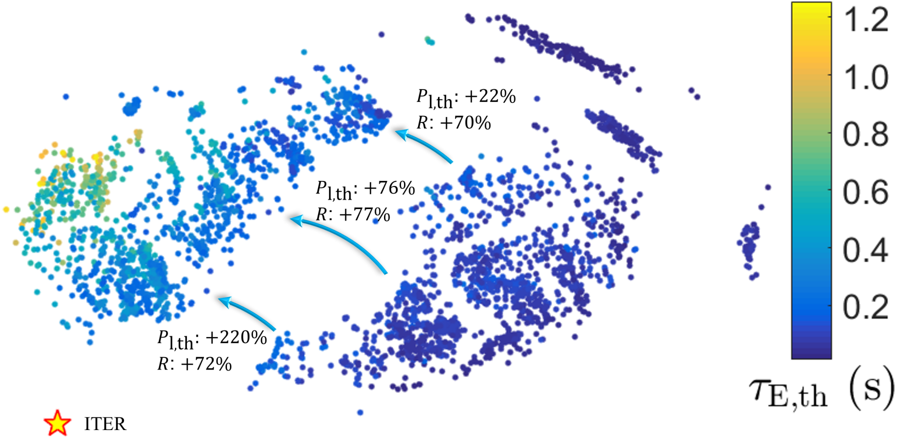
- AUG + JET: new data with high gas injection

[1] F. Ryter *et al.*, Nucl. Fusion, **61**, 046030, 2021

[2] M. Maslov *et al.*, Nucl. Fusion, **60**, 036007, 2020

Database visualization by projection

- Multidimensional scaling
- Distance measure: Rao GD between Gaussian PDFs

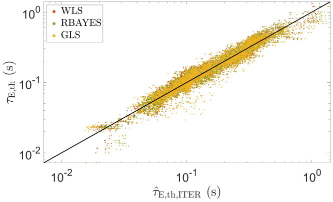


Multi-machine engineering scaling

STD5-IL ELMy H-mode (error bars from Bayesian analysis)

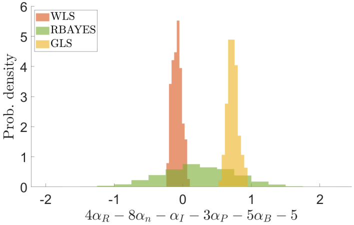
Engineering scaling **ITPA20-IL**

$$\tau_{\mathrm{E,th}} = (0.067 \pm 0.060)\, I_{\mathrm{p}}^{1.29\pm0.17}\, B_{\mathrm{t}}^{-0.13\pm0.17}\, \bar{n}_{\mathrm{e}}^{0.15\pm0.10}\, P_{\mathrm{l,th}}^{-0.644\pm0.060}\, R_{\mathrm{geo}}^{1.19\pm0.29} \times (1+\delta)^{0.56\pm0.35}\, \kappa_{\mathrm{a}}^{0.67\pm0.65}\, M_{\mathrm{eff}}^{0.30\pm0.17}$$

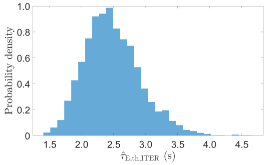


$\mathrm{RMSE} = 0.17$

$R^2 = 0.95$



Quasi-neutral high-β Fokker-Planck with Landau collision operator + Ampére

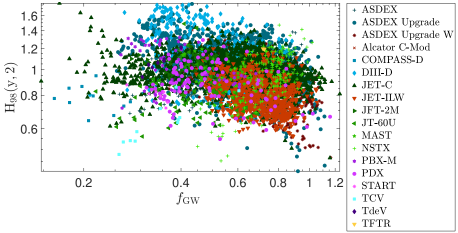


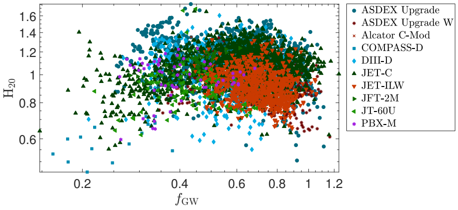
$\hat{\tau}_{\mathrm{E,th,ITER}} = 2.90 \pm 0.46\,\mathrm{s}$

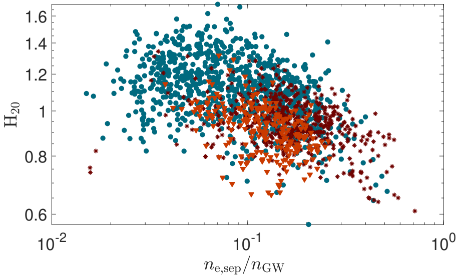
G. Verdoolaege *et al.*, Nucl. Fusion, **61**, 076006, 2021

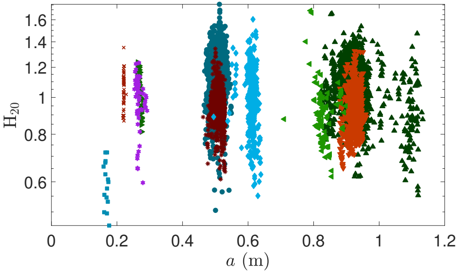
Residuals

$\mathrm{H}_{20} \equiv \tau_{\mathrm{E,th}} / \hat{\tau}_{\mathrm{E,th}}$









Multi-machine dimensionless scaling

- Transformation introduces large errors:

$$\boldsymbol{\alpha}_{\mathrm{D}} = A^{-1}\left(\frac{\boldsymbol{\alpha}}{1+\alpha_P} - \boldsymbol{a}\right)$$

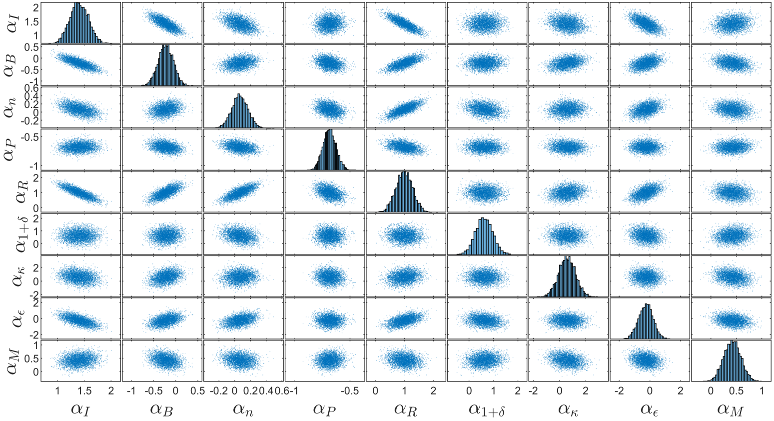
- Regression in dimensionless space

Dimensionless scaling **ITPA20-IL-dim**

$$\Omega_{\mathrm{i}}\tau_{\mathrm{E,th}} = (2.0 \pm 6.2 \times 10^{-6})\, \rho_*^{-1.97\pm0.33}\, \beta_{\mathrm{t}}^{0.12\pm0.21}\, \nu_*^{-0.425\pm0.062}\, q_{\mathrm{cyl}}^{-1.03\pm0.42} \times (1+\delta)^{0.14\pm0.63}\, \kappa_{\mathrm{a}}^{1.90\pm0.81}\, \epsilon^{-0.26\pm0.62}\, M_{\mathrm{eff}}^{0.61\pm0.37}$$

Error analysis

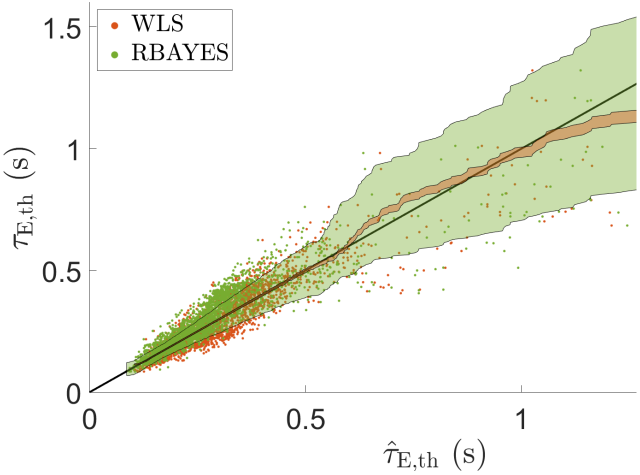
- Sensitivity analysis under multicollinearity: $\sim 10\times$ larger error bars
- Observed uncertainty $\gg$ modeled uncertainty



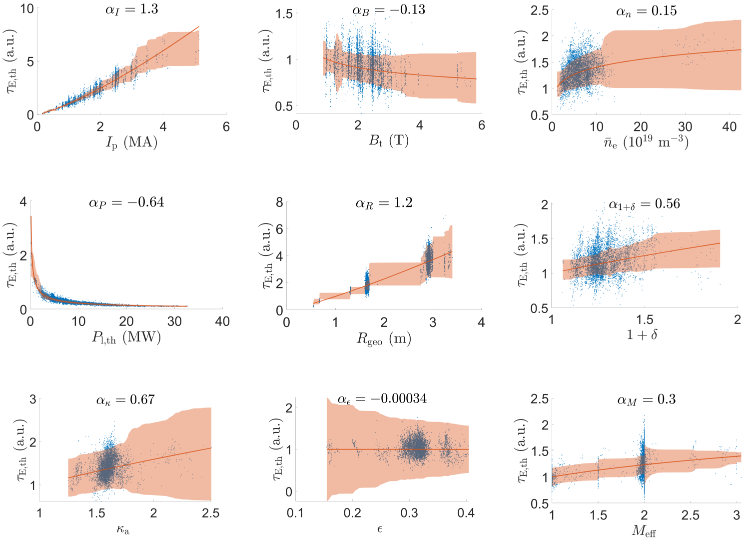
MCMC sampler with 4 chains

Uncertainty calibration

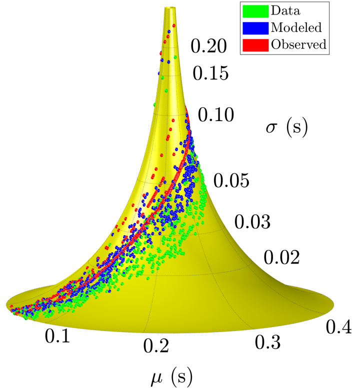
- Predictions for JET
- Confidence bands (smoothed):

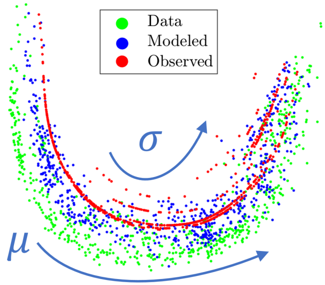


Confinement trends



Visualization on pseudosphere and projection



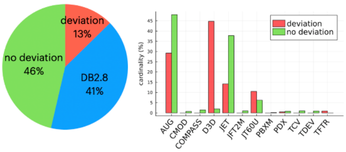


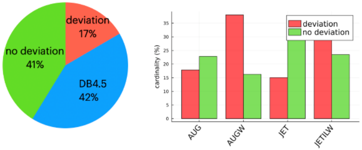
	AUG	AUG-W	Alcator C-Mod	COMPASS-D	DIII-D	JET-C	JET-ILW	JFT-2M	JT-60U	PBX-M
γ_k	0.96	0.79	0.51	3.8	0.85	1.0	0.89	0.69	1.1	0.85
$\sigma_{\text{obs},k}/\sigma_{\text{mod},k}$	1.2	1.1	1.2	2.0	1.2	1.2	1.2	1.1	1.2	2.0

Origin of reduced size scaling

- Clustering 'deviation' vs. 'no-deviation' set:

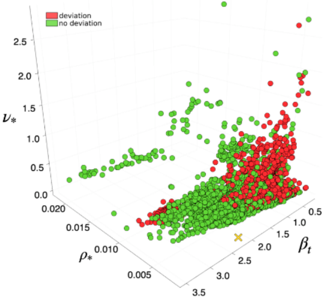
DB2.8 → DB4.5 → DB5.2.3





- Scaling for 'no deviation' set:

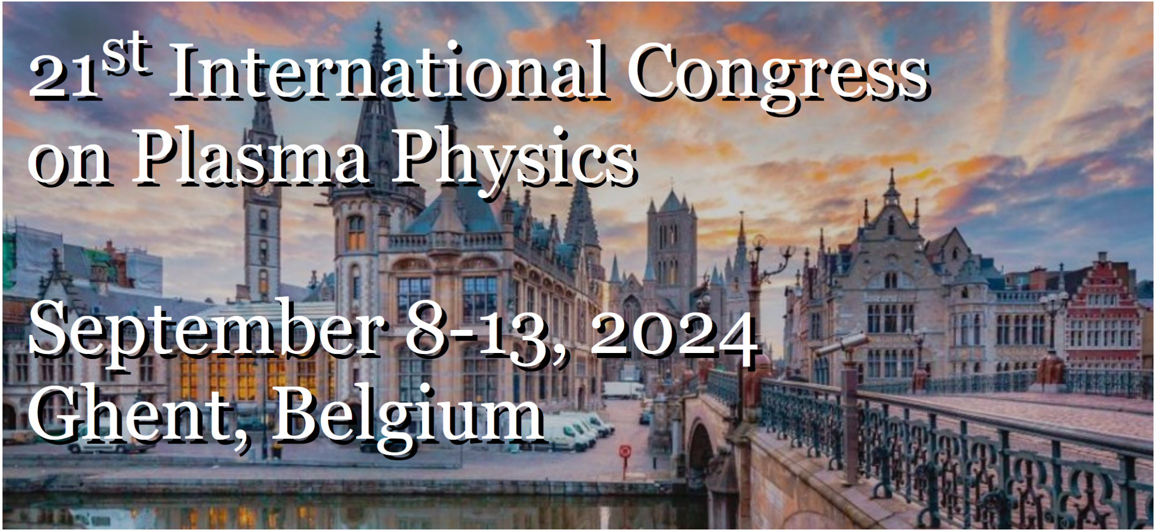
$$\tau_{\mathrm{E,th}} = 0.07\, I_{\mathrm{p}}^{0.87}\, B_{\mathrm{t}}^{0.20}\, \bar{n}_{\mathrm{e}}^{0.38}\, P_{\mathrm{l,th}}^{-0.74}\, R_{\mathrm{geo}}^{2.0} \times \kappa_{\mathrm{a}}^{0.41}\, \epsilon^{0.59}\, M_{\mathrm{eff}}^{0.24}$$



Overview

Conclusions

- Probability distributions maximally descriptive for stochastic processes
- Information geometry $\rightarrow$ distance between PDFs
- Classification and regression (GLS) on probabilistic manifolds:
    - Simple, fast and robust algorithms
    - Visual interpretation
- Applications:
    - Classification and regression of ELM distributions
    - Revision of global confinement scaling with realistic error estimates
- Wider applicability to parameter estimation, model validation, etc.



https://icpp2024.ugent.be

Number of database entries

	DB2.8	DB5.2.3			
		STD5		STD5 ITER-like	
Device	ELMy H	All H	ELMy H	ELMy H	ELM-free H
ASDEX	431	575	431	0	0
AUG	102	1 385	1 377	1 370	8
AUG-W	0	767	767	767	0
Alcator C-Mod	37	82	45	45	37
COMPASS-D	0	21	16	16	5
DIII-D	270	502	388	383	114
JET-C	246	2 211	1 762	1 606	426
JET-ILW	0	866	866	855	0
JFT-2M	59	348	69	59	197
JT-60U	9	100	100	100	0
MAST	0	43	43	0	0
NSTX	0	230	185	0	0
PBX-M	59	214	59	59	155
PDX	97	119	97	0	0
START	0	8	8	0	0
T-10	0	4	0	0	0
TCV	0	17	11	0	0
TdeV	0	7	7	0	0
TEXTOR	0	0	0	0	0
TFTR	0	2	2	0	0
TUMAN-3M	0	36	0	0	0
Total	**1 310**	**7 537**	**6 233**	**5 260**	**942**

Data selection

- 14 153 points from 19 tokamaks
- 'Standard set' ***STD5*** $\rightarrow$ 7 537 points from 18 devices:
  - Quasi steady-state H-modes, no pellets, no ITBs
  - Limited P_{rad}, W_{fast}, l_{i}; minimum q_{95}
  - Limited $T_{\text{e}} \neq T_{\text{i}}$
- 'ITER-like' ***STD5-IL*** $\rightarrow$ 6 202 points from 8 devices:
  - $q_{95} > 2.8$
  - $1.3 < \kappa < 2.2$
  - $\epsilon = a/R_{\text{geo}} < 0.5$
  - $Z_{\text{eff}} < 5$
- High-Z: Alcator C-Mod, AUG-W, JET-ILW

D. McDonald *et al.*, Nucl. Fusion, **47**, 147–174, 2007