Institute of Mechatronic Systems
imes Leibniz Universität Hannover

Leibniz Universität Hannover

International RAVE Workshop 2022

Improved domain adaptation for condition monitoring of wind turbines

Kortmann, Karl-Philipp (Speaker)
Ziaukas, Zygimantas

Location: Online
Date: 2022/02/03

RAVE
RESEARCH AT ALPHA VENTUS

Source

Target

Source: https://rave-offshore.de, ©DOTI / Matthias Ibeler

Agenda

- Motivation of Domain Adaptation (DA)
- Previous work in wind energy
- Methodology
  - Proposed DA model
  - Baseline DA models
  - Raw data description
- Results
- Next steps

Motivation of DA: Examples

Transfer of an existing estimator from a previous (training) domain to a new target domain.

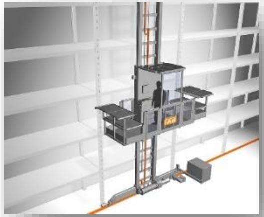
Examples:

- Regression of tire pressure (new vehicle type)
- Regression of belt tension (individual stacker crane)
- Condition monitoring in wind farms

Problems with traditional approches:

- Presumably poor results for non-adapted estimator
- Costly:
  - recording of labelled data of the new (target) domain
  - data pre-processing and model engineering

Source: https://www.autozeitung.de



Source: https://www.lenze.com

Source: https://rave-offshore.de, ©DOTI / Matthias Ibeler

Motivation of DA: Taxonomy

Sub-task of Transfer Learning

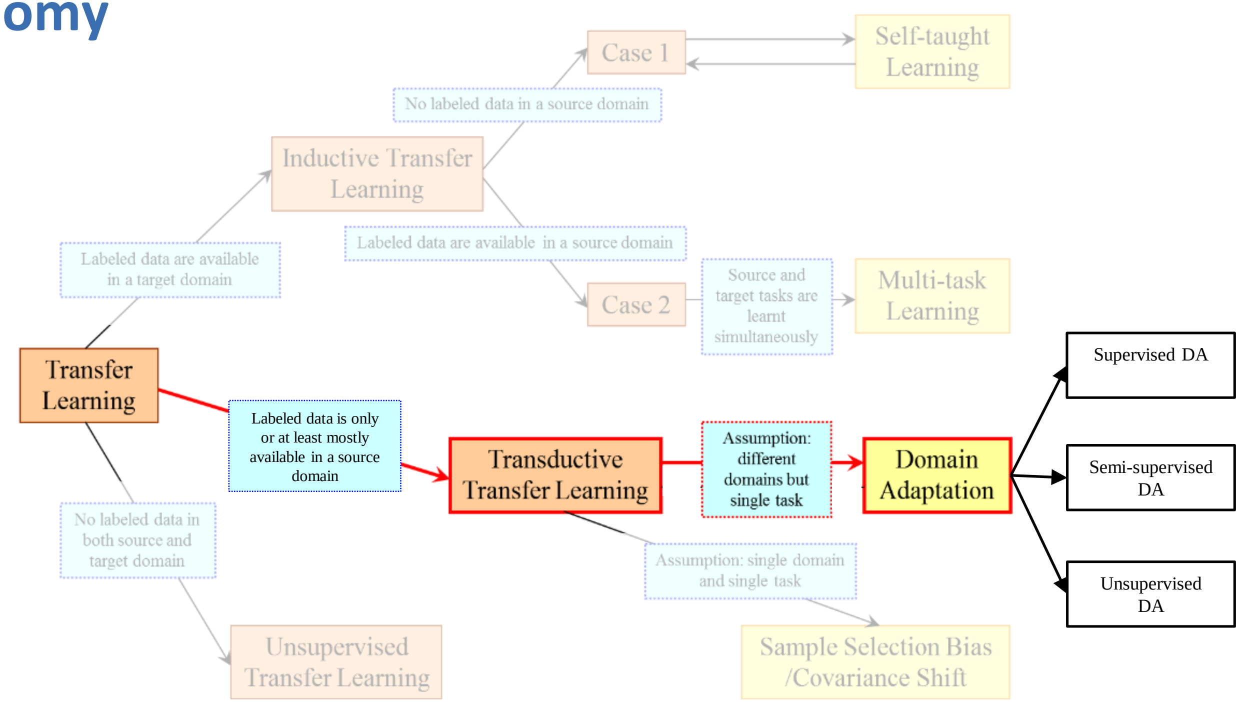
Requirements / Assumptions:

1. same task for both domains
2. high amount of labeled data from source domain

3.a) only unlabeled data from target domain or

3.b) partly labeled data from target domain.

Methodological categorization by *D. Tuia et al. (2016)*:



adapted from S. J. Pan and Q. Yang (2010)

Motivation of DA: Taxonomy

Sub-task of Transfer Learning

Requirements / Assumptions:

1. same task for both domains
2. high amount of labeled data from source domain

3.a) only unlabeled data from target domain or

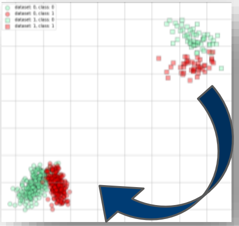
3.b) partly labeled data from target domain.

Methodological categorization by *D. Tuia et al. (2016)*:

Learning/selection of invariant features (unsup.):

$$\underset{v\in V, v\neq\{\}}{\arg\min}\ \mathrm{Div}(P(X_s^v, Y_s^v), P(X_t^v, Y_t^v))$$

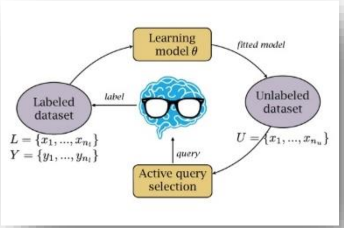
Data alignment (unsup):



Estimator adaption (sup.):

e.g. retrain last layers of ANN

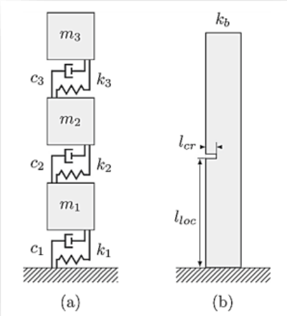
Active learning (sup.):

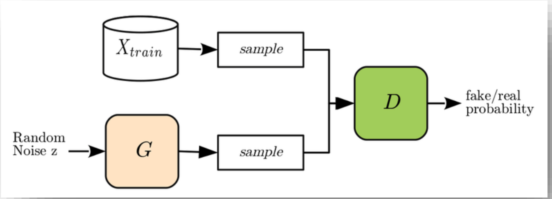


Deep learning approaches…

Previous work in wind energy

- DA so far only rarely applied to estimation tasks for wind turbines (first publications towards SHM since 2017)
- P. Gardner et al. (2019) compared two data alignment methods (transfer component analysis (TCA) and joint domain adaptation (JDA)) in order to classify cracks in WT blades.
- W. Juang and J. Jin (2021) used SCADA data to detect blade icing across two domains using a generative adverserial networks (GAN).





Methodology

Basic estimation tasks:

- Regression of generator power using blades‘ strain measurements (*„academic“* task)
- Regression of strain in tower and blade segments using SCADA data
- ~~Ice detection on blades~~

Wind turbines used (a.k.a. domains):

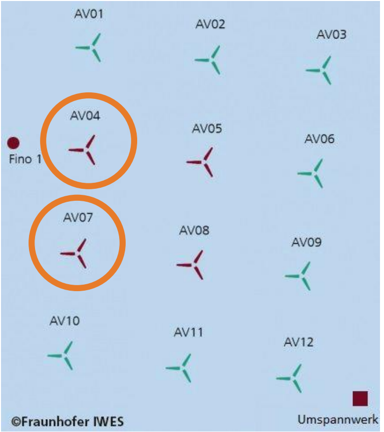
- AV-07 (Adwen)
- AV-04 (Senvion)

RAVE data used:

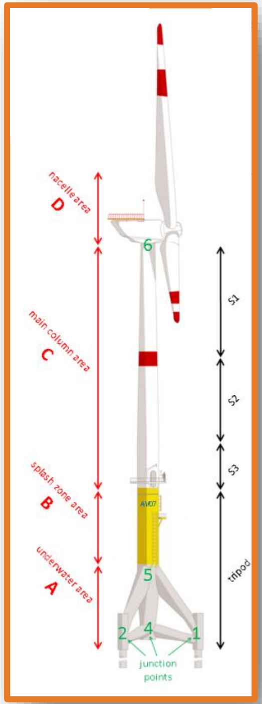
- SCADA and blade / tower strain measurements from RAVE research archive

Estimation algorithm:

- MLP (tanh activation, MSE objective, SGD optimization)

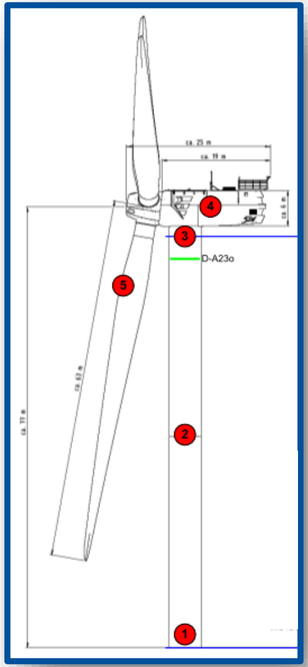


Methodology: Data



Domains: AV-07 (Adwen), AV-04 (Senvion)

- tower strain data (DMS, 10 min)
  - lowest tower segment
  - four gauges around tower
- blade strain data (DMS, 10 min)
  - measurment location next to hub
  - flap- and edgewise
- SCADA data (10 min, normalized)
  - rotor speed, wind speed
  - pitch angle
  - eff. generator power
- FINO data (10 min)
  - Environmental and weather data
- Derived variables (planned)
  - Short-term 10 min damages [1]

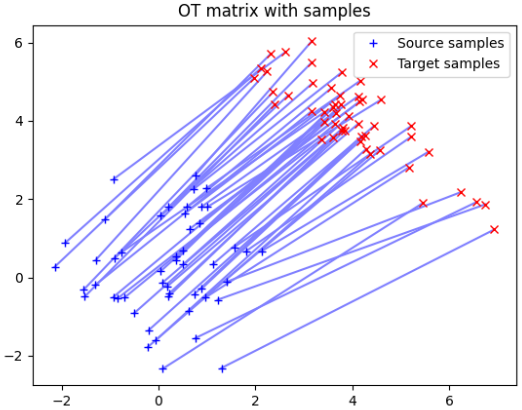


[1]: Barradas-Berglind, J. J., and Rafael Wisniewski. "Representation of fatigue for wind turbine control." *Wind Energy* 19.12 (2016): 2189-2203.

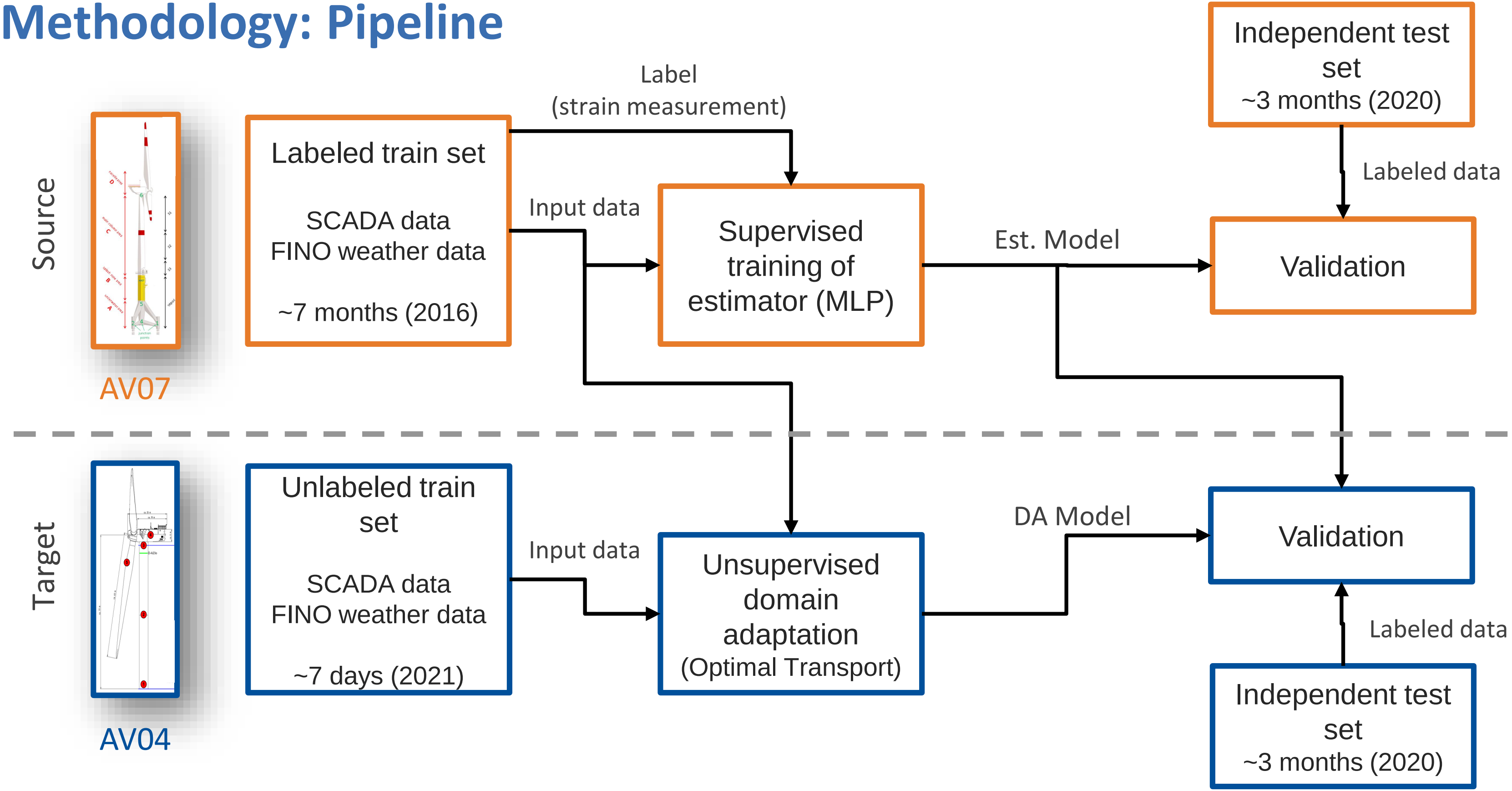
Methodology: Proposed DA model

Optimal Transport (OT)

- Originally used to measure distances between distributions (e.g. Wasserstein distance)
- Growing attention in domain adaptation since ~2015
- Idea: Find (optimal) mapping from $P(X_s)$ to $P(X_t)$ (or vice versa) with respect to the shift expense
- Assumptions:
  - homogenious task ($\gamma_s = \gamma_t$)
  - no / small target value imbalance $P(Y_s|X_s) \approx P(Y_t|X_t)$
- Advantage over other data alignment techniques (e.g. PCA, TCA, JDA):
  - non-linear mapping
  - both unsupervised **and** semi-supervised applications possible
- Algorithm used: Convex group-lasso regularized OT (Sinkhorn-Knopp algorithm)

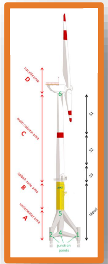


Methodology: Pipeline



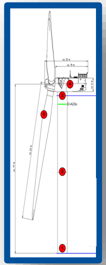
Methodology: Pipeline

Source

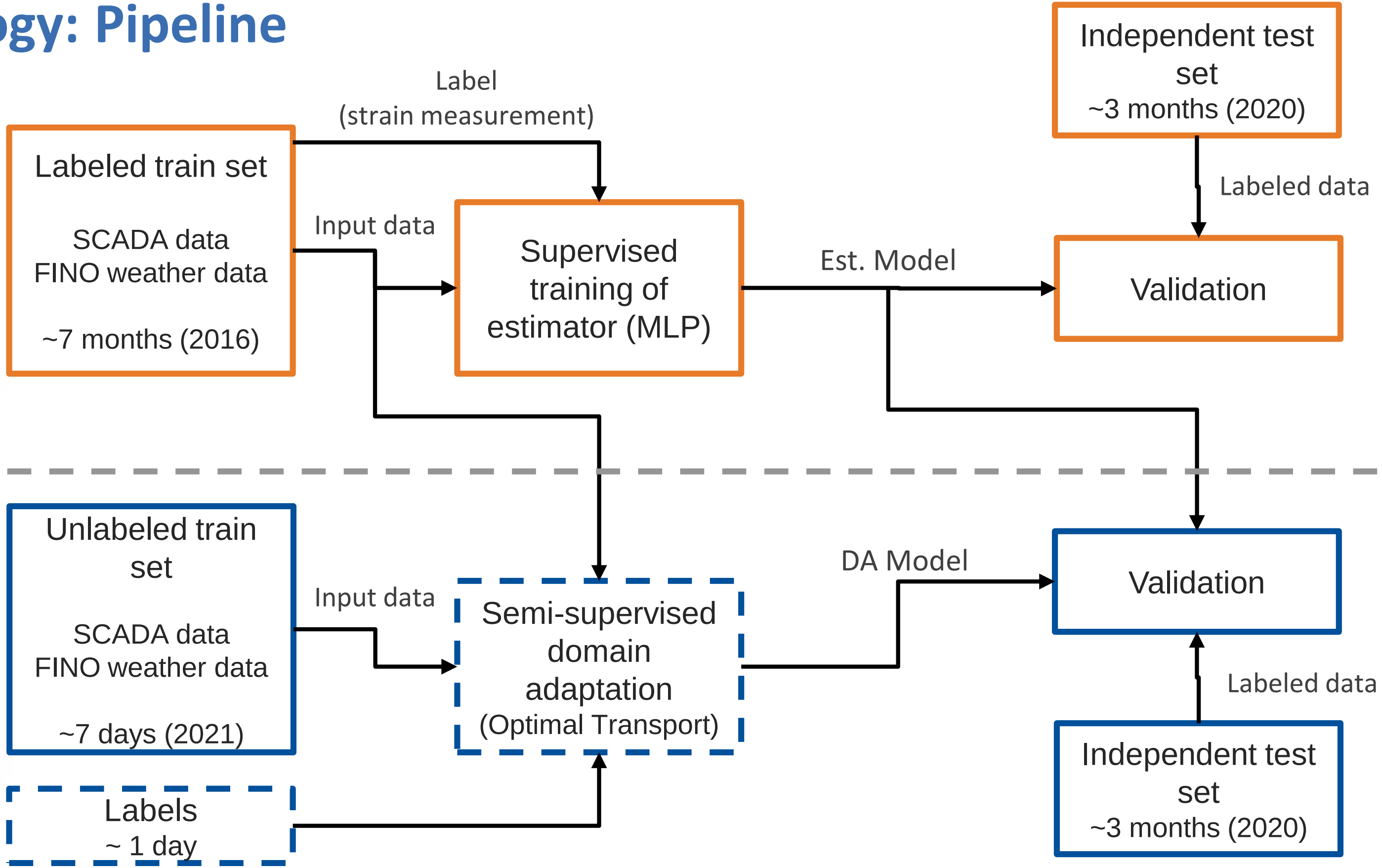


AV07

Target



AV04



Methodology: Baseline DA models

Other DA methods used for benchmarking:

- No domain adaptation at all:
  Applying the source domain trained estimator directly to the target domain
- Feature normalization:
  Normalizing each input variable i.o.t. equalize value range across the domains
- Principle component analysis (PCA):
  Linear data mapping between two domains

Results: Strain estimation

Same estimator (MLP) for all models.

Training data amount: Source (7 months), Target (7 days (unlabeled), 1 day (labeled))

Table 1: Exemplary results (Normalised RMSE) for the estimation of mean tower strain (10 min) using un- and semi-supervised domain adaptation.

Domain Adaptation (direction)	**NRMSE for domain adaptation method:**				
	no data alignment	**Normalization**	**PCA**	**Sinkhorn OT (unsupervised)**	**Sinkhorn OT (semi-supervised)**
AV04 -> AV07	1.699	1.503	0.836	**0.178**	**0.102**
AV07 -> AV04	1.544	0.819	0.922	**0.165**	**0.093**
AV07 only (no DA)	0.058	-	-	-	-
AV04 only (no DA)	0.076	-	-	-	-

Results: Power estimation

Same estimator (MLP) for all models.

Training data amount: Source (7 months), Target (7 days (unlabeled), 1 day (labeled))

Table 2: Exemplary results (Normalised RMSE) for the estimation of mean effective power generation (10 min) using un- and semi-supervised domain adaptation.

Domain Adaptation (direction)	**NRMSE for domain adaptation method:**				
	no data alignment	**Normalization**	**PCA**	**Sinkhorn OT**	**Sinkhorn OT (semi-supervised)**
AV04 -> AV07	0.945	0.252	0.179	**0.094**	**0.059**
AV07 -> AV04	0.697	0.115	0.153	**0.088**	**0.075**
AV07 only (no DA)	0.046	-	-	-	-
AV04 only (no DA)	0.065	-	-	-	-

Next steps

- Application to classification tasks (seeking for good examples within RAVE)
- Comparison to state of the art deep learning DA methods (e.g. GAN-based approaches)
- Investigating the influence of the amount of data available (both target and source domain)
- Domain adaptation across different wind farms, not only different wind turbines
  - Stronger target value imbalance $P(Y_s|X_s) \neq P(Y_t|X_t)$
  - Different sensors $\Omega_s \neq \Omega_t$

Institute of Mechatronic Systems
Leibniz Universität Hannover

Leibniz Universität Hannover

Thanks for your attention!

Questions?

Contact details:

Kortmann, Karl-Philipp:
kortmann@imes.uni-hannover.de

RAVE
RESEARCH AT ALPHA VENTUS

Appendix: Data periods

WT	Start	End	Duration	Function
AV-04	2019-04-01	2020-03-31	366	train
AV-04	2020-08-01	2021-07-31	365	test
AV-07	2016-08-01	2017-07-31	365	train
AV-07	2015-03-01	2015-07-31	153	test
AV-07	2016-11-01	2017-02-28	120	test
AV-07	2020-10-01	2021-02-28	151	train

Appendix: Notation

Domain D of

- d-dimensional (feature) space $\Omega \in \mathbb{R}^d$ with marginal prob. dist. $P(X)$
- a task T defined by a label space γ and cond. prob. dist. $P(Y|X)$ with (multivariate) random variables X and Y.

Source domain $D_s = \{\Omega_s, P(X_s)\}$ with $T_s = \{\gamma_s, P(Y_s|X_s)\}$

Target domain $D_t = \{\Omega_t, P(X_t)\}$ with $T_t = \{\gamma_t, P(Y_t|X_t)\}$

Given a Dataset $\boldsymbol{X} = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\} \in \chi$ with labels $\boldsymbol{Y} = \{y_1, \dots, y_n\} \in \gamma$, we try finding an estimator $f(\cdot) = P(\boldsymbol{Y}|\boldsymbol{X}) \approx P(Y|X)$.

Traditional ML: $D_s = D_t$ and $T_s = T_t$

Transfer Learning: $D_s \neq D_t$ or $T_s \neq T_t$, examples:

- Class Imbalance: Different label distribution $P(Y_s) \neq P(Y_t)$, but at least $P(X_s|Y_s) = P(X_t|Y_t)$.
- Covariate Shift: $P(Y_s|X_s) = P(Y_t|X_t)$, but (small) difference in data distribution $P(X_s) \neq P(X_t)$.

adapted from N. Courty et al. (2017) and G. Csurka (2017)