

Vorhersagen für grüne Energie: Warum Transfer Learning den Unterschied macht

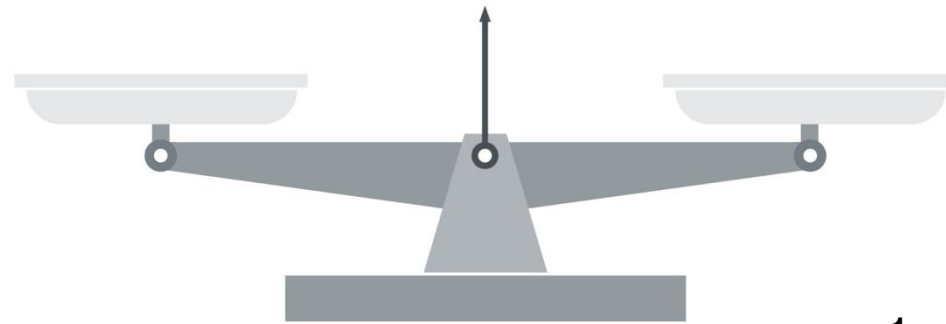
TODO ADD QR CODE

Präsentiert von Jens Schreiber



Blackout

Wodurch entstehen Blackouts?



1.

49,8 Hz - Einsatz der noch nicht mobilisierten Kraftwerksleistung

48,4 Hz - Sofortiger automatischer Lastabwurf von 15-20% der Netzlast

47,5 - Blackout

50,2 - ältere Solaranlagen schalten ab, neue reduzieren ihre Einspeisung

51,5 - Abtrennen der Kraftwerke vom Netz, alle Solaranlagen schalten ihre Einspeisung ab

Problem mit erneuerbaren Energien für die Netzstabilität

Herausforderung: Ausgleich zwischen Last und Erzeugung

- Traditionelle Kraftwerke: planbare Leistung mit Vorlaufzeit
- Erneuerbare Energien: wetterabhängig und weniger steuerbar

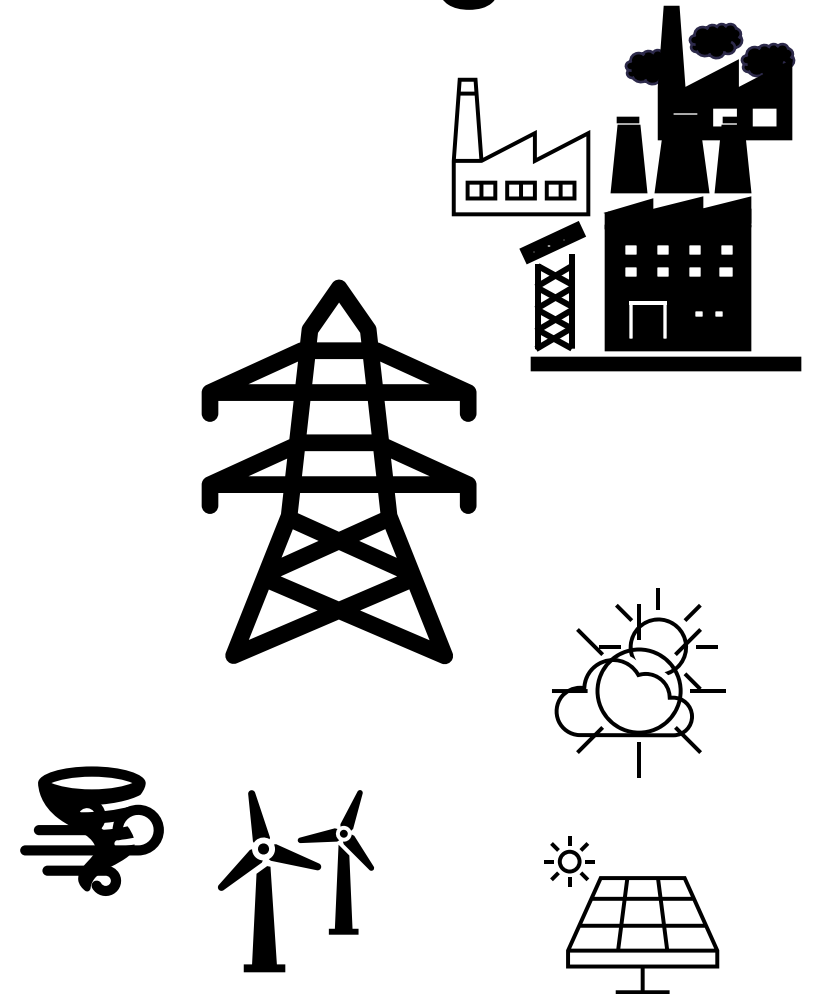
Besonderheiten erneuerbarer Energien

- **Variabilität:** Sonnenschein und Wind schwanken stark
- **Geografische Diskrepanz:** Strom wird nicht immer dort erzeugt, wo er benötigt wird

→ **Infrastrukturbedarf:** Netzausbau und Speichersysteme

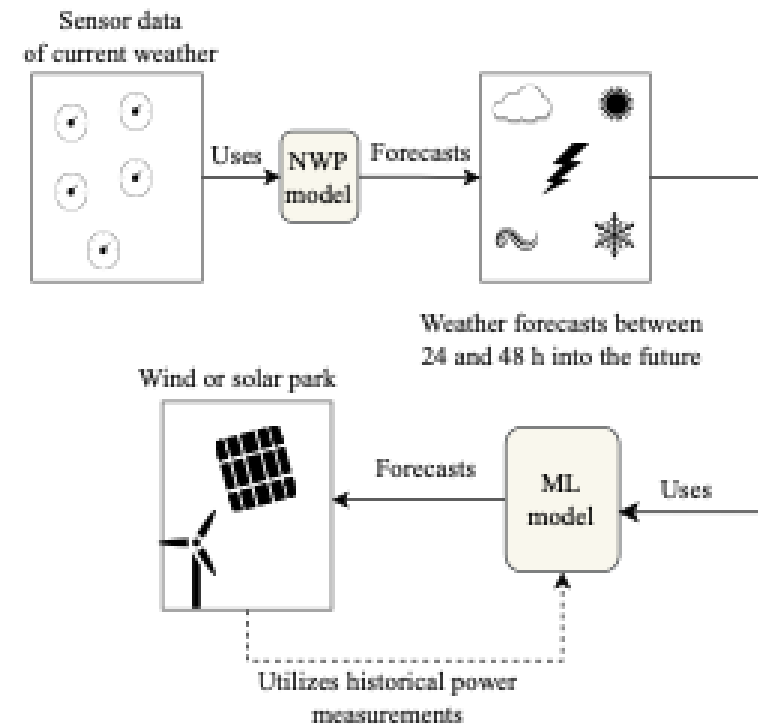
→ **Planung ist entscheidend**

Prognosen für heutige und morgige Leistung notwendig
Grundlage für Netzstabilität und effiziente Stromnutzung



Wie werden Vorhersagen für erneuerbare Energie erstellt?

- Die maschinelle Lernverfahren werden auf Basis historischer **Erzeugungsdaten** eines Parks und **Wettervorhersagen** als Eingabe trainiert, um die zukünftige Stromerzeugung abzuschätzen.
- Physikalische Modelle werden eingesetzt, wenn nicht genügend historische Daten für maschinelles Lernen vorhanden sind.
 - Für Windenergie sind Faktoren wie Windgeschwindigkeit (kubischer Zusammenhang), Rotorfläche, Luftdichte und Effizienz entscheidend; für Solarenergie sind Azimut und Neigung der Solarmodule wichtig.
- **Numerischen Wettermodellen (NWP) Modelle prognostizieren meteorologische Größen** wie Windgeschwindigkeit und Strahlung auf Basis physikalischer Gleichungen und aktueller Wetterdaten.
- Der Prognosehorizont wird je nach Anwendung gewählt, z.B. Folgetagsprognosen (24–47 Stunden) oder Intraday (6–24 Stunden).

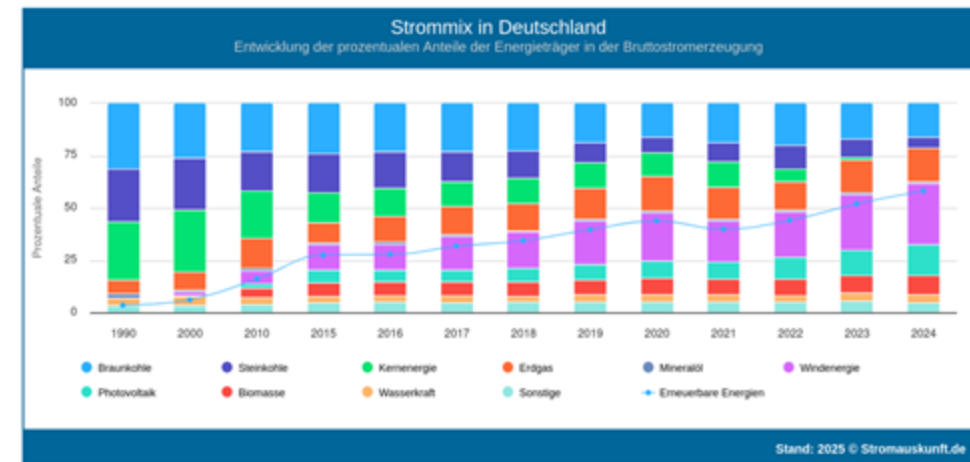
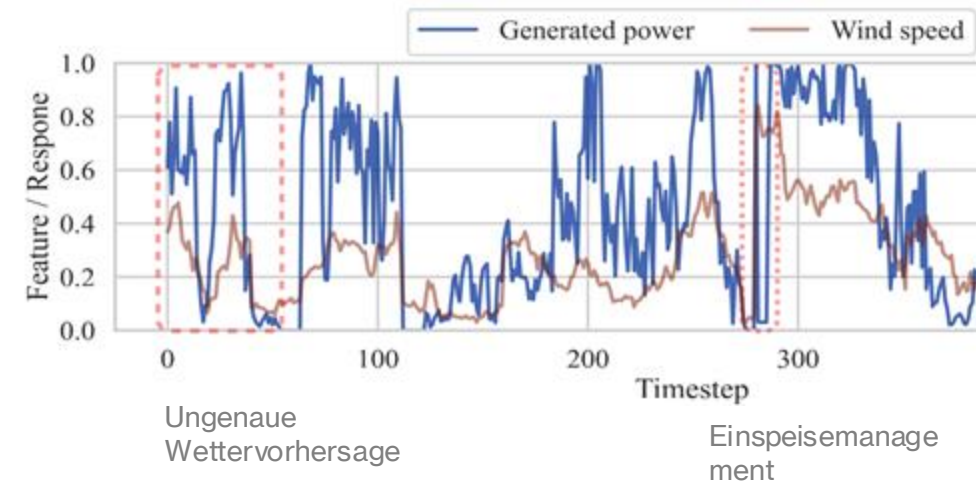


Was ist das Problem mit dem Stand der Technik?

- Prognosen für erneuerbare Energien sind **stark von Wettervorhersagen abhängig**, was zu erheblicher Unsicherheit führt.
- Wettervorhersagen gelten oft für große Gebiete, wodurch **lokale Besonderheiten** von Wind- und PV-Parks nicht ausreichend berücksichtigt werden.
- Zusätzliche Unsicherheiten entstehen durch **Netzinterventionen** (z.B. Einspeisemanagement) und spezielle Betriebsbedingungen wie nächtliche Drosselung von Windturbinen.
- Es gibt eine **natürliche Grenze für die Prognosegenauigkeit**
- Bei **neuen Anlagen mit wenig historischen Daten** sind Prognosen besonders schwierig, da wichtige Muster und saisonale Einflüsse noch nicht ausreichend gelernt werden können. Hier werden aktuelle physikalische Modelle genutzt.

⚡ Je mehr erneuerbare Energie eingesetzt werden, desto relevanter wird es jede dieser Fehlerquellen zu reduzieren.

→ Techniken der **Wissensübertragung** (Transfer Learning) **nutzt bestehendes Wissen** in Form von Modellen und Daten um die Prognosegenauigkeit zu erhöhen und die “Probleme” zu reduzieren.

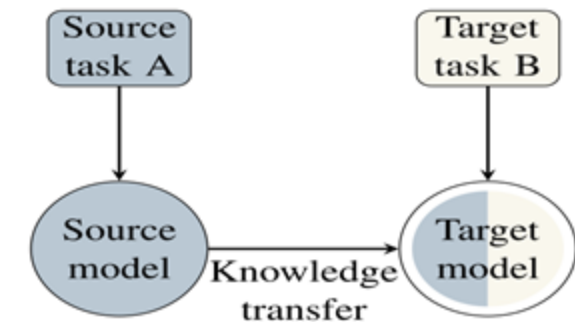
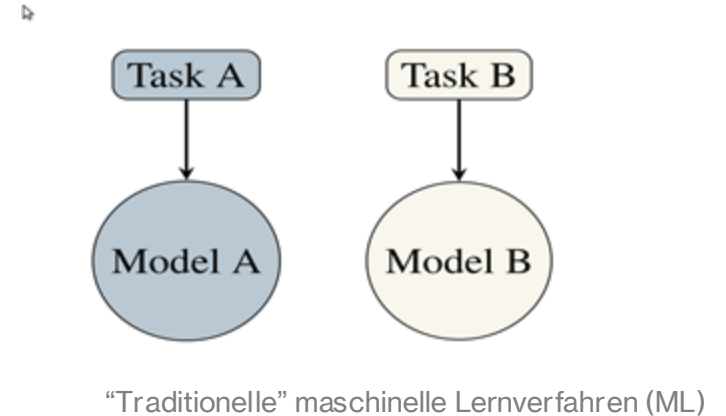


Grafik: Strommix in Deutschland seit 1990

<https://www.stromauskunft.de/strompreise/strommix-in-deutschland/>

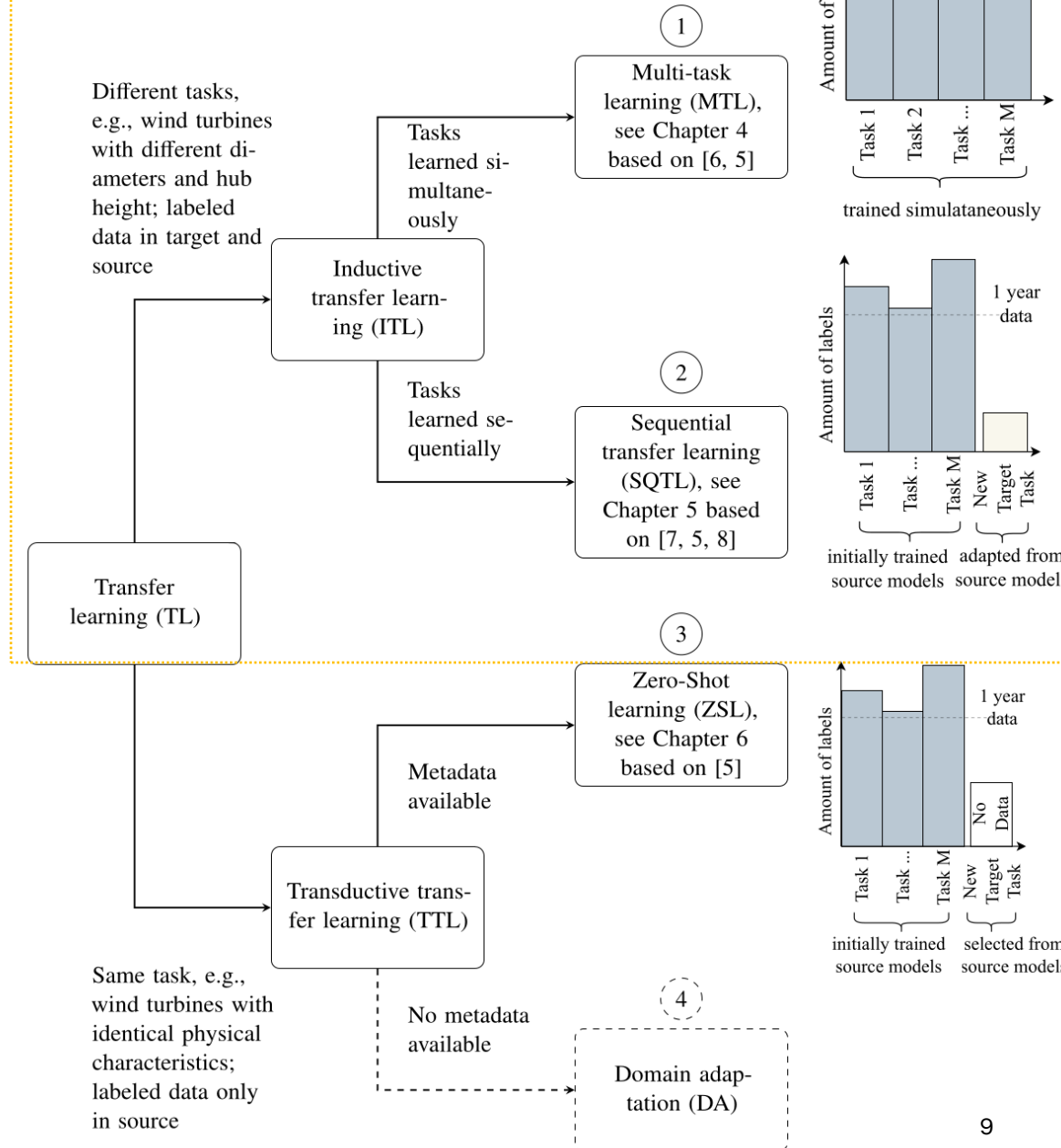
Was ist die Idee von Transfer Learning?

- In traditionellen ML-Verfahren wird für jede Aufgabe („Task“) ein eigenes Modell **unabhängig voneinander trainiert**, z.B. Task A (Windpark mit nächtlicher Drosselung) und Task B (neuer Windpark mit ähnlichem Verhalten) werden separat gelernt.
- **Transfer Learning (TL)** ermöglicht die **Übertragung von Wissen**
 - Eine „**Domain**“ ist definiert durch die Eingabemerkmale (Feature Space) und deren Verteilung; eine „**Task**“ beschreibt die Vorhersagefunktion und die Zielgröße (Label Space).
 - Im Gegensatz zu traditionellem ML-Verfahren wird beim TL **Wissen aus ähnlichen Tasks** genutzt, um mit begrenzten Trainingsdaten bestmögliche Vorhersagequalität zu erreichen.
 - TL kann die **Prognosegenauigkeit verbessern und den Trainingsaufwand reduzieren** z.B. durch Multi-Task, Sequential-Transfer oder Zero-Shot-Learning.
 - Risiken bestehen bei der Übertragung zwischen unähnlichen Aufgaben (**Negative Transfer**), daher ist die Auswahl passender Source Tasks entscheidend.



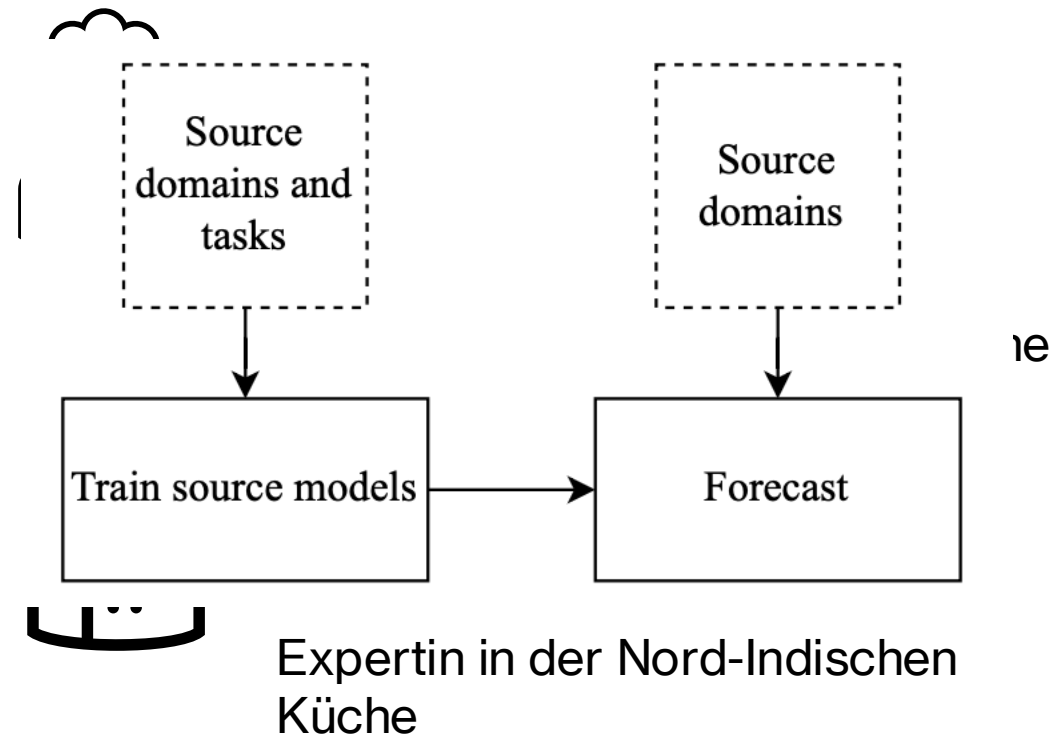
Übersicht Wissenstransfer

- **Induktives TL** nutzt gelabelte Daten in Quelle und Ziel, während **transduktives TL** nur gelabelte Daten in der Quelle und keine im Ziel verwendet.
- **Multi-Task-Learning (MTL)**: Mehrere Aufgaben werden gleichzeitig mit einem gemeinsamen Modell gelernt, um Ähnlichkeiten und Unterschiede zwischen Parks auszunutzen und die Prognosegenauigkeit zu verbessern.
- **Sequential-Transfer-Learning (SQTl)**: Ein vortrainiertes Modell aus einer Quelle wird für eine neue Aufgabe mit wenig Daten angepasst, z.B. für neu gebaute Parks mit nur einer Saison an Daten.
- **Zero-Shot Learning**: Für neue Parks ohne historische Daten wird ein ähnliches Modell aus der Quelle direkt angewendet, basierend auf Meta-Daten wie physikalischen Eigenschaften.
- **Domain Adaption**: Modelle werden für unterschiedliche Eingabedaten angepasst, z.B. wenn Modelle trainiert mit Messdaten der Windgeschwindigkeit auf Wettervorhersagen als Eingabe angepasst werden. Üblicherweise nicht der Fall, da ausreichend Wettervorhersagen vorhanden sind.



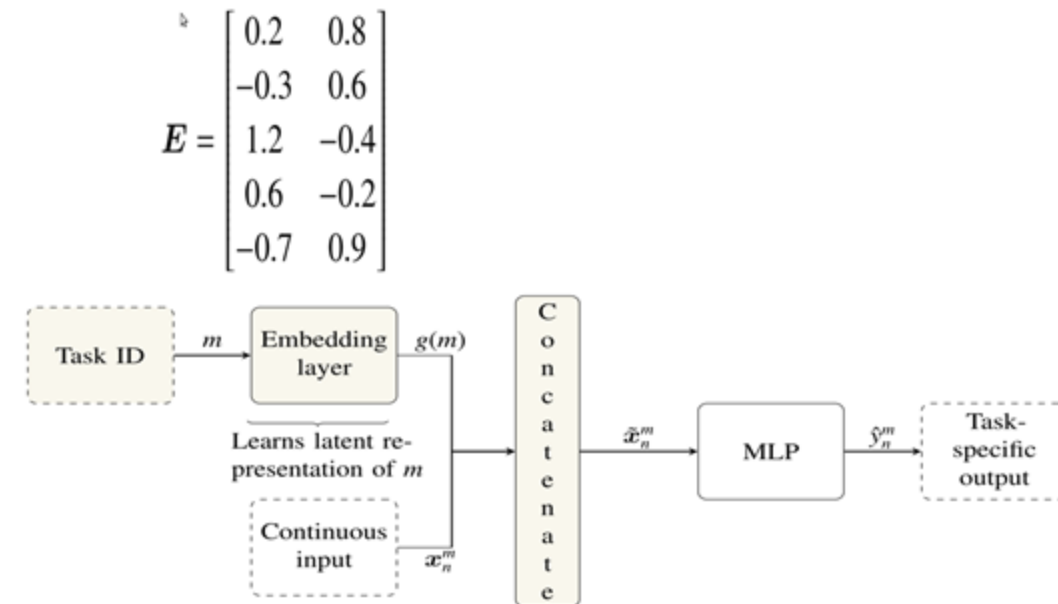
Multi-Task-Learning

- Multi-Task Learning (MTL) **trainiert ein gemeinsames Modell für mehrere Tasks aus unterschiedlichen Source-Domains** gleichzeitig und nutzt dabei Ähnlichkeiten und Unterschiede zwischen den Tasks.
- Durch das gemeinsame Lernen werden die **Prognosegenauigkeit und die Generalisierungsfähigkeit** erhöht, da Wissen aus verschiedenen Source-Domains und Tasks geteilt wird.
- Beispiel: In der erneuerbaren Energie werden mit MTL gleichzeitig Prognosemodelle für mehrere Wind- oder PV-Parks trainiert, sodass das Modell **saisonale Muster, extreme Wetterereignisse und Abschaltungen** (zum Beispiel Fledermausabschaltung) über verschiedene Tasks hinweg besser erfassen kann.



Task-Embedding für MLP und TCN

- **Jedem Park (Task)** wird eine **eindeutige ID** zugeordnet, die als Schlüssel für das Task Embedding dient.
- Die **Task-ID** wird durch eine **Embedding-Schicht** in einen latenten Vektor **transformiert**, der Ähnlichkeiten zwischen Parks abbildet.
- Im MLP-Ansatz wird **der Embedding-Vektor mit anderen Eingabefeatures verkettet**, um task-spezifische Vorhersagen zu ermöglichen.
- Der Ansatz ist auch erweiterbar für Zeitreihenmodell wie Temporal Convolutional Networks
- Das Task Embedding erlaubt die Erweiterung auf neue Parks **ohne Catastrophic Forgetting**, da für jeden neuen Task einfach ein neuer Vektor gelernt wird.



Task-Embedding für Multi Layer Perceptron

Zusammenfassung Multi-Task-Learning

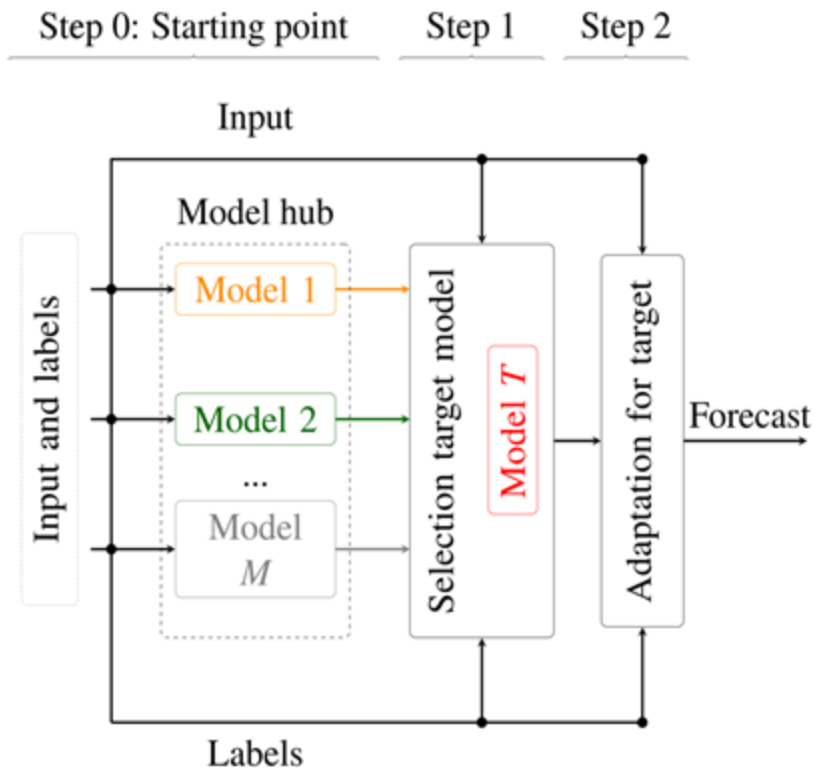
- MTL-Architekturen (Task Embeddings) **verbessern die Prognose** signifikant zwischen **7,8% und 14.6% für Wind und zwischen 0,8% und 25.3% für PV** je nach Datensatz.
- Sie reduzieren die **benötigte Anzahl an Modellparametern pro Park bis zu 43-fach für PV und 8-fach für Wind** je nach Datensatz.
- Temporal Convolutional Networks-basierte Methoden sind Multi Layer Perceptron-basierten Ansätzen überlegen, da sie zeitliche Zusammenhänge besser erfassen und mehr relevante Informationen aus den Zeitreihen nutzen.
- Die Modelle sind flexibel einsetzbar, unabhängig von verfügbaren Kategorischen-Merkmalen, und liefern in allen Fällen mindestens so gute Ergebnisse wie die Baseline.
- Die Task Embeddings ermöglichen eine direkte Analyse und Interpretation der latenten Repräsentationen, z.B. zeigen sie Korrelationen zu physikalischen Eigenschaften wie dem Azimutwinkel in PV-Parks, ohne zusätzliche XAI-Algorithmen.

Sequential-Transfer -Learning

- SQTL nutzt Wissen aus einer Source Domain und einer Source Task für einen neuen **Target Task mit wenigen historischen Daten**.
- Die **Source Domain umfasst die Parks und deren historische Daten**, während die Target Domain die neue Anlage mit begrenzten Daten beschreibt.
- Ein **vortrainiertes Quellmodell** dient als Ausgangspunkt für die Anpassung an die Zielaufgabe.
- Die **Auswahl des passenden Quellmodells erfolgt anhand von Ähnlichkeiten zwischen Source und Target Task**, z.B. durch Metadaten oder latente Embeddings.
- Ziel von SQTL ist eine **schnelle und effiziente Übertragung von Wissen**, um auch mit wenigen Daten zuverlässige Vorhersagen für die neue Aufgabe zu ermöglichen.



SQTL - Modellauswahl



Algorithm 1 Model Selection in Sequential Transfer Learning.

Require: Source models $\{f_1, f_2, \dots, f_M\}$

Require: Target data X_T, Y_T

Ensure: Best model for the target task

Initialize $bestModel \leftarrow \text{null}$

Initialize $highestSimilarity \leftarrow -\infty$

for f_m in $\{f_1, f_2, \dots, f_M\}$ **do**
 $similarity \leftarrow \text{SimilarityMeasure}(f_m, X_T, Y_T)$
 if $similarity > highestSimilarity$ **then**
 $highestSimilarity \leftarrow similarity$
 $bestModel \leftarrow f_m$

end if

end for

return $bestModel$

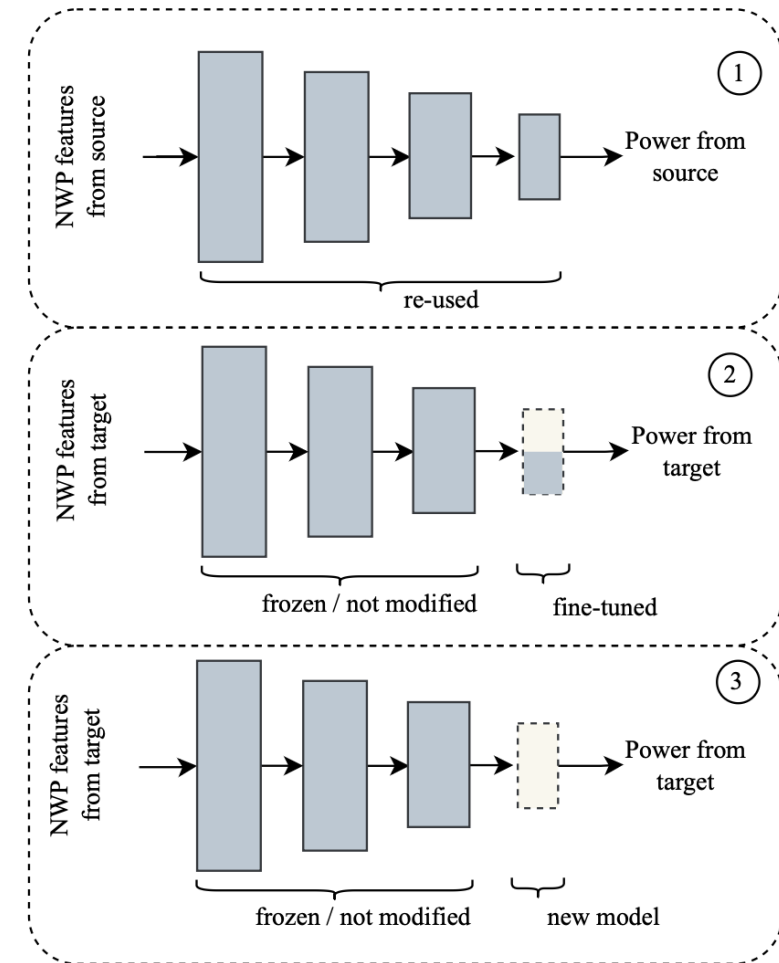
SQTL – Modellanpassung für Single-Task-Learning

Model Anpassung durch Fine-tuning

- Die **Gewichte eines vortrainierten Modells** werden für die **Zielaufgabe weiter angepasst**, meist durch Training der letzten Schichten mit Ziel-Daten (frühe Schichten werden oft "gefroren").
- **Fine-tuning ist effizient** und nutzt vorhandenes Wissen, birgt aber das **Risiko von "Negative Transfer"**, falls das Quellmodell nicht gut zur Zielaufgabe passt.

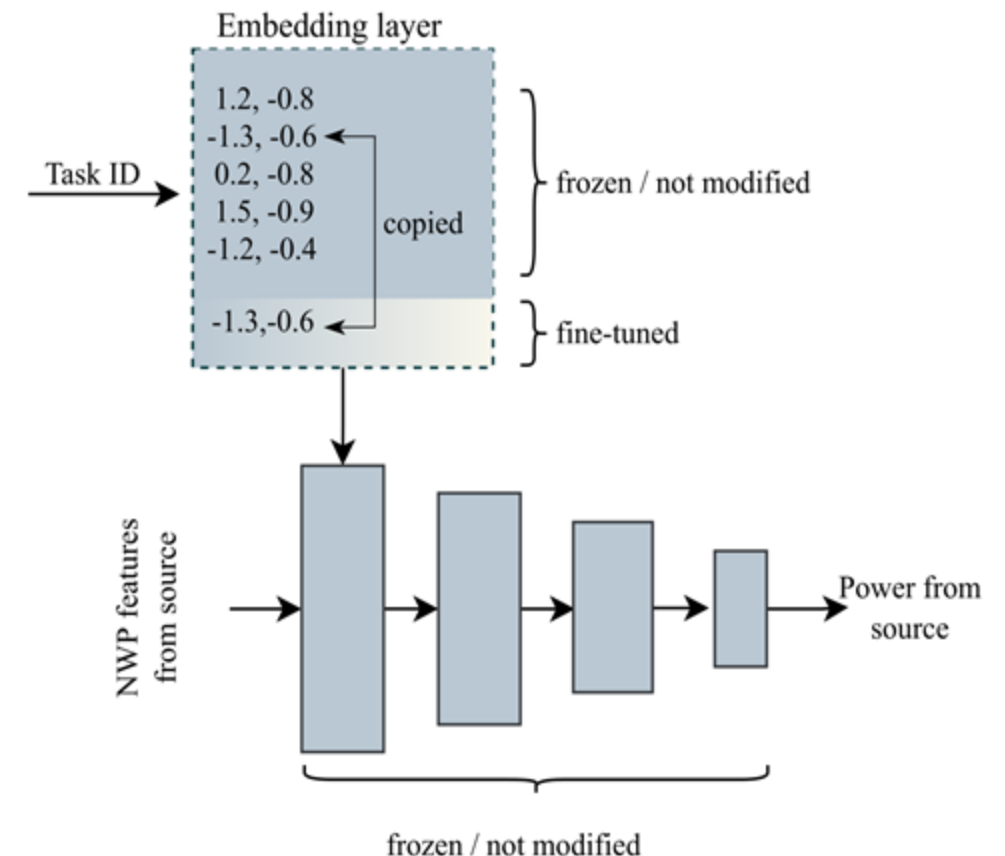
Model Anpassung durch Feature Extraction

- Die ersten Schichten des Quellmodells werden eingefroren und deren Ausgaben als Features für ein **neues Zielmodell verwendet** (z.B. lineare Regression).
- Feature Extraction **reduziert das Risiko von "Negative Transfer"**, nutzt aber das Quellwissen weniger effizient als Fine-tuning.



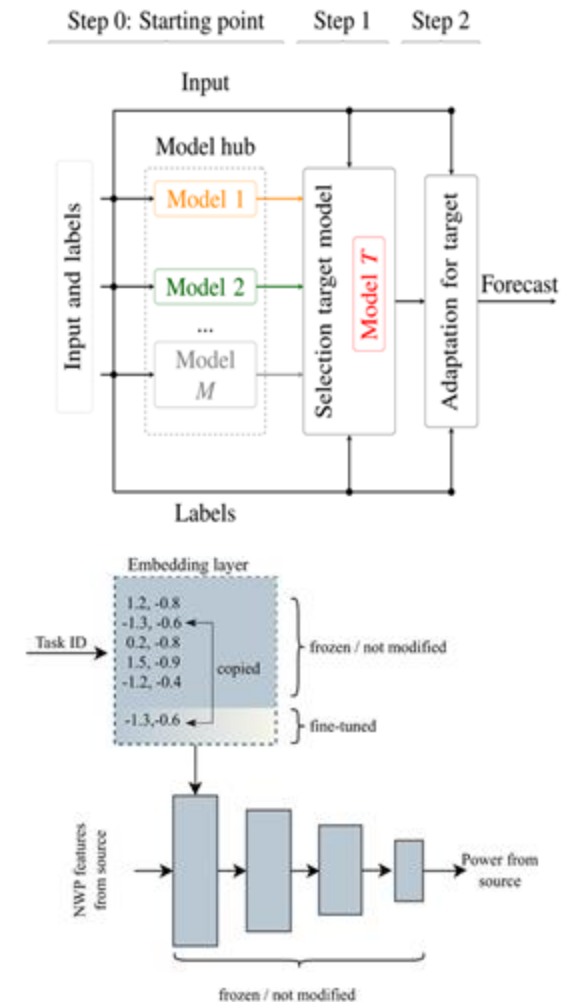
Anpassung beim Task-Embedding

- Für eine neue **Zielaufgabe** wird der **Embedding-Vektor** eines ähnlichen Quellparks **kopiert** und an die Embedding-Matrix angehängt.
- **Nur der Embedding-Vektor für die neue Aufgabe wird feinjustiert** (fine-tuning), alle anderen Modellparameter bleiben unverändert ("frozen"), wodurch das Wissen der Quellmodelle erhalten bleibt.
- Diese **Methode vermeidet "Catastrophic Forgetting"** und ermöglicht eine effiziente Erweiterung des Modells für neue Parks mit minimalem zusätzlichem Parameteraufwand.



Zusammenfassung Sequential-Transfer-Learning

- **SQTL ermöglicht zuverlässige Prognosen für neu angeschlossene Parks mit wenig historischen Daten**, indem Wissen aus bestehenden Parks genutzt wird; dies führt zu besseren Ergebnissen als physikalische Modelle.
- Die **Auswahl und Anpassung eines geeigneten Quellmodells ist entscheidend**, um die Prognose auf die Zielanlage zu spezifizieren und Überanpassung zu vermeiden.
- Modelle mit **Task Embedding** können effizient um neue Parks erweitert werden, **ohne dass zuvor gelerntes Wissen verloren geht** (Catastrophic Forgetting).
- Task Embedding und ähnliche Ansätze können bereits mit 7 Tagen Trainingsdaten die Genauigkeit von physikalischen Modellen erreichen. **Mit 60 Trainingstagen werden Ergebnisse, die sonst nur mit einem Jahr an historischen Daten erreichbar sind erzielt.**
- **Für die Praxis sind (Bayesche) lineare Modelle ein erster guter Ansatz**, da sie bereits mit 7 Tagen für Wind und 14 Tagen historische Daten für PV vergleichbare Ergebnisse wie physikalische Modelle erreichen.
- **Langfristig sollten MTL-Architekturen genutzt**, da diese skalierbar sind und mit minimalem Zusatzaufwand auf neue Parks angepasst **werden** können.

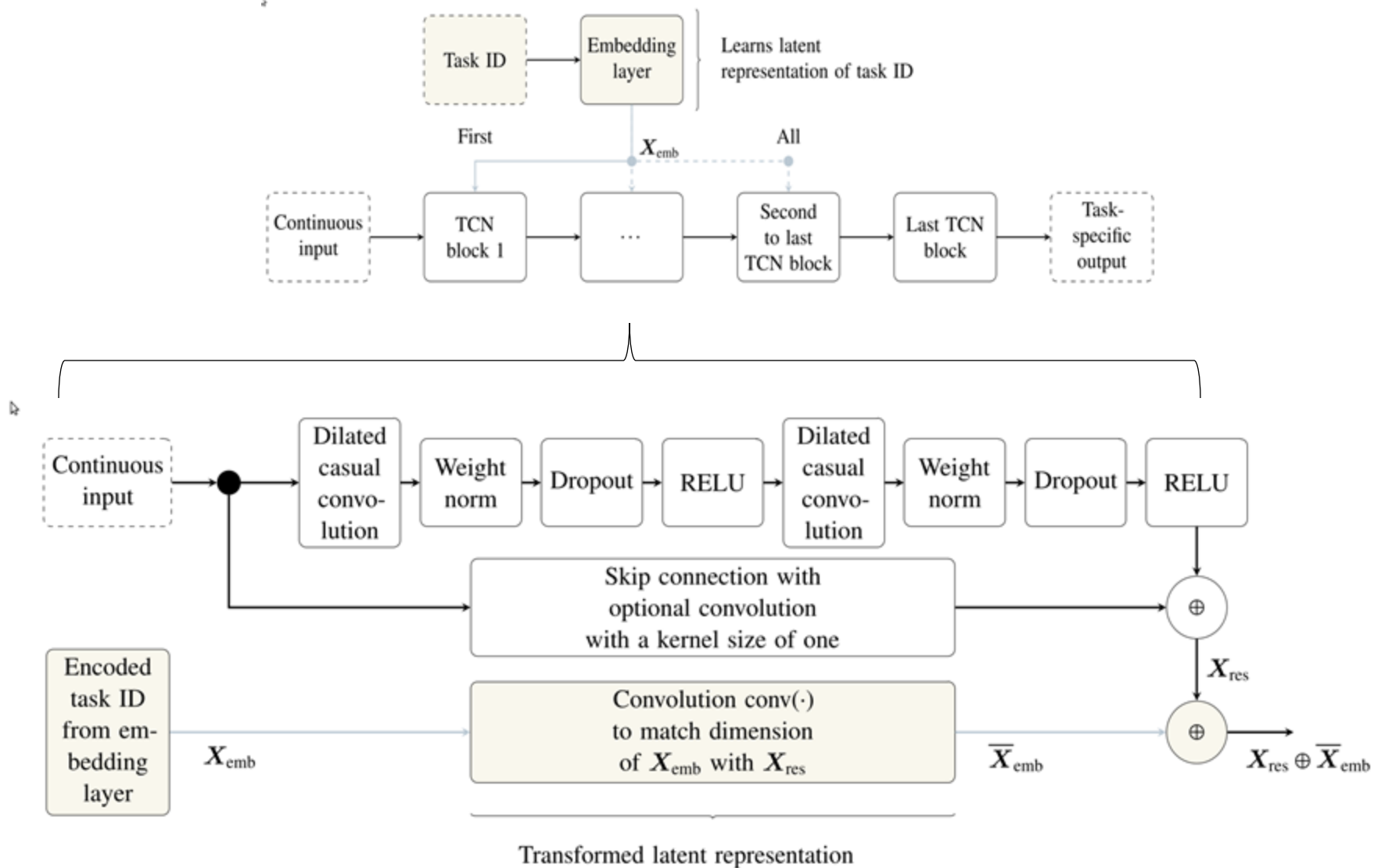


**DANKE FÜR EURE
AUFMERKSAMKEIT! QR CODE MIT
LINK ZUR PDF**



BACKUP

Task-TCN



ERGEBENISSE MULTI-TASK- LEARNING - RANK

Table 4.6: Rank summary for all STL and MTL models. GBRT is the *baseline*, and all models are tested to see if the prediction error is significantly ($\alpha = 0.01$) better (\vee), worse (\wedge), or not significantly different (\diamond). We perform this hypothesis test for all parks in a dataset. The colors indicate the rank of the models. Blue represents a lower (better) rank and ocher represents a higher (worse) rank. We highlighted the best rank of a dataset in bold.

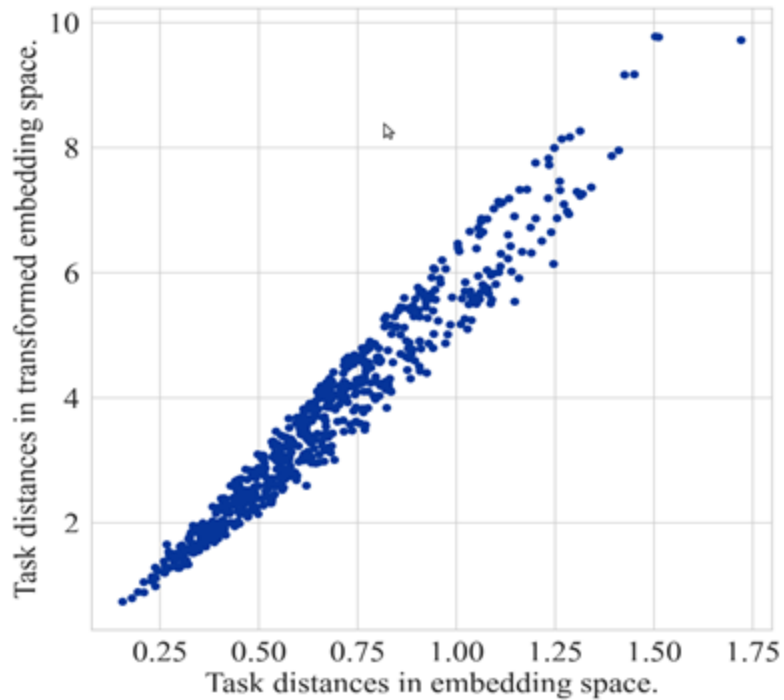
Data Type	GBRT-STL Base-line	BELM-STL	MLP-STL	TCN-STL	MLP-MTL-UNI	MLP-MTL-TASK-First	MLP-MTL-META-First	MLP-MTL-TASK-First	TCN-MTL-UNI	TCN-MTL-TASK-First	TCN-MTL-TASK-All	TCN-MTL-META-First	TCN-MTL-META-All	TCN-MTL-TASK-First	TCN-MTL-TASK-All
PVOPEN	5.190	6.429 \diamond	7.333 \wedge	6.857 \wedge	1.476\vee	2.667 \vee	-	-	5.333 \diamond	4.238 \diamond	5.476 \diamond	-	-	-	-
PVREAL	12.238	12.833 \diamond	13.024 \diamond	11.452 \diamond	11.619 \diamond	7.310 \vee	7.452 \vee	7.667 \vee	10.643 \diamond	2.619 \vee	7.024 \vee	3.833 \vee	4.143 \vee	5.571 \vee	2.571\vee
PVSYN	5.596	8.272 \wedge	7.737 \wedge	7.439 \wedge	14.149 \wedge	9.202 \wedge	9.518 \wedge	8.912 \wedge	13.667 \wedge	7.351 \wedge	4.684 \vee	5.930 \wedge	4.246\vee	5.474 \vee	5.763 \wedge
WINDOPEN	3.607	5.956 \wedge	5.822 \wedge	5.133 \wedge	6.556 \wedge	2.444\vee	-	-	6.556 \wedge	4.000 \wedge	2.889 \vee	-	-	-	-
WINDREAL	7.212	10.449 \wedge	10.400 \wedge	9.038 \wedge	13.335 \wedge	7.303 \wedge	8.027 \wedge	7.319 \wedge	11.351 \wedge	5.686 \vee	4.978 \vee	4.492 \vee	7.930 \wedge	6.341 \vee	3.908\vee
WINDSYN	8.054	10.308 \wedge	10.327 \wedge	9.588 \wedge	11.692 \wedge	3.977 \vee	9.427 \wedge	5.177 \vee	12.523 \wedge	3.396\vee	4.785 \vee	10.662 \wedge	9.938 \wedge	5.135 \vee	5.012 \vee
STL or MTL	STL	STL	STL	STL	MTL	MTL	MTL	MTL	MTL	MTL	MTL	MTL	MTL	MTL	MTL
Timeseries Model	no	no	no	yes	no	no	no	no	yes	yes	yes	yes	yes	yes	yes
Position Encoded Information	-	-	-	-	-	First	First	First	-	First	All	First	All	First	All
Categorical Input	-	-	-	-	-	Task ID	Meta Data	Task ID & Meta Data	-	Task ID	Task ID	Meta Data	Meta Data	Task ID & Meta Data	Task ID & Meta Data

ERGEBENISSE MULTI-TASK- LEARNING - NRMSE

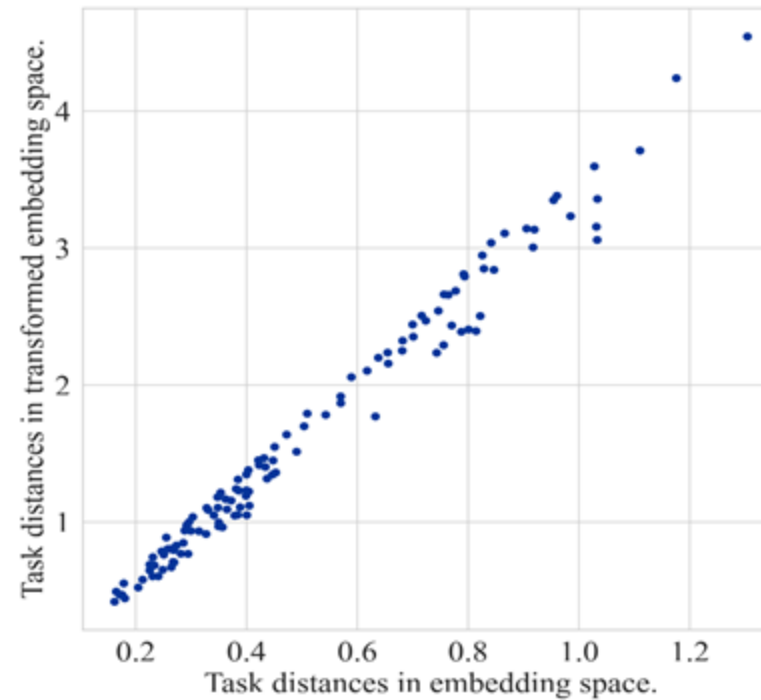
Table 4.7: Median nRMSE for all STL and MTL models. GBRT is the *baseline*, and all models are tested if the rank of the forecast error prediction error is significantly ($\alpha = 0.01$) better (\vee), worse (\wedge), or not significantly different (\diamond). We perform this hypothesis test for all parks in a dataset. The colors encode the error values. Blue represents lower (better) nRMSE and ocher represents higher (worse) nRMSE.

Data Type	GBRT-STL Base-line	BELM-STL	MLP-STL	TCN-STL	MLP-MTL-UNI	MLP-MTL-TASK-First	MLP-MTL-META-First	MLP-MTL-META-TASK-First	TCN-MTL-UNI	TCN-MTL-TASK-First	TCN-MTL-TASK-All	TCN-MTL-META-First	TCN-MTL-META-All	TCN-MTL-META-TASK-First	TCN-MTL-META-TASK-All
PVOPEN	0.065	0.066 \diamond	0.065 \wedge	0.068 \wedge	0.044 \vee	0.046 \vee	-	-	0.057 \diamond	0.055 \diamond	0.056 \diamond	-	-	-	-
PVREAL	0.090	0.091 \diamond	0.092 \diamond	0.089 \diamond	0.084 \diamond	0.079 \vee	0.080 \vee	0.080 \vee	0.085 \diamond	0.067 \vee	0.071 \vee	0.069 \vee	0.069 \vee	0.069 \vee	0.067 \vee
PVSYN	0.073	0.080 \wedge	0.079 \wedge	0.079 \wedge	0.105 \wedge	0.080 \wedge	0.080 \wedge	0.080 \wedge	0.098 \wedge	0.073 \wedge	0.072 \vee	0.073 \wedge	0.071 \vee	0.072 \vee	0.072 \wedge
WINDOPEN	0.133	0.144 \wedge	0.147 \wedge	0.142 \wedge	0.154 \wedge	0.122 \vee	-	-	0.155 \wedge	0.133 \wedge	0.127 \vee	-	-	-	-
WINDREAL	0.119	0.123 \wedge	0.123 \wedge	0.123 \wedge	0.138 \wedge	0.118 \wedge	0.119 \wedge	0.118 \wedge	0.127 \wedge	0.104 \vee	0.103 \vee	0.104 \vee	0.107 \wedge	0.105 \vee	0.104 \vee
WINDSYN	0.129	0.136 \wedge	0.138 \wedge	0.133 \wedge	0.140 \wedge	0.110 \vee	0.131 \wedge	0.110 \vee	0.153 \wedge	0.104 \vee	0.106 \vee	0.140 \wedge	0.138 \wedge	0.106 \vee	0.106 \vee
STL or MTL	STL	STL	STL	STL	MTL	MTL	MTL	MTL	MTL	MTL	MTL	MTL	MTL	MTL	MTL
Timeseries Model	no	no	no	yes	no	no	no	no	yes	yes	yes	yes	yes	yes	yes
Position Encoded Information	-	-	-	-	-	First	First	First	-	First	All	First	All	First	All
Categorical Input	-	-	-	-	-	Task ID	Meta Data	Task ID & Meta Data	-	Task ID	Task ID	Meta Data	Meta Data	Task ID & Meta Data	Task ID & Meta Data

Evaluierung Task-TCN-Embedding mit Distanzmaßen



(a) WINDOPEN dataset.



(b) PVOPEN dataset.

Evaluierung Task-TCN-Embedding mit

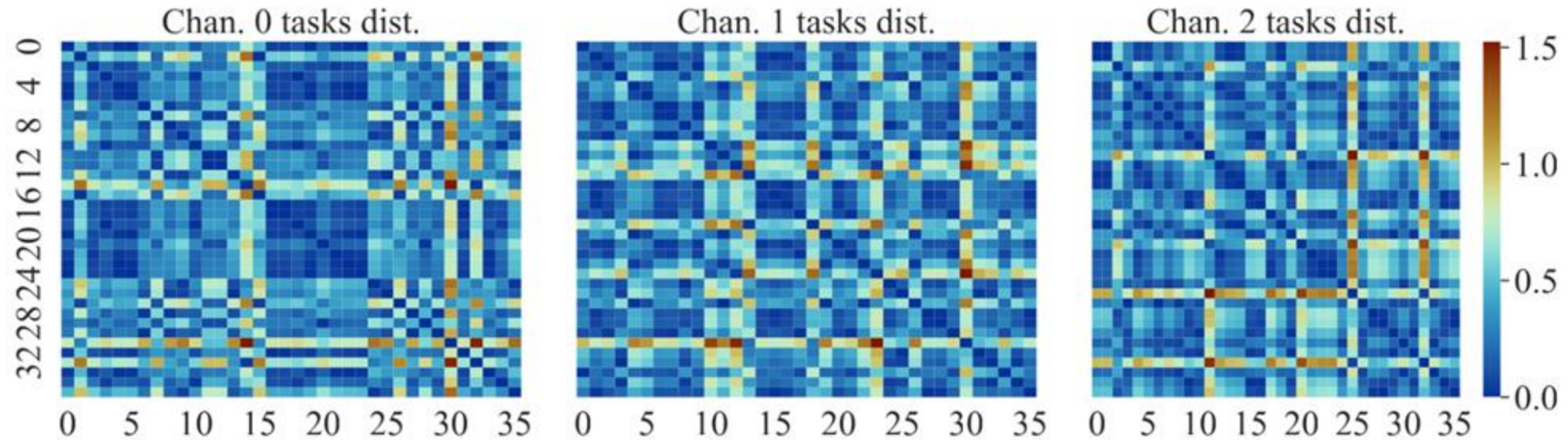
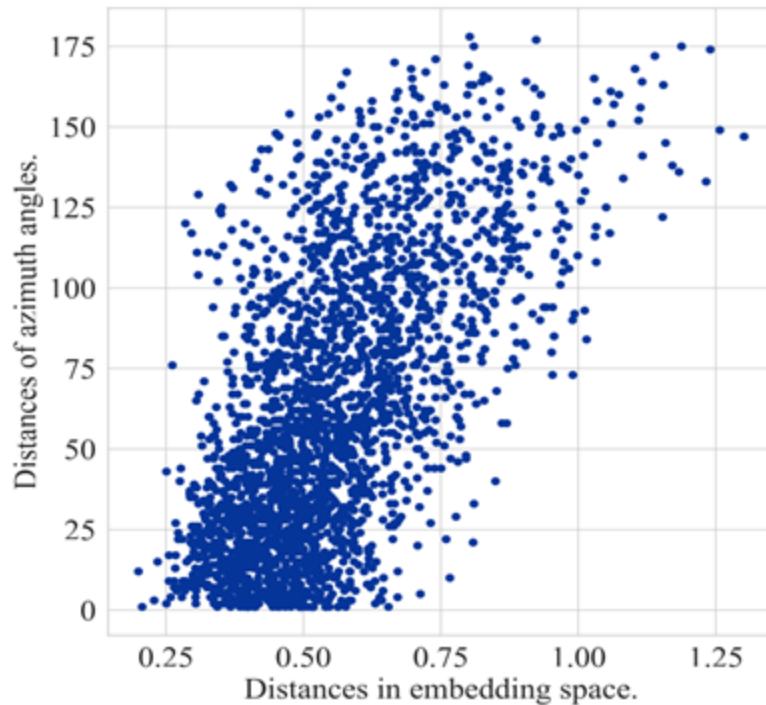
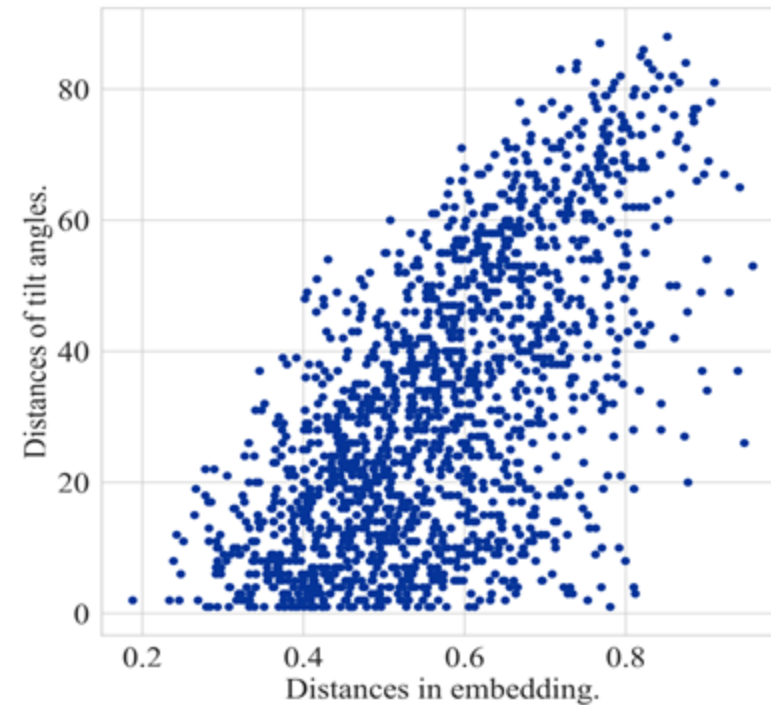


Figure 4.8: Euclidean distances of tasks in the transformed latent embedding for the WIND-OPEN dataset. Values closer to zero indicate a more substantial similarity.

Evaluierung Task-Embedding mit Physikalischen Größen



(a) Azimuth angle.



(b) Tilt angle.

Ergebnisse Sequential - Transfer - Learning - PV

Table 5.7: Rank summary for best models, selection, and adaptation strategies on the PV datasets, cf. Table 5.6. Only those that are in the top six rank of a dataset at least once are included. Those not included in the table are presented in Appendix E. Those not included in the table are presented in Appendix E. The rank in this table is calculated only on those presented here. GBRT is the *baseline*, and all models are tested if the rank of the forecast error is significantly ($\alpha = 0,01$) better (\vee), worse (\wedge), or not significantly different (\diamond). We conduct this hypothesis test on the test error for all parks in a dataset for the given number of days of training data (#days) from all four seasons. The colors denote the respective rank. Blue indicates a smaller (better) rank and ocher a higher (worse) rank. The best rank is highlighted in bold.

		GBRT-STL Baseline	BELM-STL- EV	BELM-STL- RM	MLP-STL- EV- DI	MLP-STL- EV- DILI	MLP-STL- RM- DI	TCN-STL- EV- DI	TCN-STL- EV- DILI	TCN-STL- EV- WDS	TCN-STL- RM- DI	TCN-STL- RM- WD	TCN-STL- RM- WDS
Data Type	#Days												
PVOPEN	7	7.929	4.464 \vee	5.071 \vee	7.262 \diamond	6.976 \vee	4.750 \vee	6.202 \vee	5.500 \vee	7.036 \diamond	4.107\vee	7.036 \diamond	6.869 \vee
PVREAL	7	10.357	6.357 \vee	6.208 \vee	5.685 \vee	7.095 \vee	4.863 \vee	4.940 \vee	5.536 \vee	6.321 \vee	3.679\vee	6.036 \vee	6.113 \vee
PVSYN	7	10.368	4.428 \vee	4.362 \vee	6.086 \vee	7.557 \vee	4.969 \vee	4.787 \vee	6.072 \vee	6.952 \vee	4.057\vee	6.649 \vee	6.695 \vee
PVOPEN	14	7.107	3.869\vee	3.869\vee	7.083 \diamond	6.774 \diamond	4.750 \vee	6.048 \diamond	5.940 \vee	8.060\diamond	4.202 \vee	7.726 \diamond	7.631 \diamond
PVREAL	14	9.869	6.565 \vee	6.375 \vee	5.435 \vee	6.089 \vee	4.732 \vee	5.536 \vee	4.750 \vee	6.935 \vee	3.423\vee	6.827 \vee	6.744 \vee
PVSYN	14	9.842	4.513 \vee	3.719 \vee	6.184 \vee	6.715 \vee	4.763 \vee	4.546 \vee	4.991 \vee	8.029 \vee	3.656\vee	7.662 \vee	7.616 \vee
PVOPEN	30	6.333	4.202 \vee	4.143\vee	7.595 \wedge	6.595 \diamond	5.929 \diamond	6.333 \diamond	5.702 \diamond	7.488\diamond	4.417 \vee	7.393 \diamond	7.524 \wedge
PVREAL	30	9.417	7.060 \vee	6.756 \vee	6.393 \vee	5.304 \vee	5.071 \vee	5.012 \vee	3.512\vee	6.042 \vee	3.327\vee	6.792 \vee	6.595 \vee
PVSYN	30	9.529	4.761 \vee	3.761\vee	6.673 \vee	6.197 \vee	5.189 \vee	4.939 \vee	4.816 \vee	7.660 \vee	3.976 \vee	7.116 \vee	7.265 \vee
PVOPEN	60	6.107	4.381 \vee	4.369\vee	7.869 \wedge	6.286 \diamond	5.881 \diamond	7.083 \diamond	5.143 \diamond	7.786\wedge	4.726 \vee	7.095 \diamond	7.024 \diamond
PVREAL	60	9.577	7.583 \vee	7.024 \vee	6.673 \vee	5.012 \vee	5.179 \vee	5.464 \vee	3.155\vee	6.417 \vee	3.214 \vee	6.220 \vee	6.220 \vee
PVSYN	60	9.139	4.664 \vee	3.789\vee	7.074 \vee	6.433 \vee	5.186 \vee	4.677 \vee	4.643 \vee	7.381 \vee	4.157 \vee	7.300 \vee	7.408 \vee
PVOPEN	90	5.560	4.071\vee	4.726 \diamond	8.071 \wedge	6.869 \wedge	6.155 \diamond	6.667 \diamond	5.131 \diamond	7.429 \wedge	4.786 \diamond	6.845 \wedge	7.119 \wedge
PVREAL	90	9.173	7.637 \vee	7.071 \vee	7.060 \vee	5.083 \vee	5.250 \vee	5.548 \vee	3.351\vee	6.173 \vee	3.143\vee	6.220 \vee	6.226 \vee
PVSYN	90	8.678	4.514 \vee	3.732\vee	7.111 \vee	6.386 \vee	5.546 \vee	4.835 \vee	4.516 \vee	7.308 \vee	4.395 \vee	7.305 \vee	7.324 \vee
Timeseries Model		no	no	no	no	no	no	yes	yes	yes	yes	yes	yes
Selection Strategy		Evidence	Evidence	RMSE	Evidence	Evidence	RMSE	Evidence	Evidence	Evidence	RMSE	RMSE	RMSE
Adaptation Strategy		-	Online	Online	Direct	Direct linear	Direct	Direct	Direct linear	Weight decay source	Direct	Weight decay	Weight decay source
Utilizes BLR		no	yes	yes	yes	yes	no	yes	yes	yes	no	no	no
Fine-Tuning (FT) or Feature Extraction (FE)		-	FT	FT	-	FE	-	-	FE	FT	-	FT	FT

Ergebnisse Sequential - Transfer - Learning - PV

Table 5.9: Median nRMSE for best models, selection, and adaptation strategies on the PV datasets, cf. Table 5.6. Only those that are in the top six rank of a dataset at least once are included. Those not included in the table are presented in Appendix E. GBRT is the *baseline*, and all models are tested if the rank of the forecast error is significantly ($\alpha = 0,01$) better (\vee), worse (\wedge), or not significantly different (\diamond). We conduct this hypothesis test on the test error for all parks in a dataset for the given number of days of training data (#days) from all four seasons. The rank for the hypothesis test and the color coding is calculated only on those presented here. The colors denote the respective nRMSE. Blue indicates a smaller (better) nRMSE, and ocher indicates a higher (worse) nRMSE. The best error rate is highlighted in bold.

Data Type	#Days	GBRT-STL Baseline	BELM-STL-EV	BELM-STL-RM	MLP-STL-EV-DI	MLP-STL-EV-DILI	MLP-STL-RM-DI	TCN-STL-EV-DI	TCN-STL-EV-DILI	TCN-STL-EV-WDS	TCN-STL-RM-DI	TCN-STL-RM-WD	TCN-STL-RM-WDS
PVOPEN	7	0.084	0.070 \vee	0.074 \vee	0.077 \diamond	0.080 \vee	0.071 \vee	0.080 \vee	0.077 \vee	0.081 \diamond	0.074 \vee	0.080 \diamond	0.080 \vee
PVREAL	7	0.129	0.099 \vee	0.101 \vee	0.097 \vee	0.105 \vee	0.096 \vee	0.097 \vee	0.100 \vee	0.101 \vee	0.094 \vee	0.100 \vee	0.099 \vee
PVSYN	7	0.120	0.093 \vee	0.092 \vee	0.096 \vee	0.101 \vee	0.092 \vee	0.093 \vee	0.097 \vee	0.098 \vee	0.091 \vee	0.096 \vee	0.097 \vee
PVOPEN	14	0.080	0.069 \vee	0.070 \vee	0.076 \diamond	0.078 \diamond	0.070 \vee	0.077 \diamond	0.072 \vee	0.087 \diamond	0.071 \vee	0.085 \diamond	0.084 \diamond
PVREAL	14	0.112	0.098 \vee	0.100 \vee	0.095 \vee	0.101 \vee	0.094 \vee	0.097 \vee	0.098 \vee	0.104 \vee	0.094 \vee	0.102 \vee	0.101 \vee
PVSYN	14	0.109	0.092 \vee	0.088 \vee	0.094 \vee	0.096 \vee	0.090 \vee	0.089 \vee	0.090 \vee	0.098 \vee	0.087 \vee	0.097 \vee	0.098 \vee
PVOPEN	30	0.073	0.068 \vee	0.067 \vee	0.076 \diamond	0.073 \diamond	0.071 \diamond	0.075 \diamond	0.075 \diamond	0.080 \diamond	0.071 \vee	0.080 \diamond	0.078 \wedge
PVREAL	30	0.109	0.097 \vee	0.097 \vee	0.096 \vee	0.097 \vee	0.093 \vee	0.093 \vee	0.093 \vee	0.097 \vee	0.091 \vee	0.097 \vee	0.097 \vee
PVSYN	30	0.099	0.091 \vee	0.086 \vee	0.092 \vee	0.092 \vee	0.087 \vee	0.088 \vee	0.088 \vee	0.094 \vee	0.086 \vee	0.091 \vee	0.092 \vee
PVOPEN	60	0.074	0.067 \vee	0.072 \vee	0.072 \wedge	0.071 \diamond	0.070 \diamond	0.080 \diamond	0.071 \diamond	0.080 \diamond	0.071 \vee	0.076 \diamond	0.078 \diamond
PVREAL	60	0.105	0.097 \vee	0.096 \vee	0.095 \vee	0.094 \vee	0.092 \vee	0.092 \vee	0.091 \vee	0.096 \vee	0.090 \vee	0.096 \vee	0.095 \vee
PVSYN	60	0.095	0.089 \vee	0.086 \vee	0.091 \vee	0.091 \vee	0.086 \vee	0.086 \vee	0.086 \vee	0.090 \vee	0.085 \vee	0.090 \vee	0.090 \vee
PVOPEN	90	0.071	0.066 \vee	0.072 \diamond	0.076 \wedge	0.072 \wedge	0.070 \diamond	0.077 \diamond	0.071 \diamond	0.077 \wedge	0.068 \diamond	0.076 \wedge	0.075 \wedge
PVREAL	90	0.103	0.097 \vee	0.096 \vee	0.095 \vee	0.094 \vee	0.092 \vee	0.093 \vee	0.091 \vee	0.095 \vee	0.089 \vee	0.095 \vee	0.094 \vee
PVSYN	90	0.093	0.088 \vee	0.085 \vee	0.090 \vee	0.090 \vee	0.085 \vee	0.086 \vee	0.085 \vee	0.090 \vee	0.084 \vee	0.090 \vee	0.090 \vee
Timeseries Model		no	no	no	no	no	no	yes	yes	yes	yes	yes	yes
Selection Strategy		Evidence	Evidence	RMSE	Evidence	Evidence	RMSE	Evidence	Evidence	Evidence	RMSE	RMSE	RMSE
Adaptation Strategy		-	Online	Online	Direct	Direct linear	Direct	Direct	Direct linear	Weight decay source	Direct	Weight decay	Weight decay source
Utilizes BLR		no	yes	yes	yes	yes	no	yes	yes	yes	no	no	no
Fine-Tuning (FT) or Feature Extraction (FE)		-	FT	FT	-	FE	-	-	FE	FT	-	FT	FT

Ergebnisse Sequential Transfer Learning - WIND

Table 5.8: Rank summary for best models, selection, and adaptation strategies on the wind datasets, cf. Table 5.6. Only those that are in the top six rank of a dataset at least once are included. Those not included in the table are presented in Appendix E. The rank in this table is calculated only on those presented here. GBRT is the *baseline*, and all models are tested if the rank of the forecast error is significantly ($\alpha = 0,01$) better (\vee), worse (\wedge), or not significantly different (\diamond). We conduct this hypothesis test on the test error for all parks in a dataset for the given number of days of training data (#days) from all four seasons. The colors denote the respective rank. Blue indicates a smaller (better) rank and ocher a higher (worse) rank. The best rank is highlighted in bold.

Data Type	#Days	GBRT-STL Baseline	BELM-STL-EV	BELM-STL-RM	MLP-STL-EV-DI	MLP-STL-EV-DILI	MLP-STL-RM-DI	TCN-STL-EV-DI	TCN-STL-EV-DILI	TCN-STL-EV-WDS	TCN-STL-RM-DI	TCN-STL-RM-WD	TCN-STL-RM-WDS
WINDOPEN	7	8.584	4.092 \vee	4.549 \vee	8.081 \diamond	7.376 \vee	5.555 \vee	7.509 \vee	5.792 \vee	6.509 \vee	4.873 \vee	6.445 \vee	6.272 \vee
WINDREAL	7	9.751	6.502 \vee	5.739 \vee	7.687 \vee	7.535 \vee	4.642 \vee	7.286 \vee	5.253 \vee	6.204 \vee	3.893 \vee	5.459 \vee	5.517 \vee
WINDSYN	7	9.395	5.815 \vee	4.666 \vee	7.653 \vee	7.369 \vee	5.070 \vee	7.300 \vee	5.842 \vee	6.694 \vee	4.370 \vee	5.572 \vee	5.746 \vee
WINDOPEN	14	7.676	3.925 \vee	4.688 \vee	8.306 \diamond	6.116 \vee	5.954 \vee	7.798 \diamond	5.191 \vee	7.098 \diamond	4.884 \vee	6.954 \diamond	7.052 \diamond
WINDREAL	14	8.905	5.827 \vee	5.462 \vee	8.181 \vee	6.413 \vee	4.473 \vee	7.337 \vee	4.429 \vee	7.015 \vee	3.823 \vee	6.639 \vee	6.747 \vee
WINDSYN	14	8.836	5.097 \vee	4.391 \vee	8.125 \vee	6.407 \vee	5.144 \vee	7.299 \vee	4.941 \vee	7.470 \vee	4.337 \vee	6.422 \vee	6.894 \vee
WINDOPEN	30	6.180	4.012 \vee	5.072 \vee	9.000 \wedge	5.994 \diamond	6.042 \diamond	8.216 \wedge	5.000 \vee	6.922 \diamond	5.317 \vee	6.778 \diamond	6.725 \diamond
WINDREAL	30	7.968	6.435 \vee	6.272 \vee	8.830 \wedge	5.808 \vee	4.857 \vee	7.396 \vee	4.093 \vee	6.823 \vee	4.025 \vee	6.341 \vee	6.394 \vee
WINDSYN	30	7.912	4.822 \vee	4.412 \vee	8.712 \wedge	6.118 \vee	5.537 \vee	7.793 \diamond	4.830 \vee	7.104 \vee	4.847 \vee	6.244 \vee	6.615 \vee
WINDOPEN	60	5.118	4.224 \vee	5.388 \diamond	8.941 \wedge	6.151 \wedge	6.355 \wedge	8.211 \wedge	4.954 \diamond	7.263 \wedge	5.454 \diamond	6.487 \wedge	6.572 \wedge
WINDREAL	60	6.543	6.610 \diamond	6.593 \diamond	9.142 \wedge	5.459 \vee	5.084 \vee	7.807 \wedge	4.110 \vee	6.836 \diamond	4.219 \vee	6.380 \diamond	6.436 \diamond
WINDSYN	60	7.249	4.918 \vee	4.312 \vee	8.845 \wedge	5.943 \vee	5.993 \vee	8.065 \wedge	4.844 \vee	7.093 \diamond	4.965 \vee	6.293 \vee	6.557 \vee
WINDOPEN	90	4.586	4.094 \diamond	5.773 \wedge	8.891 \wedge	6.281 \wedge	6.656 \wedge	8.180 \wedge	5.125 \diamond	7.234 \wedge	5.297 \diamond	6.180 \wedge	6.344 \wedge
WINDREAL	90	5.805	6.739 \wedge	6.647 \wedge	9.471 \wedge	5.471 \vee	5.253 \vee	7.958 \wedge	4.015 \vee	6.861 \wedge	4.322 \vee	6.446 \wedge	6.568 \wedge
WINDSYN	90	6.263	4.590 \vee	4.373 \vee	9.177 \wedge	5.700 \vee	6.127 \diamond	8.229 \wedge	4.804 \vee	7.264 \wedge	5.277 \vee	6.565 \wedge	6.723 \wedge
Timeseries Model		no	no	no	no	no	no	yes	yes	yes	yes	yes	yes
Selection Strategy		Evidence	Evidence	RMSE	Evidence	Evidence	RMSE	Evidence	Evidence	Evidence	RMSE	RMSE	RMSE
Adaptation Strategy		-	Online	Online	Direct	Direct linear	Direct	Direct	Direct linear	Weight decay source	Direct	Weight decay	Weight decay source
Utilizes BLR		no	yes	yes	yes	yes	no	yes	yes	yes	no	no	no
Fine-Tuning (FT) or Feature Extraction (FE)		-	FT	FT	-	FE	-	-	FE	FT	-	FT	FT

Ergebnisse Sequential - Transfer - Learning - WIND

Table 5.10: Median nRMSE for best models, selection, and adaptation strategies on the wind datasets, cf. Table 5.6. Only those that are in the top six rank of a dataset at least once are included. Those not included in the table are presented in Appendix E. GBRT is the *baseline*, and all models are tested if the rank of the forecast error is significantly ($\alpha = 0,01$) better (\vee), worse (\wedge), or not significantly different (\diamond). We conduct this hypothesis test on the test error for all parks in a dataset for the given number of days of training data (#days) from all four seasons. The rank for the hypothesis test and the color coding is calculated only on those presented here. The colors denote the respective nRMSE. Blue indicates a smaller (better) nRMSE, and ocher indicates a higher (worse) nRMSE. The best error rate is highlighted in bold.

Data Type	#Days	GBRT-STL Baseline	BELM-STL- EV	BELM-STL- RM	MLP-STL- EV- DI	MLP-STL- EV- DILI	MLP-STL- RM- DI	TCN-STL- EV- DI	TCN-STL- EV- DILI	TCN-STL- EV- WDS	TCN-STL- RM- DI	TCN-STL- RM- WD	TCN-STL- RM- WDS
WINDOPEN	7	0.190	0.158 \vee	0.158 \vee	0.192 \diamond	0.183 \vee	0.167 \vee	0.192 \vee	0.162 \vee	0.171 \vee	0.158 \vee	0.163 \vee	0.164 \vee
WINDREAL	7	0.181	0.151 \vee	0.145 \vee	0.163 \vee	0.158 \vee	0.140 \vee	0.162 \vee	0.145 \vee	0.150 \vee	0.138 \vee	0.144 \vee	0.146 \vee
WINDSYN	7	0.208	0.173 \vee	0.164 \vee	0.189 \vee	0.189 \vee	0.168 \vee	0.193 \vee	0.172 \vee	0.180 \vee	0.165 \vee	0.169 \vee	0.171 \vee
WINDOPEN	14	0.186	0.153 \vee	0.153 \vee	0.194 \diamond	0.169 \vee	0.159 \vee	0.192 \diamond	0.156 \vee	0.170 \diamond	0.156 \vee	0.169 \diamond	0.166 \diamond
WINDREAL	14	0.157	0.142 \vee	0.141 \vee	0.161 \vee	0.144 \vee	0.135 \vee	0.157 \vee	0.138 \vee	0.149 \vee	0.134 \vee	0.143 \vee	0.146 \vee
WINDSYN	14	0.185	0.161 \vee	0.156 \vee	0.182 \vee	0.170 \vee	0.161 \vee	0.182 \vee	0.163 \vee	0.176 \vee	0.159 \vee	0.168 \vee	0.170 \vee
WINDOPEN	30	0.160	0.147 \vee	0.150 \vee	0.196 \wedge	0.159 \diamond	0.157 \diamond	0.179 \wedge	0.151 \vee	0.164 \diamond	0.149 \vee	0.166 \diamond	0.161 \diamond
WINDREAL	30	0.145	0.141 \vee	0.139 \vee	0.159 \wedge	0.136 \vee	0.133 \vee	0.151 \vee	0.133 \vee	0.140 \vee	0.131 \vee	0.137 \vee	0.136 \vee
WINDSYN	30	0.166	0.155 \vee	0.151 \vee	0.177 \wedge	0.160 \vee	0.156 \vee	0.176 \diamond	0.153 \vee	0.162 \vee	0.153 \vee	0.159 \vee	0.160 \vee
WINDOPEN	60	0.148	0.145 \vee	0.146 \diamond	0.187 \wedge	0.154 \wedge	0.153 \wedge	0.170 \wedge	0.147 \diamond	0.160 \wedge	0.147 \diamond	0.153 \wedge	0.153 \wedge
WINDREAL	60	0.137	0.137 \diamond	0.137 \diamond	0.161 \wedge	0.134 \vee	0.133 \vee	0.151 \wedge	0.132 \vee	0.138 \diamond	0.132 \vee	0.136 \diamond	0.137 \diamond
WINDSYN	60	0.161	0.154 \vee	0.149 \vee	0.173 \wedge	0.156 \vee	0.154 \vee	0.173 \wedge	0.151 \vee	0.158 \diamond	0.150 \vee	0.154 \vee	0.155 \vee
WINDOPEN	90	0.141	0.140 \diamond	0.144 \wedge	0.179 \wedge	0.150 \wedge	0.149 \wedge	0.165 \wedge	0.142 \diamond	0.149 \wedge	0.143 \diamond	0.146 \wedge	0.144 \wedge
WINDREAL	90	0.134	0.136 \wedge	0.137 \wedge	0.158 \wedge	0.133 \diamond	0.133 \vee	0.147 \wedge	0.132 \vee	0.136 \wedge	0.131 \vee	0.136 \wedge	0.136 \wedge
WINDSYN	90	0.155	0.150 \vee	0.147 \vee	0.168 \wedge	0.152 \vee	0.152 \vee	0.169 \wedge	0.148 \vee	0.156 \wedge	0.150 \vee	0.152 \wedge	0.154 \wedge
Timeseries Model		no	no	no	no	no	no	yes	yes	yes	yes	yes	yes
Selection Strategy		Evidence	Evidence	RMSE	Evidence	Evidence	RMSE	Evidence	Evidence	Evidence	RMSE	RMSE	RMSE
Adaptation Strategy		-	Online	Online	Direct	Direct linear	Direct	Direct	Direct linear	Weight decay source	Direct	Weight decay	Weight decay source
Utilizes BLR		no	yes	yes	yes	yes	no	yes	yes	yes	no	no	no
Fine-Tuning (FT) or Feature Extraction (FE)		-	FT	FT	-	FE	-	-	FE	FT	-	FT	FT

Zero-Shot-Learning

