Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

EURAMET

Reference Data for Trustworthy Machine Learning/ AI

Markus Bär

Mathematical Modelling and Data Analysis, PTB

EMN for Mathematics and Statistics, EURAMET

MATHMET

Metrology for Digital Transformation – Online

September, 23rd, 2021

Content

- Motivation: Reference data for AI/machine learning

- Example: Metrology of automated ECG analysis – EMPIR Project 18HLT07 (2019 – 2022)

Background:

Future research in mathematics and statistics for metrology

Machine learning and artificial intelligence:

- Application in sensor networks, advanced manufacturing, medical physics (e. g. PTB AI for Health) <--> large data sets

Virtual metrology and digital twins:

- Competence VirtMet@PTB

- Intl. Workshop of VirtMet, EMNs MATHMET & AdvManu (Sept. 21-22, 2021, 120 participants from > 20 countries)

- Simulation of realistic physics-based models
- Proper treatment of uncertainties, uncertainty quantification
- Validation by comparison to measurements

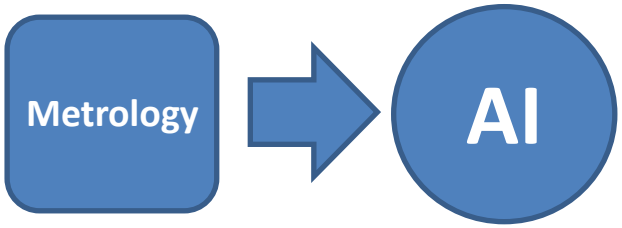
AI Survey among 14 European NMIs

Research Priorities ?

- **Reliability** (19)

  Robustness, Repeatability, **Training data quality**

Metrology Support ?



- Uncertainty, traceability (12)

- **Reference data bases** (8)
  - Measurements, simulation data (“digital twins”), quality ?

- Standardization concepts (6)

Metrology for ECG data analysis

Objectives (EMPIR Project MedalCare, 2019 – 2022)

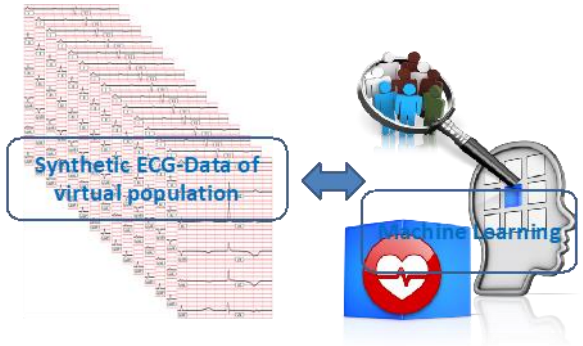
- **Metrology for advanced data analysis:**
  Benchmark data analysis & AI methods

- **Reference data for AI :**

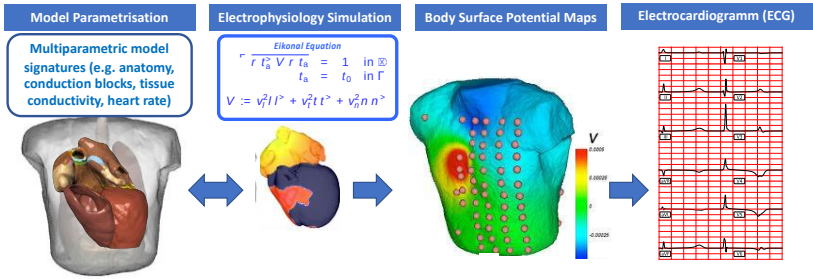
  i. Clinical ECG data base (here PTB-XL) with diagnosis

  ii. Synthetic data base with „ground truth“

Uncertainty Analysis



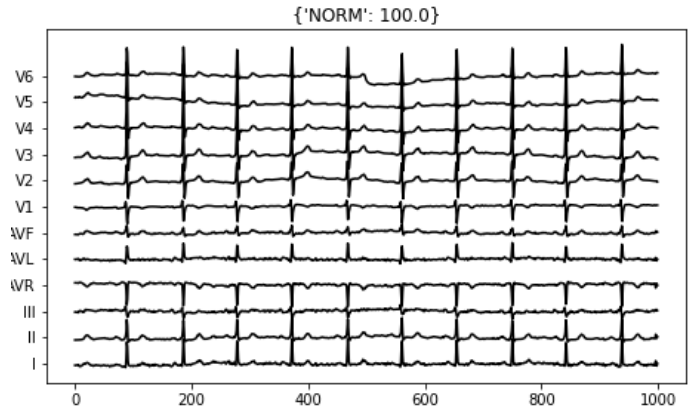
Computational Model



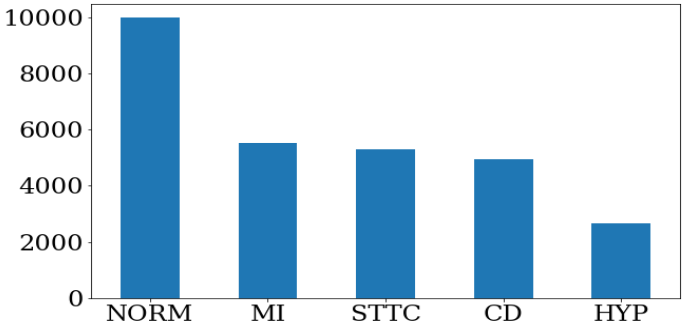
Measurements

PTB – XL clinical reference data base

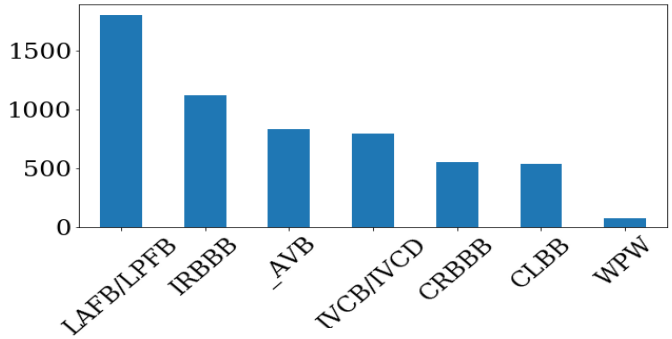
- 22.000 ECG-recording (10s)
- 12-lead ECG measurements
- Diagnostic (62) and rhythm (24) statements according SCP ISO-standard



Superclasses



Conduction disturbance (CD)



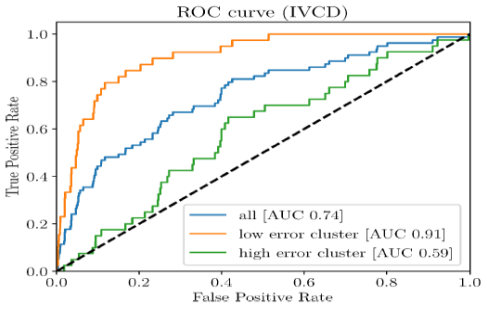
P. Wagner, N. Strodthoff, et al. Scientific Data (2020)

Benchmarking methods

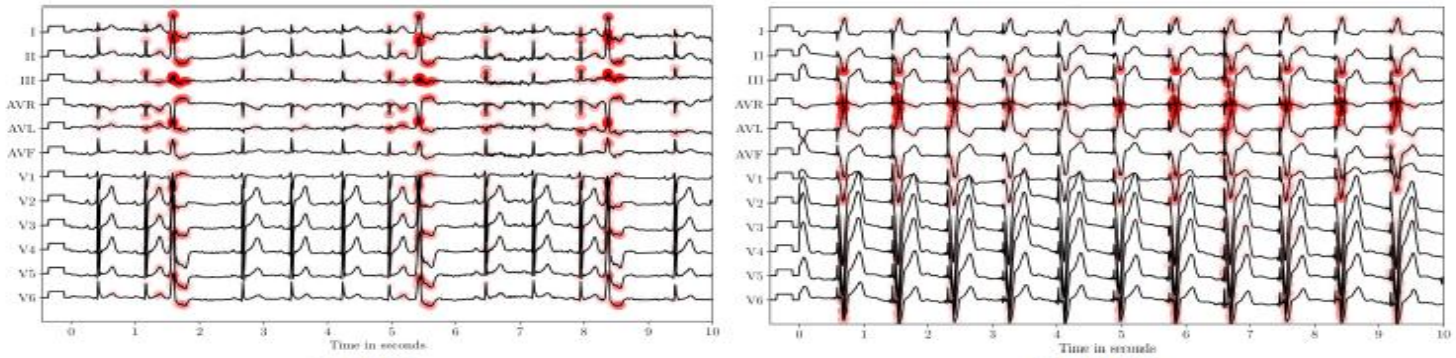
Machine Learning for ECG Classification

Benchmarking: Diagnosis

Method	all	diag.	sub-diag.	super-diag.	form	rhythm
inception1d	**.925(08)**	**.931(09)**	**.930(10)**	.921(06)	**.899(22)**	**.953(13)**
xresnet1d101	**.925(07)**	**.937(08)**	**.929(14)**	**.928(05)**	**.896(12)**	**.957(19)**
resnet1d_wang	.919(08)	**.936(08)**	**.928(10)**	**.930(05)**	.880(15)	**.946(10)**
fcn_wang	.918(08)	.926(10)	**.927(11)**	.925(06)	.869(12)	.931(08)
lstm	.907(08)	.927(08)	**.928(10)**	.927(05)	.851(15)	**.953(09)**
lstm_bidir	.914(08)	.932(07)	**.923(12)**	.921(06)	.876(15)	**.949(11)**
Wavelet+NN	.849(13)	.855(15)	.859(16)	.874(07)	.757(29)	.890(24)
ensemble	**.929(07)**	**.939(08)**	**.933(11)**	**.934(05)**	**.907(12)**	**.965(07)**



„Heatmapping“ / Explainability



(a) PVC (b) PACE

N. Strodthoff, et al. IEEE Trans. Biomed. Health Informatics (2020)

„Digital twin“

WP1: **Modelling pipeline** – Ventricles

REACTION-EIKONAL

Bidomain Formulation

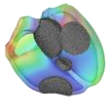
$$\beta^{-1}\nabla\cdot(G_i\nabla\phi_i) = C_m\partial_t V_m + I_{ion}(V_m,n)$$
$$\beta^{-1}\nabla\cdot(G_e\nabla\phi_e) = -C_m\partial_t V_m - I_{ion}(V_m,n)$$
$$V_m = \phi_i - \phi_e \qquad +I_{foot}$$
$$\partial_t n = F(V_m,n)$$

Assumption: $G_e = \lambda G_i$

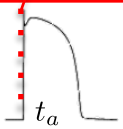
Monodomain

$$C_m\partial_t V_m = \beta^{-1}\nabla\cdot(G_m\nabla V_m) - I_{ion}(V_m,n) \qquad +I_{foot}$$
$$G_m = \frac{G_i G_e}{G_i + G_e}$$

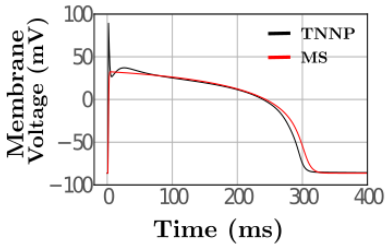
Eikonal Activation



$$\begin{cases} \sqrt{t_a^T \mathbf{V}\nabla t_a} = 1 & \text{in } \Omega_i \\ t_a = 1 & \text{in } \Gamma \end{cases}$$



Ionic Model: TT2 vs. Mitchell Schaeffer

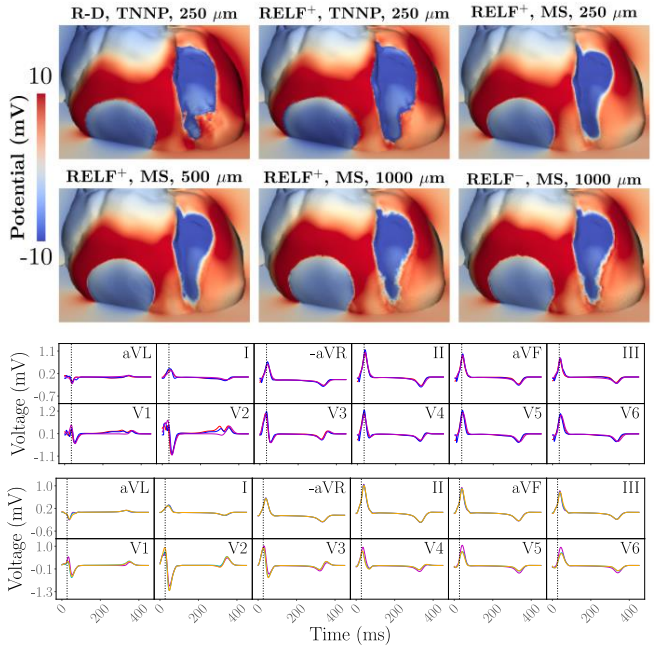


Lead-Field Mapping

$$\phi_{ecg}(t) = \int \nabla Z(x)\cdot G_i\nabla V_m dx$$
$$\nabla\cdot((G_i+G_e)\nabla Z(x)) = \sum_j c_j\delta(x-x_j)$$

Simulation with 450 ms (10):
Bidomain: 3.29 hours
Pseudo-bidomain: 1.77 hours
R-E+ with Lead Field: 63 s

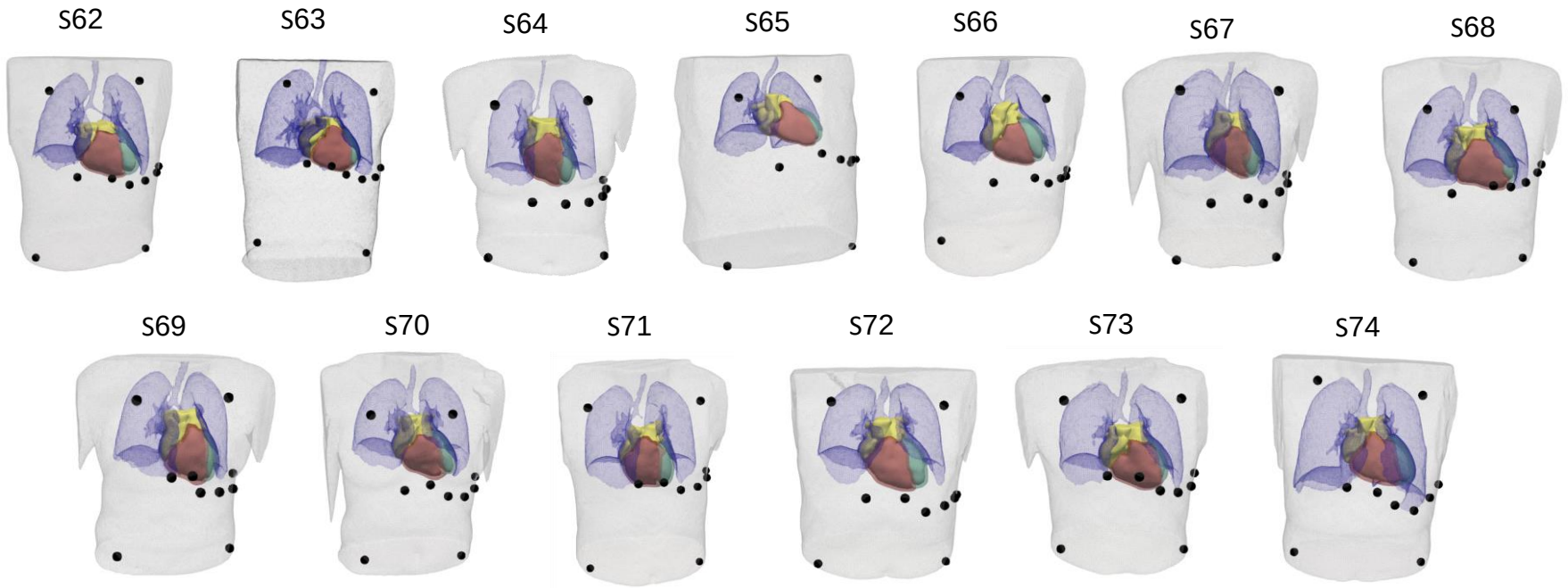
Validation



K. Gillette, A. Prassl, G. Plank et al. (2021)

Digital Twin II

WP1: **Modelling pipeline** – Model cohort

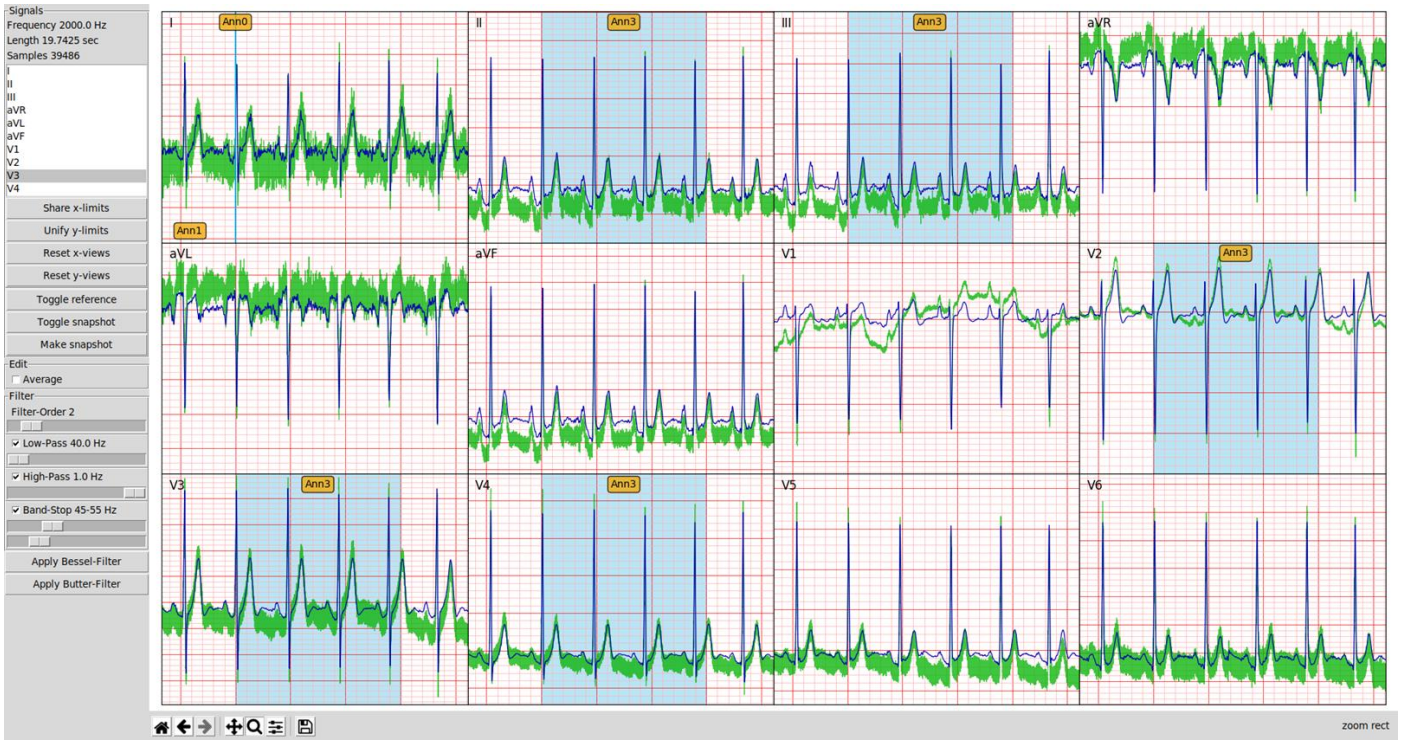


Models: 13 (9 male, 4 female)
Model Resolution: 1201 +/- 56.21 microns

K. Gillette, A. Prassl, G. Plank et al. (2021)

Simulated ECGs

WP1+3: **Modelling pipeline** – ECG simulation + processing



Validation by statistical comparison of synthetic data base with PTB-XL

Summary

Metrology approach to Trustworthy ML/ AI:
Uncertainty evaluation, robustness, explainability

Training data are key: Lack of high-quality references

-> Well characterized measurement data
-> Validated synthetic reference data from „digital twins“

Input to standardization ?

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Abbestraße 2-12
10587 Berlin

Prof. Dr. Markus Bär
Department 8.4 Mathematical Modelling & Data Analysis

Telefon: +49 30 3481 7687
E-Mail: markus.baer@ptb.de
www.ptb.de

PTB Strategy “Metrology for AI in Medicine”

Autumn 2020: Strategy paper analysing demands regarding Standardisation & Regulation of AI products (inter alia) from

- German Government‘s Strategy Artificial Intelligence
- DIN “AI Standardisation Roadmap”
- EUROLAB position paper
- FDA Regulatory Framework Proposal

Key messages

- Core requirements:
  - Functionality and performance → Uncertainty
  - Robustness
  - Explainability
- Essential: Selection & assessment of data used

Implementation

- Metrology for AI in medicine as a nucleus

  (coordinators Hans Rabus, David Auerbach)

- Focus on generic methods, transferable to other areas (autonomous driving, data evaluation, ...)
- Collective call for 10 PhD / PostDoc positions from a portfolio of 13 projects related to

  - Basic research on trust in AI
  - Application of AI in medicine
  - Type testing for medical devices/methods with AI components

List of Projects

1	Towards standardized quality control for artificial intelligence systems in critical care [B]
2	ML and uncertainty quantification for bioelectromagnetic inverse solutions and signal separation methods [B]
3	Advancing the theory and practice of machine learning model explanations in biomedicine [B]
4	Invertible neural networks for resolving the hemodynamic inverse problem [B]
5	Robust machine learning-based quantitative magnetic resonance imaging [B]
6	Active learning using Fisher information [B]
7	Uncertainty in deep learning versus conventional statistics [B]
8	Artificial intelligence and metabolite markers in diagnosis and prognosis of Parkinson's disease [BS]
9	AI-based image enhancement for reduced radiation exposure in computed tomography imaging [BS]
10	Deep learning-based dosimetry in medical x-ray imaging [BS]
11	Uncertainty of artificial intelligence-based dose prediction compared to Monte Carlo methods [B]
12	Incorporation of spatial regularization and uncertainty estimations into magnetic-resonance parametric mapping [B]
13	Accelerating radiation transport simulations in radiation medicine by machine learning [B]