

Prognose über die Stromerzeugung einer Photovoltaikanlage mittels maschinellem Lernen

Bachelorarbeit

Vorgelegt von

Leopold Schmid

Matrikelnummer:79776

Fakultät für Elektronik und Informatik

Hochschule Aalen

Datum

Erklärung

Verzeichnis häufig verwendeter Symbole und Abkürzungen

Inhaltsverzeichnis

1	Einführung	4
2	Stand der Technik	6
2.1	Strombedarf in Deutschland	6
2.2	Schwankungen im Stromnetz	6
2.3	Maschinelles Lernen	6
3	Eigene Fragestellung und methodisches Vorgehen	7
4	Wichtige Faktoren bei der Stromerzeugung durch Solarenergie	9
4.1	Wetter-unabhängige Faktoren	9
4.2	Nachgeführte Photovoltaikanlagen	11
4.3	Wetter-abhängige Faktoren	13
5	Zusammenhänge der Merkmale	15
6	Datenvorverarbeitung	16
7	Maschinelles Lernen - Auswahl des Modells	17

1 Einführung

In der grössten bisher durchgeführten Studie zu Klimaangst von jungen Menschen behaupten 45% der Befragten, dass ihre Gefühle bezüglich des Klimawandels ihr tägliches Leben negativ beeinflussen. Ohne jede Zweifel stellt der Klimawandel uns und die uns folgenden Generationen vor eine nicht zu unterschätzende Herausforderung. Angefangen mit der Tatsache, dass die alleinige Diskussion darüber nicht selten zu einer Spaltung und Polarisierung der Gesellschaft führt. Folglich leitet die hitzige und emotionale Debatte dazu, dass die gegensätzlichen Lager sich immer weiter voneinander entfernen und somit jegliche Grundlage für eine zielführende Diskussion entreissen.

Auf der einen Seite werden die Befürchtungen der anderen für übertrieben, unwichtig und paranoid gehalten. Wissenschaftliche Unsicherheiten werden verwendet, um gesamte Ergebnisse von Studien als unzuverlässig einzustufen. Modelle werden in der Wissenschaft oft vereinfacht, denn die Isolation und Fokussierung auf bestimmte Aspekte kann helfen, um ein besseres Verständnis von Phänomenen in der Natur zu erlangen. Auf Grund von dieser Vereinfachung werden die Klimamodelle als zu ausdruckslos betitelt, um die Komplexität des Klimas von unserem Planeten widerzuspiegeln. Dementsprechend seien die Schlussfolgerungen aus den Studien unzutreffend. Investitionen mehrerer Milliardenbeträge seien nicht gerechtfertigt und politische Vorgaben schaden dem eigenem Land mehr, als dass sie dem Planeten helfen würden.

Am gegenüberliegenden Ufer wird gemahnt, die Ernsthaftigkeit der Situation nicht zu unterschätzen. Gewarnt wird, dass die Konsequenzen des Klimawandels irreversibel seien, weswegen Treibhausgase unverzüglich auf ein Minimum reduziert werden sollten. Um Folgen wie das Abschmelzen der Eisschilde und Gletscher, das Aussterben verschiedenster Tierarten und der Anstieg des Meeresspiegels zu verhindern, spielen erneuerbare Energien eine entscheidende Rolle. Im Gegensatz zu fossilen Energieträgern erzeugen die Erneuerbaren keine oder nur kaum Treibhausgasemissionen. Bis 2030 sollen die regenerativen Energien 80% des Strombedarfs Deutschlands decken. Da die Möglichkeiten der Energieerzeugung in Deutschland durch Wasserkraft und Biomasseverbrennung schon

heute nahezu ausgeschöpft sind, bilden Solar- und Windenergie einen wichtigen Grundpfeiler um die Ausbauziele zu erreichen.

Ebenso komplex wie die Klimamodelle, die die Verstrickungen verschiedener Phänomene in unserer Natur berücksichtigen sollen, ist allerdings die Transformation unseres Stromnetzes. In einem Netz, welches dafür ausgelegt wurde, dass wenige gröSSere Kraftwerke Strom einspeisen, werden künftig immer mehr kleinere, dezentrale Kraftwerke mitwirken. Als Folge kommen diverse Herausforderungen auf uns zu. Unter anderem verträgt nicht jedes elektronische Bauteil in einem Niederspannungsnetz Rücklaufstrom. (Quelle?) Allerdings könnte es genau zu einem solchen Rücklaufstrom kommen, wenn in dem Niederspannungsnetz mehr Strom erzeugt als verbraucht wird. Des Weiteren reagiert ein Stromnetz äusserst empfindlich auf Spannungsschwankungen. Auf Grund dessen müssen Netzbetreiber darauf achten, dass genauso viel Strom verbraucht wie erzeugt wird. Solar und Windenergie zählen zu den fluktuierenden Energieerzeugern, ergo ist die Stromproduktion nur sehr begrenzt regulierbar und somit nicht an die aktuelle Marktnachfrage anpassbar.

Die Relevanz der Abschlussarbeit basiert auf der Annahme, dass eine genauere Prognose über die erzeugte Strommenge helfen würde das Stromnetz zu stabilisieren und Netzschwankungen, ergo das Risiko eines Brown- oder sogar Black-Outs zu minimieren. Anhand der neu gewonnenen Information könnten kurzfristige Stromimporte beziehungsweise -exporte besser reguliert werden. Ebenso wäre es denkbar, dass man für die Industrie und/ oder Endverbraucher mehr Anreize schafft ihren Stromverbrauch zu einem gewissen Grad an die Verfügbarkeit des Stroms anzupassen, wobei die Information über die Verfügbarkeit von Strom gleichfalls hilfreich wäre.

2 Stand der Technik

2.1 Strombedarf in Deutschland

2.2 Schwankungen im Stromnetz

Dementsprechend ist es für uns unverzichtbar, die Wichtigkeit einer stabilen Stromerzeugung richtig einzuschätzen und anhand dessen MaSSnahmen zu ergreifen, um Spannungsschwankungen im Netz so gut wie möglich zu reduzieren.

2.3 Maschinelles Lernen

3 Eigene Fragestellung und methodisches Vorgehen

Um die zukünftigen Herausforderungen der Energieversorgung zu bewältigen, befasst sich diese Arbeit damit den fluktuierenden Stromerzeuger, die Photovoltaikanlage, so gut wie möglich in unser bestehendes Stromnetz zu integrieren. Die zentrale Fragestellung lautet hierbei, wie man präzise den erzeugten Solarstrom prognostizieren kann. Um dieses Ziel zu erreichen, sollen zuerst entscheidende Parameter ermittelt werden, die die Stromerzeugung einer Solaranlage beeinflussen. Allerdings liegt der Fokus ebenfalls darauf, dass die Berechnung der Prognose für beliebige Photovoltaikanlagen anwendbar ist, indem nur wenige spezifische Informationen über die Anlage in die Berechnung einfließen. Selbstverständlich spielen verschiedenste Eigenschaften der Solaranlage eine entscheidende Rolle, jedoch werden diese im Datenmodell nicht berücksichtigt. Zunehmende Rechenkapazitäten und kostengünstigere Speichermöglichkeiten ermöglichen es, dass für jede einzelne Solaranlage ein eigenes Datenmodell angelegt wird. Somit können die Gegebenheiten der Photovoltaikanlage, wie Quantität und Qualität der Solarpanele, vernachlässigt werden. Durch die Trainingsdaten ist das Datenmodell selbstständig in der Lage die Leistungsfähigkeit der Solaranlage zu beurteilen. Ein weiterer Vorteil ist, dass dadurch schwer zu beschaffende Datensätze nicht weiter benötigt werden. Stattdessen wird das Modell fast ausschließlich von Wetterdaten trainiert, welche von verschiedensten Institutionen flächenmäßig und umfangreich aufgezeichnet werden. Die für das Trainieren des maschinellen Lernen Modells benötigten Wetterdaten werden von der Website <https://www.visualcrossing.com> extrahiert. Die Website bietet einen kostenfreien Zugang zu detaillierten historischen Wetterdaten, bis zu 50 Jahren vor heute.

HIER: QUELLE SOLARDATEN

Durch die Literaturrecherche sind bereits die relevanten Faktoren eingegrenzt wurden, allerdings ist eine exakte Analyse der Daten unumgänglich. Hierbei soll vor allem beachtet werden, ob die Aufnahme des Merkmals einen solch starken Einfluss auf die Stromerzeugung hat, sodass die zunehmende Komplexität des Modells gerechtfertigt ist. Herausfordernd ist dabei, ob die Informationen in man-

chen Merkmalen nicht bereits in anderen Merkmalen enthalten sind. Schließlich korrelieren einige Wetterdaten sehr stark miteinander. Ebenso gilt es zu überprüfen, ob sich die ausgewählten Merkmale auf unterschiedliche Solaranlagen auf die gleiche Art und Weise auswirken. Um dies zu überprüfen, sollen mehrere Solaranlagen an verschiedenen Standorten ausgesucht und getestet werden.

Sofern es gelingt, sämtliche sich variierende Parameter herauszufinden, die die Solarproduktion beeinträchtigen, ist maschinelles Lernen eine vielversprechende Technologie um die Stromerzeugung zu prognostizieren. Insbesondere ist zu erwarten, dass sich das Spektrum der Trainingsdaten und das Spektrum während der Produktion (Inbetriebnahme) sich nicht groß voneinander unterscheiden werden.

Eine Unvollkommenheit des Datenmodells ist jedoch die Alterung der Solaranlage, welche dazu führt, dass die Leistungsfähigkeit über die Lebenszeit sich reduziert. Wie stark die Alterung jedoch die Prognose beeinträchtigt, lässt sich über den Zeitraum dieser Arbeit von vier Monaten nur schwerlich beurteilen. Mögliche Konsequenzen wären ältere Daten kontinuierlich auszusortieren und mit das Modell mit den Neusten zu aktualisieren. Es ist davon auszugehen, dass der Alterungsprozess sich nur langsam in den Datensätzen bemerkbar machen wird, weshalb er die Ergebnisse über die kurze Zeitspanne kaum verfälschen dürfte.

4 Wichtige Faktoren bei der Stromerzeugung durch Solarenergie

Bevor der Stromertrag einer Photovoltaikanlage mit Hilfe von künstlicher Intelligenz berechnet werden soll, ist es von immenser Bedeutung die maßgeblichen Faktoren herauszufinden. Wenn ein Modell eingelernt werden soll, können zu viele Faktoren zum *Fluch der Dimensionen* führen. Dieser besagt, dass zu viele Merkmale dazu verleiten, dass Muster und Strukturen sich schwerer erkennen lassen. Die Datenpunkte sind durch die Größe des mehrdimensionalen Raums weiter voneinander entfernt, wodurch die Interpretierbarkeit komplexer wird. Dementsprechend sind die Eingabedaten behutsam auszuwählen.

4.1 Wetter-unabhängige Faktoren

In Bezug auf die Rahmenparameter der Solaranlage ist die reine Größe selbstverständlich ein maßgeblicher Faktor. Hinzu kommt, dass sich dieses Merkmal für jede individuelle stark unterscheiden kann. Die Dimensionierung der Solaranlage hängt schließlich ebenfalls von verschiedenen Faktoren ab, darunter die verfügbare Fläche am Standort, dem Energiebedarf und dem Verwendungszweck.

Durch die Aggregation von Metadaten der Solaranlagen könnte man die zuvor genannten Faktoren ebenfalls in das Datenmodell einfließen lassen. Die Komplexität des Modells würde erheblich zunehmen, jedoch wäre es somit auf verschiedene Solaranlagen verallgemeinerbar. Zwei Faktoren überdehnen jedoch nicht nur die Möglichkeiten der Anwendungen von maschinellem Lernen - oder würden die Vorhersagen des Modells zumindest signifikant beeinträchtigen. Des Weiteren stellen sie teilweise eine große Herausforderung beim Schritt der Datenerfassung dar. Erster Faktor ist die Umgebung der Solaranlage, die sich folglich nur auf jene Solaranlage auswirken. Umliegende höhere Gebäude, Hügel und Bäume können Schatten auf die Solarmodule werfen, wodurch die Sonnenstrahlung blockiert und die Stromproduktion reduziert wird. Gesondert anspruchsvoll wird es dadurch, dass der geworfene Schatten von der Elevation der Sonne abhängig ist. Die Elevation der Sonne wiederum ist von der Jahreszeit abhängig. Die

Wahrscheinlichkeit, dass die Solarpaneele durch umliegende Strukturen verdeckt wird, ist somit im Winter höher.

Der zweite Faktor ist die Ausrichtung der Solarmodule bezogen auf die Himmelsrichtung. Sofern die Module statisch befestigt sind, sollten die Module Richtung Süden ausgerichtet sein, um den höchsten Stromertrag zu erzielen. Allerdings lassen sich die örtlichen Gegebenheiten bei der Installation der Solarpaneele nicht ignorieren. Vor allem Solaranlagen, die für den Eigenbedarf installiert wurden, sind häufig auf Dächern befestigt. Allein aus Sicherheitsgründen werden die Solarmodule in aller Regel flach auf den Dachziegeln des Schrägdachs montiert, da ansonsten starker Wind die Module aus ihrer Befestigung reißen könnte. Folglich gibt die Ausrichtung des Gebäudes in vielen Fällen die Ausrichtung der Solarmodule ohne großen Spielraum vor. Konsequenz dessen ist, dass die Leistungsfähigkeit einer Solaranlage so vielfältig sein kann, wie die Bedachung von Häusern individuell ist.

Die Ausrichtung der Solarmodule hat nicht nur auf die gesamte erzeugte Strommenge Auswirkungen, des Weiteren führt sie dazu, dass Solaranlagen mit vergleichbarer Gesamtleistung zu unterschiedlichen Uhrzeiten kontrastiert Strom produzieren. Für die Netzstabilität ist es jedoch zwingend notwendig, dass kontinuierlich für eine gleichbleibende Spannung gesorgt wird. Folglich interessieren wir uns nicht nur für die kumulativ erzeugte Strommenge einer Photovoltaikanlage, weitaus spannender sind die Echtzeit-Vorhersagen.

Sowohl die Ausrichtung der Solarmodule als auch die Umgebung der Anlage soll in das Modell einfließen, indem Uhr- und Jahreszeit zu der entsprechenden Strommenge festgehalten werden. Die Logik dahinter ist mit der Annahme verbunden, dass sich die beiden Faktoren durch Uhr- und Jahreszeit repräsentieren lassen, weil die Konstellationen wiederkehrend sind. Ein Haus oder ein Baum, dass zwischen 11:14 und 11:46 Uhr Schatten auf die Solaranlage wirft, wird mit hoher Wahrscheinlichkeit die Stromproduktion der Solaranlage am nächsten Tag auf die gleiche Weise beeinträchtigen. Das Ziel der Forschung ist, dass das Datenmodell diesen Zusammenhang erkennt und eine niedrigere Stromproduktion prognostiziert als am Nachmittag, wenn sonst die gleichen Bedingungen vor-

liegen.

Die Jahreszeit bündelt mehrere Faktoren und nimmt somit Komplexität aus dem Modell, ohne dabei entscheidende Rahmenparameter zu vernachlässigen. Wie bereits erwähnt können sich die Umgebungsfaktoren über die Jahreszeiten hinweg verändern, zum anderen wandert die Sonne in einem anderen Winkel über die Solaranlage. Der Einstrahlungswinkel und die Umgebung sind entscheidend für die Stromproduktion und werden durch die Jahreszeit repräsentiert.

4.2 Nachgeführte Photovoltaikanlagen

Sogenannte nachgeführte Photovoltaikanlagen folgen selbstständig und automatisiert dem Sonnenstand, wodurch die Solarstromproduktion gegenüber stationären Anlagen verbessert wird. Optimalerweise trifft das Sonnenlicht fortwährend senkrecht auf die Solarmodule. Um dies zu bewerkstelligen, wird der Neigungswinkel und/ oder die Ausrichtung nach der Himmelsrichtung an den aktuellen Sonnenstand angepasst.

Die im letzten Unterkapitel genannten Faktoren werden aus dem Datenmodell mit der Begründung ausgeschlossen, weil für jede Solaranlage ein eigenes Modell angelegt wird und somit bei gleichen Wetterbedingungen dieselben Ergebnisse erzielt werden. Sowohl die Leistungsfähigkeit, der Wirkungsgrad des Wechselrichters als auch die Umgebungsfaktoren sind nahezu feste Rahmenparameter. Bei nachgeführten Photovoltaikanlagen haben wir nun den Fall, dass sich ein wichtiger Faktor, die Position des Solarmoduls in Bezug auf die Sonne, stets verändert, ohne dass dies in den Eingabedaten des Modells bemerkbar ist. Um die Konsequenzen für unser Vorhersagemodell zu beurteilen, müssen wir zwischen zwei verschiedenen Methoden, um eine Photovoltaikanlage nachzuführen, unterscheiden. Zum einen gibt es die astronomische und die sensorische Steuerung von PV-Anlagen. Bei der astronomischen Steuerung werden die Solarmodule kontinuierlich zur Sonne hin ausgerichtet, unabhängig von der Wolkendecke. In diesem Szenario ist die Ausrichtung der Solarmodule zwar dynamisch, allerdings wird mit Hilfe dieser Methode die Ausbeute der Stromproduktion stets auf die gleiche Art und Weise verbessert. Insofern macht es für das maschinelle

Lernen keinen Unterschied, ob die Solaranlage stets zur Sonne gewandt oder statisch montiert ist. Den Sonderfall, dass die Getriebemotoren beschädigt sind und ausfallen, sei an dieser Stelle auSSen vorgelassen.

Spannender wird es bei der aufwendigeren Technologie, die einen lokal installierten Sensor die optimale Ausrichtung der Solarpanele ermitteln lässt. Der hellste Punkt am Himmel wird durch die Sensorsteuerung wahrgenommen und dementsprechend werden die Sonnenkollektoren ausgerichtet. Der hellste Punkt am Himmel kann sich jedoch sehr schnell ändern, da er von der aktuellen Wolkendecke abhängt. Vor allem bei einer durchwachsenen Wolkendecke lässt sich nur schwer ermitteln, ob die Solarpanele im Moment von einer Wolke bedeckt werden.

HIER: WOLKENKAMERA ERMÖGLICHT DIES, WEBSITE VERLINKEN

Folglich wird die Ausrichtung der Sonnenkollektoren durch den Sensor kontinuierlich an die aktuellen Gegebenheiten angepasst. Die exakte Wolkensituation unterliegt einer stetigen Veränderung und wird keineswegs durch die Eingabedaten der Realität entsprechend repräsentiert, wodurch Ungenauigkeiten bei der Prognose entstehen können.

Die gleiche Problematik besteht auch bei herkömmlichen Solaranlagen, deren Module nicht sensorgesteuert ausgerichtet werden. Insbesondere werden die für dieses Projekt zur Verfügung stehenden Wetterdaten nicht in dem Intervall aktualisiert, indem sich die Wetterlage in der Wirklichkeit verändert. Da das Ziel dieser Arbeit ist, die gemittelte, erzeugte Strommenge stündlich vorauszusagen, soll uns die ständig schwankende Leistungsabgabe einer Solaranlage nicht störend beeinflussen.

Inwiefern die Prognose bei durch einen Sensor nachgeführten Photovoltaikanlage verschlechtert wird, gilt es zu untersuchen. Da vor allem die sensorgesteuerte Variante der nachgeführten Photovoltaikanlagen auch mehrere Nachteile, wie höhere Installations- und Wartungskosten mit sich bringt, wird der Marktanteil solcher Anlagen auf eher gering eingeschätzt. Folglich werden die Auswirkungen auf die Prognose nicht in dieser Arbeit untersucht.

4.3 Wetter-abhängige Faktoren

Verschiedene Wolkentypen wirken sich unterschiedlich auf die Sonneneinstrahlung aus. Zudem gehören Wolken zu den unbeständigen Faktoren, die sich sprunghaft auf die Stromerzeugung auswirken. Selbst wenn exakte, detailreiche Daten über die Bewölkung für die Forschung dieser Arbeit nicht vorliegen, soll der Effekt der unterschiedlichen Wolkentypen im folgenden kurz erläutert werden.

Stratuswolken sind flache, graue Wolken, die den Himmel oft bedecken. Sie bestehen aus Wassertröpfchen und liegen in niedriger Höhe. Stratuswolken blockieren die Sonneneinstrahlung und reduzieren die Helligkeit des Tageslichts erheblich. Sie haben eine kühlende Wirkung, da sie einen groSSen Teil der Sonnenenergie reflektieren.

Cumuluswolken sind groSSe, weiSSe, flauschige Wolken mit einer flachen Basis und einer kuppelförmigen Oberseite. Sie treten oft an sonnigen Tagen auf. Cumuluswolken können die Sonneneinstrahlung beeinflussen, indem sie sie teilweise reflektieren und teilweise absorbieren. Dadurch entstehen Schatten und Sonnenflecken auf der Erdoberfläche.

Cirruswolken sind dünne, faserige Wolken, die in groSSen Höhen schweben. Sie bestehen aus Eiskristallen und erscheinen oft als Federwolken oder Schleierwolken. Cirruswolken lassen viel Sonnenlicht durch und haben daher eine geringere Auswirkung auf die Sonneneinstrahlung. Sie können jedoch einen Schleier vor der Sonne bilden und das Licht diffus erscheinen lassen.

Nimbostratuswolken sind dichte, graue Wolken, die mit starkem Niederschlag verbunden sind. Sie erstrecken sich über groSSe Gebiete und sind oft mit anhaltendem Regen oder Schneefall verbunden. Nimbostratuswolken blockieren die Sonneneinstrahlung weitgehend und führen zu trüben, düsteren Bedingungen.

Im Allgemeinen lässt sich sagen, dass die Teilchen (Wassertröpfchen, Eiskristalle, Staub, Pollen, Meeressalze und Schadstoffe), aus denen die Wolken bestehen, die Sonnenstrahlen reflektieren oder absorbieren. Auf die gleiche Weise können Wasser- und Luftpartikel in der unteren Troposphäre die Strahlung beeinflussen. Folglich ist die gemessene Luftfeuchtigkeit und der Luftdruck ein

weiteres Kennzeichen dafür, wie viele Teilchen sich in der Troposphäre befinden.

5 Zusammenhänge der Merkmale

6 Datenvorverarbeitung

7 Maschinelles Lernen - Auswahl des Modells