

Anwendung von Modellen des maschinellen Lernens für die Stromtragfähigkeitsberechnung eines 400 kV-Kabelsystems und deren Validierung

Florian Ainhirn

Wiener Netze GmbH, Erdbergstraße 236, 1110 Wien, florian.ainhirn@wienernetze.at,
<https://orcid.org/0000-0002-8983-1657>

Kurzfassung: Dieser Beitrag fasst die Untersuchungen der Wiener Netze zum Einsatz von künstlicher Intelligenz in Form des maschinellen Lernens als Alternative zu analytischen und numerischen Methoden für die dynamische Stromtragfähigkeitsberechnung eines 400 kV-Kabelsystems zusammen. Die Untersuchungen zeigen, dass mit einer geeigneten Validierungsmethodik auch simple Algorithmen wie Entscheidungsbäume für gute Vorhersagen der Kabeltemperatur der untersuchten 400 kV-Kabelanordnung verwendet werden können. Weiters sind komplexere Modelle wie rekurrente neuronale Netzwerke für die Zeitreihenmodellierung in der Lage, die Kabeltemperaturen für verschiedene Zeithorizonte zwischen einer und vierundzwanzig Stunden mit hoher Genauigkeit zu prognostizieren.

Keywords: Hochspannungskabel, Maschinelles Lernen, Stromtragfähigkeitsberechnung, Modellvalidierung

1 Einleitung

Die Schaffung ausreichender Übertragungskapazitäten im Stromnetz ist nicht nur eine technische, sondern auch eine wirtschaftliche und soziale Anforderung. Um die Energie- und Klimawende in der Europäischen Union (EU) zu bewerkstelligen, fordert die EU die effiziente Nutzung neuer und bestehender Netzinfrastrukturen zu erhöhen [1]. Dies gilt insbesondere für Hoch- und Höchstspannungskabel¹ als bestehende und zunehmende Netzkomponenten, besonders aufgrund ihrer dominierenden Rolle bei der Offshore-Windkraft und als Verbindungsleitungen in Form von Interkonnektoren. Eine der Möglichkeiten, die von der EU als entscheidend für die Erzielung von Effizienzsteigerungen identifiziert wurde, ist der Einsatz der dynamischen Stromtragfähigkeitsberechnung von Energieleitungen [2]. Die Vorteile der dynamischen Stromtragfähigkeitsberechnung für den Netzbetreiber sind dabei die Erhöhung der Übertragungskapazität, eine Steigerung der Betriebssicherheit sowie ein ökonomischer Vorteil durch die Option, Investitionen zum erforderlichen Zeitpunkt gezielt tätigen zu können.

Die etablierte analytische Methode zur thermischen Bemessung von Hochspannungskabeln für stationäre und dynamische Lasten wird durch die IEC-Normen [3 - 5] vorgegeben. Aufgrund der notwendigen Annahmen, die im Zuge der Berechnung getroffen werden müssen, kann es jedoch schwierig sein, konsistente, vergleichbare und damit vertrauenswürdige Ergebnisse zu erhalten. Dieser Umstand hat die CIGRE veranlasst, die Technische Broschüre 880 zu erstellen, welche die Grundlage für die Validierung analytischer Stromtragfähigkeitsberechnungsmethoden für Hochspannungskabel im Jahr 2022 bildet [6].

¹ Im weiteren Text werden Hoch- und Höchstspannungskabel für eine bessere Lesbarkeit als "Hochspannungskabel" zusammengefasst.

Die Finite Elemente Methode (FEM) ist seit mehr als einem Jahrzehnt eine frequent genutzte Alternative bei der thermischen Bemessung von Hochspannungskabeln. Dennoch gibt es derzeit weder Normen noch umfangreiche Guidelines oder Validierungsverfahren für die Anwendung dieser numerischen Methode. Mit [7] existiert seit 2003 ein technischer Bericht, der sich mit der FEM in der Stromtragfähigkeitsberechnung von Energiekabeln beschäftigt. Dieser dient aber ausschließlich als eine Erweiterung der analytischen Methode in [3] und stellt keine vollwertige Behandlung der numerischen Methode dar.

Die Anwendung von künstlicher Intelligenz (KI) in den Energienetzen wird bereits seit den 1990ern untersucht, beispielweise [8 - 10]. Solche datengetriebenen Modelle auf der Grundlage des maschinellen Lernens (ML) stellen eine vielversprechende Alternative für die dynamische Stromtragfähigkeitsberechnung von Hochspannungskabeln dar. Mit der Zunahme verteilter Temperaturmesssysteme (DTS) an Hochspannungskabeln, sowie der Schaffung von internen Data Lakes und den erforderlichen Server- und Cloudlösungen, finden datengetriebene Modelle jedoch erst in jüngster Zeit die nötigen Rahmenbedingungen für eine solche Anwendung auch bei den Netzbetreibern vor.

Wie bei jeder Technologie mit geringen Erfahrungswerten, müssen diese Methoden erst ordnungsgemäß verifiziert und validiert werden, bevor sie im Stromnetzbetrieb eingesetzt werden können. Aus diesem Grund wurden von der Wiener Netze GmbH mögliche KI-Anwendungen für die thermische Bewertung und Überwachung von Hochspannungskabeln untersucht. Zu diesem Zweck wurde unter anderem ein 400 kV-Kabelversuch errichtet, an welchem unter realen Bedingungen verschiedenste stationäre und dynamische Lasten aufgebracht wurden. Über einen Zeitraum von mehr als drei Jahren wurden unterschiedlichste Parameter wie Kabel- und Bodentemperaturen, Bodenfeuchte und Wetterdaten gemessen und aufgezeichnet. Die Daten wurden verwendet, um datengetriebene Modelle für die dynamische Stromtragfähigkeitsberechnung des 400 kV-Kabelsystems zu erstellen, zu verifizieren und zu validieren.

2 400 kV - Kabelversuch

Die Kabelversuchsanordnung ist in den beiden Abbildungen **Abb. 1** und **Abb. 2** dargestellt.

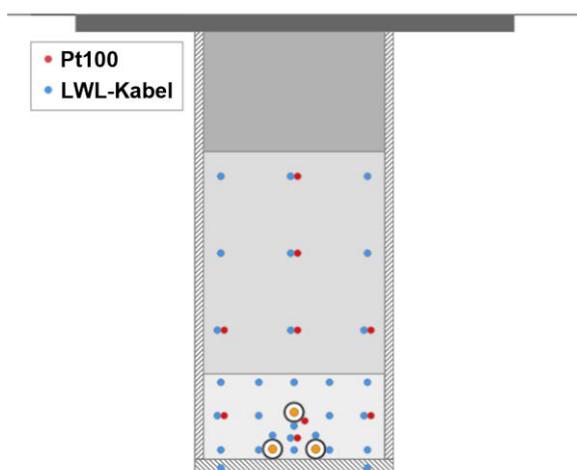


Abbildung 1: 400 kV-Kabelsystem- und Sensorenanordnung in der Kabelkүнette

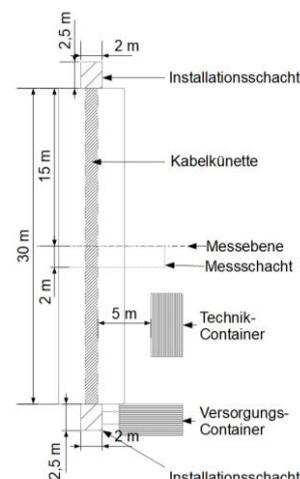


Abbildung 2: Grundriss des 400 kV - Kabelversuchs

Die Kabelversuchsanordnung besteht aus drei VPE-isolierten 400 kV – Kabeln mit je 2500 mm² Leiterquerschnitt, die als dreiphasiges Kabelsystem mit einem Phasenabstand von 270 mm in einer Dreiecksanordnung in einem thermisch stabilisierten Betonblock mit einer Breite von 1200 mm und einer Höhe von 600 mm im Erdreich verlegt sind. Der thermisch stabilisierte Betonblock ist auf einer 100 mm Sauberkeitsschicht aufgesetzt.

In der Mitte des Versuchsaufbaus befindet sich eine Messebene, in welcher über 90 Sensoren installiert sind, um neben den elektrischen Betriebsgrößen auch Kabel- und Bodentemperaturen, Bodenfeuchten sowie Klimadaten zu erfassen. Dabei wurden Temperaturen im Nahbereich des Kabelsystems als auch in mehreren Metern Entfernung zum Kabel gemessen, um das Profil des thermisch unbeeinflussten Erdreichs zu erfassen.

Um Daten über das thermische Verhalten des 400 kV-Kabelsystems zu erhalten, wurden über einen Zeitraum von 3 Jahren verschiedene Belastungsszenarien an dem 400 kV-Kabelversuch durchgeführt. Die Lastszenarien wurden hierbei so gewählt, dass diese mit den in den Normen [3 - 5] zu Grunde gelegten Belastungen vergleichbar sind. Zu diesem Zweck wurden die Lasten in stationäre Belastungen, d.h. Szenarien mit konstantem Strom über ausgedehnte Zeiträume (mehr als 90 Tage), und dynamische Belastungen in Form von zyklischen Strombeaufschlagungen und Stufentests unterteilt. Ein Beispiel für die durchgeführten stationären Belastungen mit einem abschließenden 1700 A Stufentest ist in **Abb. 3** dargestellt. Ein Auszug aus den durchgeführten dynamischen Belastungen ist in **Abb. 4** zu sehen.

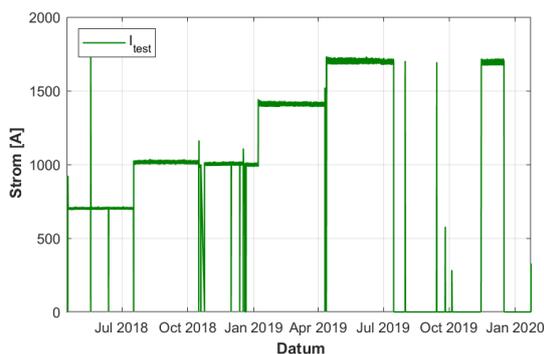


Abbildung 3: Auszug der durchgeführten stationären Belastungsszenarien

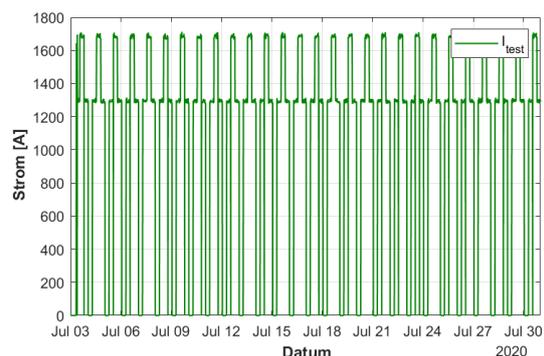


Abbildung 4: Auszug der durchgeführten dynamischen Belastungsszenarien

Für detailliertere Informationen zu dem 400 kV-Kabelversuchsaufbau, wie verwendetes Hochspannungskabel, Verlegeanordnung, installierte Sensorik, verwendete Messgeräte, untersuchte Lastprofile, etc. wird an dieser Stelle auf [11] und [12] verwiesen.

3 Daten

3.1 Datenerfassung

Die Anwendungen des maschinellen Lernens beruhen auf Daten. Daten werden benötigt, um die Modelle zu erstellen und zu trainieren, die Modelle zu validieren und zu testen, und selbst nachdem all diese Schritte abgeschlossen sind, benötigt ein ML-Modell kontinuierliche Eingabedaten zur Interpretation und Vorhersage. Dabei können Daten aus einer Vielzahl unterschiedlicher Quellen gewonnen werden. Die Datenquellen und Datentypen, die in diesem Projekt verwendet wurden, sind in **Tabelle 1** quantitativ angeführt.

Tabelle 1: Quellen und Arten der verwendeten Daten

Art der Datenquelle		Datentyp	Anzahl an Datenpunkten
Wiener Netze (Kabelbetreiber)		DTS - Temperaturdaten, Belastungsdaten (Strom), Bodentemperatur und Feuchtedaten, Klimadaten	> 70 Mio.
Extern	Behördlich	Hydrologische Daten, Klimadaten	225000
		Bodenkompositionsdaten	1396
	Öffentlich	Klimadaten	22320
	Kommerziell	Boden- und Lufttemperaturdaten	315648

3.2 Datenvorbereitung und Optimierung

Nachdem die Daten aus einer oder mehreren Quellen gewonnen wurden, sind diese Rohdaten in den meisten Fällen nicht direkt für ML-Anwendungen verwendbar. Der Hauptgrund dafür ist, dass Daten, die bestimmte Parameter enthalten, in der Regel für einen spezifischen Verwendungsfall aufgezeichnet werden und daher das Format und die Form der Daten, wie z.B. Abtastrate, Auflösungsgenauigkeit, Beschriftung usw., für diesen Fall angepasst sind. Weiters sind Daten aus realen Quellen praktisch immer mangelhaft, da diese beispielsweise Datenlücken, Offsets und Rauschen enthalten. Daher können mehrere Aufbereitungsschritte erforderlich sein, um die Daten in eine geeignete Form zu bringen, die von ML-Algorithmen und Merkmalsauswahlverfahren verwendet werden kann. Für Details zu den verwendeten Aufbereitungs- und Optimierungsschritten sei hier auf [12] verwiesen.

3.3 Datenanalyse

Bei der Datenanalyse geht es unter anderem darum, mit Methoden der deskriptiven Statistik, Informationen über die Beziehungen zwischen den Daten zu erhalten. Dieser Schritt ist beim datengetriebenen Modellierungsansatz von entscheidender Bedeutung, da eine große Anzahl von Eingangsparametern zu deutlich längeren Trainingszeiten der Modelle führt. Darüber hinaus kann die Verwendung zu vieler Merkmale die Vorhersageleistung verschlechtern, selbst wenn alle Merkmale Informationen über die Antwortvariable (Ausgangsparameter, hier die Kabeltemperatur) enthalten. Um die Beziehung zwischen den verschiedenen Merkmalen auszudrücken, können die Korrelationskoeffizienten bestimmt und in einer Matrix angeordnet werden, wie in **Abb. 5** für Umweltparameter, Strom und Kabeltemperatur exemplarisch gezeigt. Die Korrelationskoeffizienten ergeben Werte zwischen +1, was eine perfekte positive Korrelation bedeutet, und -1, was eine perfekte negative Korrelation bedeutet. Werte zwischen 0,3 und -0,3 können als unwesentlich korreliert angesehen werden, während Werte von 0 bedeuten, dass überhaupt kein linearer Zusammenhang zwischen den Merkmalen besteht. In der Diagonalen der Korrelationsmatrix nehmen die Elemente den Wert 1 an, da jedes Merkmal perfekt mit sich selbst korreliert. Eine Liste mit den Kürzeln für die verschiedenen Parameter, die für die dargestellte Parameteranalyse verwendet wurden, findet sich in **Tabelle 2**. Betrachtet man die letzte Zeile der Korrelationsmatrix in **Abb. 5**, so lässt sich für die untersuchte Kabelinstallation gut erkennen, dass beispielsweise der Strom stark positiv, die Bodenfeuchte stark negativ oder die Bodentemperatur in der Tiefe des Kabels stark positiv mit der Kabeltemperatur korreliert.

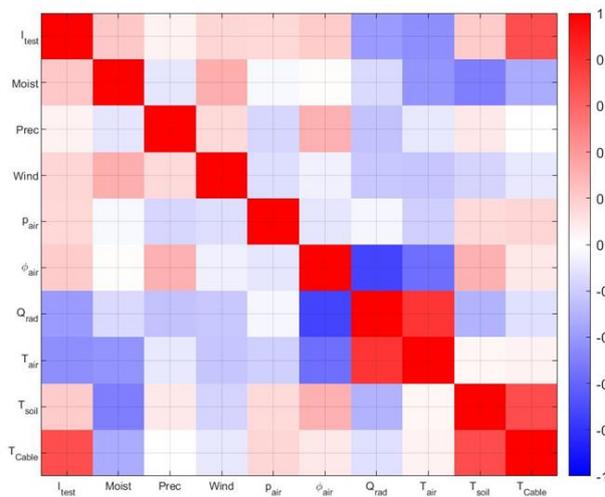


Abbildung 5: Korrelationsmatrix der Parameter

Tabelle 2: Kürzel für die Parameteranalyse

Kürzel	Beschreibung	Einheit
I _{test}	Prüfstrom	[A]
Moist	Bodenfeuchte	[%]
Prec	Niederschlag	[mm/h]
Wind	Windgeschwindigkeit	[m/s]
p _{air}	Absoluter Luftdruck	[hPa]
φ _{air}	Relative Luftfeuchtigkeit	[%]
Q _{rad}	Globalstrahlung	[W/m ²]
T _{air}	Lufttemperatur	[°C]
T _{soil}	Bodentemperatur	[°C]
T _{cable}	Kabeltemperatur	[°C]

Das Entscheidende hierbei ist, dass all diese Beziehungen zwischen den verschiedenen Parametern ausschließlich aus den Daten allein, ohne vorheriges Domänenwissen und vollständig automatisiert ermittelt werden. Dies ist einer der Hauptvorteile eines datengetriebenen Modellierungsansatzes. Eine ausführlichere Diskussion und Erläuterung der Beziehung zwischen den Parametern findet sich in [12 - 14].

3.4 Auswahl und Entwicklung von Merkmalen

Die dargestellte Datenanalyse und die Berechnung der Korrelationskoeffizienten ist Teil der sogenannten Merkmalsauswahl, d.h. der Vorauswahl von Eingangsparametern mit hoher statistischer Relevanz für die Vorhersagen. Neben dem Korrelationskoeffizienten gibt es eine Vielzahl von Methoden, beispielsweise Filter-, Wrapper- oder eingebettete Methoden, um solche Merkmale zu finden und zu bewerten. Der Hauptunterschied zwischen ihnen ist der Umfang, in dem die Modellarchitektur einbezogen wird. Der Merkmalsraum ist hierbei nicht nur auf Messdaten beschränkt. Selbst wenn es verschiedene Merkmale mit einem hohen Korrelationswert gibt, was auf eine signifikante Relevanz hinweist, kann es zwischen ihnen eine gewisse Redundanz geben. Das Problem dabei ist, dass redundante Merkmale den Lern- bzw. Trainingsprozess erschweren und zu falschen Vorhersagen führen können. Aus dem ursprünglichen Merkmalsraum können alternativ neue Merkmale generiert werden, die in hohem Maße mit der Zielgröße (Ausgangsparameter) in Verbindung stehen und gleichzeitig möglichst wenig mit den anderen Merkmalen zu tun haben. Dies kann mit Hilfe spezieller Algorithmen oder durch Anwendung von Domänenwissen geschehen. Beispielsweise können viele Regressionsmodelle, wie Entscheidungsbäume, Support Vector Machines oder neuronale Netzwerke mit Vorwärtskopplung (Feed Forward Neural Network), ihre Gegenwartsvorhersage auf der Grundlage der Werte der zu diesem Zeitpunkt verfügbaren Merkmale erzeugen. Die thermischen Vorgänge in der Energiekabelanlage hängen jedoch unter anderem vom zeitlichen Einfluss z.B. des Stroms ab und nicht nur von dessen Momentanwert. Um dies zu berücksichtigen, können verzögerte Merkmale, wie in **Tabelle 3** dargestellt, erstellt und verwendet werden. Diese verzögerten Merkmale (lagged features) bringen nicht nur eine zeitliche Komponente in das Modell ein, sondern auch eine gewisse Kausalität.

4 Anwendung von Modellen des maschinellen Lernens

4.1 Regressionsmodelle für Kabeltemperatur (Nowcasting)

Das Hauptaugenmerk dieser Untersuchungen lag auf der Klärung der Frage, ob maschinelle Lernmodelle zur Bestimmung der Kabeltemperatur eingesetzt werden können und wenn ja, mit welcher Genauigkeit und unter welchen Umständen. Zu diesem Zweck wurden mehrere datengetriebene Modelle mit dem Ziel erstellt, die Kabeltemperatur an der Kabeloberfläche zu bestimmen. Es wurden die notwendigen Merkmale, ihr Einfluss auf die Vorhersageergebnisse und verschiedene Metriken der Modellleistung bewertet. Für alle Modelle wurde der komplette Datensatz mit allen Belastungsszenarien und einer Abtastrate von zwei Minuten verwendet, wobei das Ziel immer das DTS-Messergebnis der Glasfaser in der Mitte der Kabelanordnung war. Zu diesem Zweck wurden drei Algorithmen zur Gegenwartsvorhersage (Nowcasting) der Kabeloberflächentemperatur verwendet, nämlich Entscheidungsbaum (DT), Support Vector Machine (SVM) und Feed Forward Neural Network (FFNN). Die Modelle wurden dann mit einem Trainings-, Test- und Validierungsdatensatz trainiert und validiert, wobei der Trainings-, Validierungs- und Testdatensatz im Verhältnis 40/40/20 aufgeteilt wurde. Dabei wurden Merkmale und Modellparameter, die so genannten Hyperparameter, variiert und die Genauigkeit der Modelle überwacht. Ein Beispiel für die so erzielten Ergebnisse ist in **Abb. 6** dargestellt, wo die Kabeltemperatur während des Übergangs von zwei dynamischen Belastungen mit einer SVM bestimmt und mit den DTS-Messwerten verglichen wurde.

Tabelle 3: Zeitlich verzögerte Merkmale

Merkmal	Basis	Beschreibung
I_{lag1h}	I_{test}	Strom vor 1 Stunde
I_{lag6h}		Strom vor 6 Stunden
I_{lag12h}		Strom vor 12 Stunden
I_{lag24h}		Strom vor 24 Stunden
T_{lag1h}	T_{cable}	Kabeltemperatur vor 1 Stunde
T_{lag6h}		Kabeltemperatur vor 6 Stunden
T_{lag12h}		Kabeltemperatur vor 12 Stunden
T_{lag24h}		Kabeltemperatur vor 24 Stunden

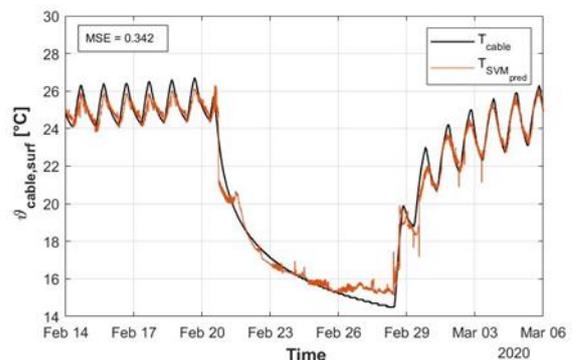


Abbildung 6: Temperaturprognose des DT - Modells für einen dynamischen Lastwechsel

Ein Vergleich zwischen DT- und SVM-Modell ist in **Tabelle 4** angeführt. Diese zeigt, dass das DT-Modell mit nur drei Merkmalen einen RMSE (Wurzel der mittlere Fehlerquadratsumme) von weniger als 1 K erreicht, während für das SVM-Modell doppelt so viele Merkmale erforderlich sind. Obwohl hier nicht dargestellt, werden die Vorhersagen des DT-Modells auch von größeren Ausreißern im Vergleich überschattet.

Tabelle 4: Vergleich Kreuzvalidierungsmetriken für Decision Tree und Support Vector Machine Modell

Merkmale (Features)		Kreuzvalidierungsergebnisse					
		DT			SVM		
Anzahl	Beschreibung	RMSE	R^2_{test}	R^2_{val}	RMSE	R^2_{test}	R^2_{val}
3	$I_{test}, T_{soil1}, M_{soil}$	0.73	98.85	98.88	3.01	84.43	84.13
4	$I_{test}, T_{soil1}, T_{soil2}, M_{soil}$	0.41	99.65	99.65	2.08	92.23	92.81
5	$I_{test}, T_{soil1}, T_{soil2}, M_{soil}, I_{lag6h}$	0.27	99.85	99.85	1.68	95.03	95.01
6	$I_{test}, T_{soil1}, T_{soil2}, M_{soil}, I_{lag6h}, T_{lag24h}$	0.14	99.96	99.96	0.34	99.69	99.72

4.2 Zeitreihen- und Prognosemodelle

4.2.1 Zeitreihenmodelle

Die bisher betrachteten Modelle sind so genannte statische Modelle. Das bedeutet, dass der vorhergesagte Output zu einem bestimmten Zeitpunkt t ausschließlich von den Eingangsgrößen zu diesem Zeitpunkt t abhängt. Die Zeitabhängigkeit des Prozesses wurde teilweise in diese Modelle integriert, indem verzögerte Merkmale (**Tabelle 3**) einbezogen wurden. Da das thermische Verhalten eines Kabelsystems jedoch eine ausgeprägte zeitliche Komponente aufweist, sind für diesen Fall sogenannte Zeitreihenmodelle geeignet. Dabei handelt es sich um dynamische Modelle, die sowohl vergangene Werte der Eingangsgrößen als auch ihre eigenen Prognosen berücksichtigen. Zu diesem Zweck wurde ein Zeitreihenmodell auf der Basis eines nichtlinearen autoregressiven exogenen neuronalen Netzwerks (NARX NN) verwendet, dessen grundlegende Architektur in **Abb. 7** dargestellt ist. Die Ergebnisse, die mit dem NARX NN, einer Form des rekurrenten neuronalen Netzwerks (RNN), für denselben dynamischen Lastübergang erzielt wurden, sind in **Abb. 8** dargestellt.

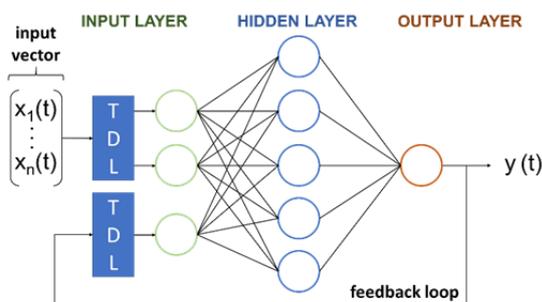


Abbildung 7: Grundlegende Architektur des NARX neuronalen Netzwerkmodells

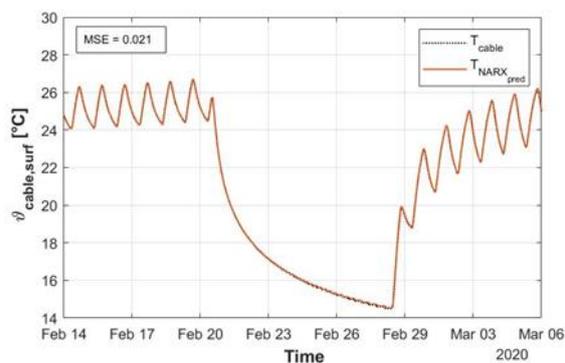


Abbildung 8: Temperaturprognose des NARX NN Modells für einen dynamischen Lastwechsel

Ähnlich wie bei den statischen Modellen wurde auch das dynamische Zeitreihenmodell mit 40 % der Daten trainiert. Da in diesem Fall jedoch die zeitliche Komponente berücksichtigt wird, kann der Trainingsdatensatz nicht über den gesamten Datenraum verteilt werden. Stattdessen wurde eine zeitliche Abfolge der Messdaten verwendet, genauer gesagt die ersten drei stationären Lastversuche mit 700 A, 1000 A und 1400 A (siehe **Abb. 3**), da diese 39,22 % der Gesamtdaten ausmachen. Wie in **Abb. 8** zu sehen ist, weist die Vorhersage des NARX NN-Modells keine Ausreißer und eine sehr hohe Genauigkeit auf. Diese Verbesserung zeigt sich insbesondere bei den dynamischen Lastwechseln und deren Übergängen.

4.2.2 Vorhersagemodelle (Intraday-Forecasting)

Auch für die Vorhersage zukünftiger Ereignisse wird das maschinelle Lernen seit Jahren intensiv genutzt, zum Beispiel zur Ermittlung von Aktienkursen oder Wettervorhersagen. Die Frage, die es zu beantworten gilt, ist jedoch, welche Algorithmen geeignet sind. Viele Algorithmen für derartige Vorhersagen basieren auf dem Deep Learning mit entsprechend komplexen Modellstrukturen. Es ist aber auch bekannt, dass Deep-Learning-Modelle, wie zum Beispiel Long-Short-Term-Memory-Modelle (LSTM), für viele Prognoseaufgaben nicht unbedingt Vorteile bringen. Vielmehr kann ihr Einsatz sogar kontraproduktiv sein, da einfachere Modelle manchmal besser abschneiden, da der für die Vorhersage erforderliche

Informationsgehalt in einigen Fällen viel geringer ist, als oft intuitiv angenommen wird [15]. In dieser Arbeit wurde ein direkter Vorhersageansatz verwendet. Dabei werden mehrere Einzelmodelle verwendet, um die Kabeltemperatur für verschiedene Prognosehorizonte vorherzusagen. In diesem Fall wurden fünf DT-Modelle und fünf RNN-Modelle verwendet, um die Kabeltemperatur für die Prognosehorizonte 1 Stunde, 3 Stunden, 6 Stunden, 12 Stunden und 24 Stunden vorherzusagen, da diese für die dynamische thermische Bewertung von Hochspannungskabeln, dem sogenannten Intraday-Forecasting, am wichtigsten sind.

5 Modelauswahl und Validierung

5.1.1 Load-Fold-Validierungsprinzip

Wie im vorangegangenen Abschnitt gezeigt wurde, sind ML-Modelle für die Vorhersage der Kabeltemperatur sehr gut geeignet. Die hohe Anpassungsfähigkeit dieser Modelle geht jedoch mit einem Mangel an Robustheit im Vergleich zu z.B. numerischen Modellen einher, die (im Idealfall nur) auf physikalischen Grundlagen beruhen. An dieser Stelle stellen sich einige Fragen: Wie viele Daten werden benötigt, um gute Vorhersagen zu erhalten? Welche Belastungsszenarien müssen abgedeckt werden? Wie weit kann ein Modell über den Trainingsdatenumfang hinaus noch prognostizieren? Wie können die erstellten Modelle validiert werden? In **Tabelle 4** wurde eine 5-fache Kreuzvalidierung verwendet, bei der auch die prozentualen Anteile der einzelnen Belastungsszenarien in das Training der Modelle einbezogen wurden. Um zu untersuchen, wie viele und welche Lastszenarien die Modelle benötigen, um gute Vorhersagen zu erzielen, wurde der gesamte Datensatz in achtzehn Segmente, sogenannte Folds, unterteilt, wobei jeder Fold ein ganzes Lastszenario enthält, wie in **Abb. 9** dargestellt. Diese Belastungs-Folds wurden dann zum Trainieren und Testen der verschiedenen Modelle verwendet und ihre Leistungskennzahlen bestimmt. Der RMSE und das Bestimmtheitsmaß (R^2) sind in den Abbildungen **Abb. 10** und **Abb. 11** am Beispiel des DT-Modells mit und ohne verzögerte Merkmale (**Tabelle 3**) dargestellt. Wie in **Abb. 10** zu sehen ist, ist die Leistung des DT-Modells ohne verzögerte Merkmale bei jedem neuen Fold unzureichend, selbst bei Training mit mehr als zehn Belastungs-Folds, obwohl das Modell bei der zufällig gewählten 5-fach Kreuzvalidierung sehr gut abschneidet, siehe **Tabelle 4**. Mit Ausnahme des Folds 3, der sich signifikant anders verhält als Fold 1 und 2, zeigt **Abb. 11** die erhöhte Genauigkeit bzw. Vorhersageleistung, die durch verzögerte Merkmale erreicht wird.



Abbildung 9: Teilung der Daten nach Belastungsszenarien für die Validierung (load fold validation)

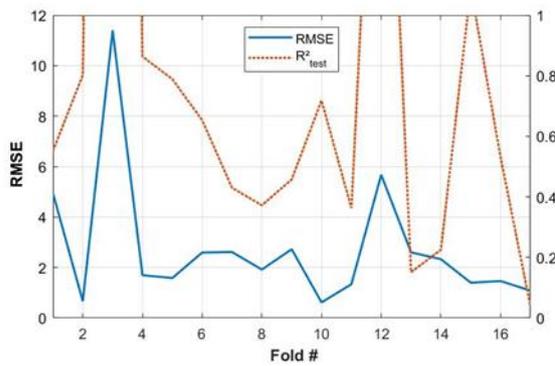


Abbildung 10: Validierungsmetriken für das DT-Modell ohne verzögerte Merkmale

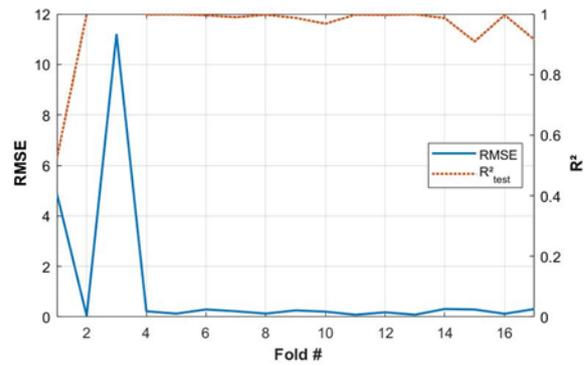


Abbildung 11: Validierungsmetriken für das DT-Modell mit zusätzlichen verzögerten Merkmalen

Durch die Verwendung der verzögerten Merkmale und des Load-Fold-Validierungsprinzips bleibt der RMSE des datengetriebenen Modells für alle weiteren Folds unter 1 K, selbst beim Wechsel von stationären Lasten zu dynamischen Lasten und für die verschiedenen Übergänge innerhalb der dynamischen Lastszenarien.

5.1.2 Validierung der Zeitreihenmodelle

Für die Validierung der Modelle mit dem direkten Prognoseansatz wurde der gesamte Belastungstestzeitraum in einzelne Tage unterteilt, was insgesamt 923 Tages-Folds ergab. Danach wurde jedes Modell auf den Daten der vorangegangenen Tage (z.B. Tag 1 bis 9) des Testzeitraums trainiert und jeder Datenpunkt des folgenden Tages (Tag 10 = Prognosetag) wurde anschließend mit den trainierten Modellen prognostiziert. Dabei wurden alle Modelle für jeden Prognosehorizont von 1 bis 24 Stunden verwendet. Mit anderen Worten, jedes Modell wurde, unabhängig vom Vorhersagehorizont, zur Vorhersage jedes einzelnen Datenpunkts jedes Tages abzüglich des anfänglichen Trainingsatzes (= Tag 1) verwendet. Fehlerkennzahlen wie RMSE und maximaler Fehler wurden für jeden einzelnen Prognosetag ermittelt und aufgezeichnet. Die resultierenden RMSEs für die RNN-Intraday-Prognosemodelle sind in **Abb. 12** dargestellt. Die maximalen resultierenden Fehler für fünf verschiedene Fehlerkategorien sind in **Tabelle 5** angegeben. Bei allen Modellen liegen über 99,5 % der Fehler unter 1 K, und keines der Modelle weist einen Fehler von mehr als 10 K auf. Insgesamt treten bei allen RNN-Vorhersagemodellen zusammengenommen nur zu acht Zeitpunkten Kabeltemperaturprognosen mit einem Fehler von mehr als 1 K auf.

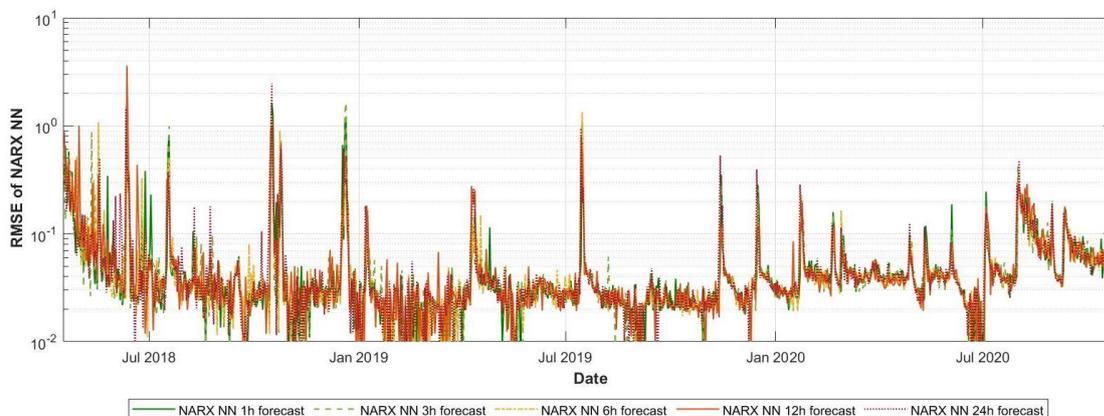


Abbildung 12: Abweichung zwischen gemessener Temperatur und Vorhersage des RNN-Intraday-Prognosemodells für 922 Prognosetage

Tabelle 5: Prozentuale Fehler des RNN für die Intraday-Prognosen

Prognose- horizont	Prozentualer Fehler [%]				
	< 1 K	1 K < Fehler < 3 K	3 K < Fehler < 5 K	5 K < Fehler < 10 K	Fehler > 10 K
1 Stunde	99.72	0.26	0.014	0.014	0
3 Stunden	99.71	0.27	0.014	0.01	0
6 Stunden	99.81	0.15	0.023	0.01	0
12 Stunden	99.67	0.29	0.014	0.02	0
24 Stunden	99.84	0.14	0	0.02	0

6 Zusammenfassung

Die bisherigen Untersuchungen an einem 400 kV-Kabelversuch der Wiener Netze haben gezeigt, dass selbst einfache datengetriebene Modelle auf Basis des maschinellen Lernens, wie Entscheidungsbäume und einfache neuronale Netzwerke, verwendet werden können, um die Kabeltemperatur mit angemessener Genauigkeit vorherzusagen. Die thermische Dynamik des Kabelsystems konnte allein aus den aufbereiteten Daten automatisiert abgeleitet werden. Allerdings muss auf die Validierungsmethode dieser datengetriebenen Modelle geachtet werden, da sie weit weniger robust sind als herkömmliche Modelle. Es hat sich gezeigt, dass die Implementierung einer zeitlichen Komponente, z.B. durch verzögerte Merkmale oder durch die Verwendung von Zeitreihenmodellen, die Robustheit und die Vorhersagegenauigkeit erheblich verbessert. So konnten Modelle für Intraday-Vorhersagen für Zeithorizonte von 1 bis 24 Stunden mit einer Abweichung von weniger als 1 K in 99,75 % der Fälle erstellt werden. Dennoch sind die hier erzielten Ergebnisse, sowohl bei der Modellierung als auch bei der Datenanalyse, auf die untersuchte spezifische 400 kV-Kabelanordnung beschränkt und keineswegs allgemein gültig. Außerdem haben sich die hier vorgestellten Untersuchungen auf nur einen räumlichen Messpunkt und nicht auf verteilte Messungen konzentriert.

7 Bibliografie

- [1] European Union Agency for the Cooperation of Energy Regulators, 2021, "Position on incentivising smart investments to improve the efficient use of electric transmission assets"
- [2] European Union, 2019, "Directive (EU) 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27/EU"
- [3] IEC 60287:SER, 2023, "Electric cables – ALL PARTS"
- [4] IEC 60853-2, 1989, "Calculation of the cyclic and emergency current rating of cables. Part 2: Cyclic rating of cables greater than 18/30 (36) kV and emergency ratings for cables of all voltages"
- [5] IEC 60853-3, 2002, "Calculation of the cyclic and emergency current rating of cables - Part 3: Cyclic rating factor for cables of all voltages, with partial drying of the soil"
- [6] CIGRE WG B1.56 – Technical Brochure 880, 2022, "Power cable rating examples for calculation tool verification"
- [7] IEC TR 62095, 2003, "Electric cables – Calculation for current ratings – Finite element method"
- [8] S. Venkata, Z. Sumic, S. Vadari and C.-C. Liu, 1993, "Applying AI systems in the T&D arena," IEEE Computer Applications in Power, vol.6, no.2, pp. 29-34
- [9] I. Dahhaghchi, R. Christie, G. Rosenwald and C.-C. Liu, "AI application areas in power systems," IEEE Expert, vol. 12, no. 1, pp. 58 - 66, 1997.
- [10] V. Viswanathan, V. Krishnan and L. Tsoukalas, "Novel AI approaches in power systems," in Proceedings 1999 International Conference on Information Intelligence and Systems, Bethesda, 1999.
- [11] R. Woschitz, F. Ainhirn und A. Bolzer, 2018, "Versuchsstrecke zur Untersuchung der magnetischen und thermischen Auswirkungen eines 400-kV-Kabelsystems", VDE-Fachtagung Hochspannungstechnik 2018, November 2018
- [12] F. Ainhirn, 2021, "Development of an Extended Calculation Procedure for the Thermal Rating of a 400 kV Cable System Based on Empirical Data", Doctoral Thesis, Graz University of Technology
- [13] F. Ainhirn, 2020, „Environmental Parameter Modelling for Thermal Rating Calculations of Power Cables in Urban Areas using Machine Learning and Big Data“, 2020 IEEE International Smart Cities Conference (ISC2), Piscataway, NJ, USA, 2020, pp. 1-7, doi: 10.1109/ISC251055.2020.9239014
- [14] F. Ainhirn and A. Bolzer, "Evaluation of Temperature Sensor Placements for Machine Learning Applications in the Thermal Rating of a 400-kV-Cable System," IEEE EUROCON 2021 - 19th International Conference on Smart Technologies, Lviv, Ukraine, 2021, pp. 503-507, doi: 10.1109/EUROCON52738.2021.9535537.
- [15] F. Gers, D. Eck and J. Schmidhuber, 2001, "Applying LSTM to Time Series Predictable through Time-Window Approaches," in Artificial Neural Networks - ICANN 2001 - Proceedings, Springer, 2001, pp. 669 – 676