

Prognose von Bleichspargelerträgen mit Neuronalen Netzen

JÜRGEN JAKI, WIESBADEN

PETER JÜRGEN PASCHOLD, GEISENHEIM

KAI VELTEN, WIESBADEN

Abstract

Neural Networks have been used to estimate the daily yield of various white asparagus types. The calculations were based on measured soil temperatures and on yield data obtained under realistic field conditions at the Forschungsanstalt Geisenheim. Various combinations of input parameters have been tested to find an empirical model for yield prognosis. In general, the Neuronal Network reproduces the yield with an accuracy of 25 kg/ha.

1 Einführung

Die Vorhersage von Spargelerträgen über einen Zeitraum von mehreren Tagen mit relativ hoher Genauigkeit stellt für die Spargelerzeuger ein wertvolles betriebswirtschaftliches Hilfsmittel zur Optimierung von Betriebsabläufen und zur Absatzplanung dar. In dieser Arbeit wird als erster Schritt in diese Richtung der Ertrag verschiedener Spargelsorten in Abhängigkeit von gemessenen Bodentemperaturen mit Hilfe von Neuronalen Netzen prognostiziert. Da solche Temperaturdaten nur für die zurückliegenden Tage gemessen werden können, müssen für eine spätere Anwendung dieser Modelle die Ergebnisse von laufenden Forschungsvorhaben einbezogen werden, in denen ausgehend von verfügbaren Wetterprognosen, die Temperaturen im Bodenbereich vorausberechnet werden.

Die hier verwendeten Daten stammen aus einem Feldversuch der Forschungsanstalt Geisenheim, bei dem unter praxisnahen Bedingungen 15 verschiedenen Sorten von Spargel angebaut werden. Für jede Sorte werden täglich der Ertrag (Gesamtmasse und Anzahl der geernteten Stangen) bestimmt, sowie die geernteten Stangen nach deren morphologischen Qualität bonitiert. Zusätzlich werden mit Hilfe von Datenloggern stündlich die Temperaturwerte der Luft und in zwei Bodentiefen sowie die Bodenfeuchte gemessen.

2 Neuronale Netze

Neuronale Netze stellen ein wichtiges und ausgereiftes Hilfsmittel zur Modellierung komplexer Zusammenhänge dar und werden zwischenzeitlich in vielen Bereichen der Forschung angewendet. Für eine einführende Beschreibung in die Arbeitsweise und die Anwendung von Neuronalen Netzen wird auf die entsprechende Literatur verwiesen [CAUDILL]. In dem hier gewählten Ansatz einem sogenannten „Feed Forward Netzwerk“ wird dem Netzwerk ein Trainingsdatensatz zur Verfügung gestellt, der die entsprechenden Eingangsgrößen sowie die dazugehörigen Ergebnisse enthält. Das Netzwerk bestimmt daraus eine Funktion, welche die Eingangsgrößen auf die Ergebniswerte abbildet. Im Wesentlichen beruht dies auf der Minimierung der quadratischen Abweichung zwischen Berechnungsergebnis und tatsächlichen Wert. Die so bestimmte Funktion wird dann auf einen Testdatensatz angewendet, um die Qualität der Prognosen zu überprüfen. Die Zahl der freien Parameter für diese Anpassung wird durch die Anzahl der Eingangsgrößen und die Netztopologie bestimmt. Hierbei ist darauf zu achten, dass eine große Anzahl von Parametern im Allgemeinen die Ergebnisse des Trainingsdatensatzes verbessert, aber unter Umständen für unbekannte Testdaten schlechtere Vorhersagen liefert. Dieses Problem (Overfitting genannt) kann durch Regularisierungsmethoden vermieden werden [FORESEE]. Dadurch werden die Ergebnisse auch weitgehend unabhängig von der gewählten Netztopologie. Die Berechnungen wurden mit dem Programm MathLab und der Neuronalen Netze Toolbox durchgeführt.

In dieser Arbeit soll der Ernteertrag Y einer bestimmten Spargelsorte (i), zu einem bestimmten Tag (j) aus den Temperaturwerten des Erntetages $T(j)$ und der vorausgegangenen Tage $T(j-1)$, $T(j-2)$, ... bestimmt werden. Zusätzlich muss auch der Erntezeitpunkt (bzw. die vorangegangene Ernte) berücksichtigt werden.

In einem ersten Modell wird davon ausgegangen, dass der Ertrag nur von den Tagestemperaturen und der Tagestemperatursumme bis zum Erntetag abhängt.

$$(1) \quad Y(i, j) = f(T_a(j), T_a(j-1), T_a(j-2), \dots, \sum_1^j T_a)$$

In einem zweiten Ansatz wird angenommen, dass der Ertrag von den Tagestemperaturen und der bisher erzielten Erntemenge abhängt.

$$(2) \quad Y(i, j) = f(T_a(j), T_a(j-1), T_a(j-2), \dots, \sum_1^{j-1} Y(i))$$

Für beide Modelle wurde untersucht, welche der zu Verfügung stehenden Temperaturen (Index a) bessere Ergebnisse liefert und wie viele zurückliegende Tage in der Berechnung berücksichtigt werden müssen.

3 Versuche

Ein Sortenversuch für Bleichspargel wurde am 23. März 2000 mit einem einheitlichen Pflanzabstand von 1,80 m x 0,4 m und 15 Sorten bzw. Neuzuchtstämmen aus Deutschland, den Niederlanden und Frankreich auf dem Versuchsfeld Ingelheim der Forschungsanstalt Geisenheim angelegt. Je Parzelle werden 20 Pflanzen beerntet, wobei randomisiert 4 Wiederholungen von jeder Sorte im Versuch zu finden sind.

Die erste Ernte über 21 Tage erfolgte im Frühjahr 2001 [Paschold et al.]. Entsprechend dem Alter der Anlage wird die Stechzeit in den folgenden Jahren schrittweise verlängert. Da inzwischen praxisüblich, werden die Dämme ab 2002 während der Erntezeit mit einer schwarz-weiß-Folie überdeckt, um die Temperaturen für den Spargel günstiger zu gestalten.

Bei der täglichen Ernte werden die Stangenzahl und -masse bezogen auf eine Länge von 22 cm sowie deren morphologische Qualität entsprechend den Handelsklassen ermittelt. Die Düngung nach der Ernte erfolgt nach Bodenanalysen und die Bewässerung in Abhängigkeit von Bodenfeuchtesollwerten. Der Aufwuchs wird variantenabhängig bonitiert.

Die Temperaturen werden mit Pt 100 und Datenloggern stündlich in Kronentiefe und 5 cm unterhalb der Erdoberfläche, die Bodenfeuchte in Kronentiefe mit Folisensoren (Netafim) erfasst.

4 Aufbereitung der Daten

Insgesamt stehen Ertragsdaten über einen Zeitraum von 21 Erntetagen zur Verfügung, wobei sich bei einigen Sorten erst 5 Tage nach Erntebeginn der frühesten Sorte erste erntefähige Triebe gebildet hatten. Die jeweiligen Erträge pro Tag und Sorte wurden auf den täglichen Ertrag pro Fläche (t/ha) umgerechnet. Der durchschnittliche Ertrag über alle Sorten betrug ca. 0,1 t/ha, der Maximalertrag der einzelnen Sorten variiert zwischen 0,16 und 0,3 t/ha. Die täglichen Ertragswerte unterliegen starken Schwankungen, die zu einem Großteil auf statistische Streuungen zurückzuführen sind (pro Sorte standen 80 Pflanzen zur Verfügung). Deshalb wurde für die Berechnungen ein gleitender Mittelwert über die Erträge zweier aufeinanderfolgender Tage gebildet.

Aus den Temperaturwerten wurde eine durchschnittliche Tagestemperatur gebildet (jeweils zwischen 12.00 des vorangegangenen Tages bis um 12.00 des laufenden Tages). Im Beobachtungszeitraum schwankte die durchschnittliche Tagestemperatur in Kronentiefe zwischen 12,2°C und 18,2°C, die durchschnittliche Lufttemperatur lag zwischen 10,1°C und 22,5°C. Temperaturwerte vor Erntebeginn liegen nicht vor.

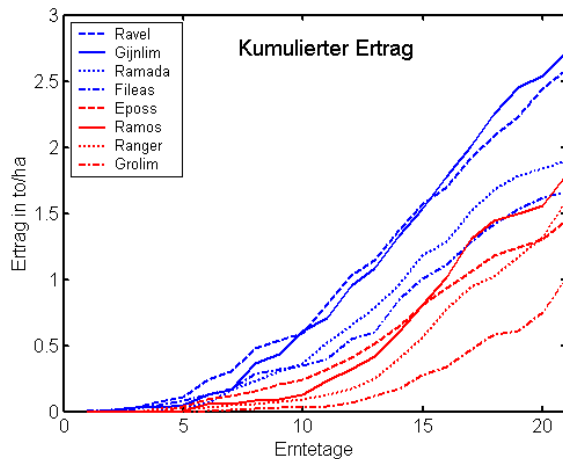


Abb. 1: kumulierte Erträge einiger Bleichspargelsorten

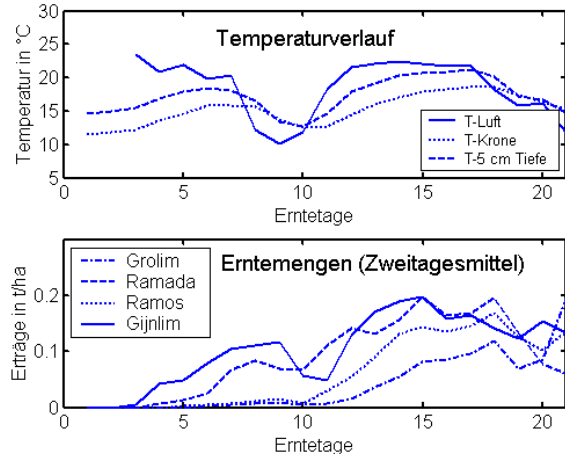


Abb. 2: Temperaturverlauf und Ertragsverläufe

Da die Erträge eines bestimmten Tages aus den Temperaturwerten der jeweils 4 zurückliegenden Tage berechnet werden sollen, konnten nur die Erträge ab dem 5. Erntetag in die Berechnungen einbezogen werden. Insgesamt liegen somit 16x15 Datensätze vor (16 Tage Ernte, 15 Sorten) welche mit Hilfe der Neuronalen Netze bearbeitet wurden. Aus diesen 240 Datensätzen wurden jeweils zufällig 190 Datensätze ausgewählt um das Neuronale Netz zu kalibrieren, d.h. an diese Daten anzupassen. Anhand der restlichen 50 Datensätze, die nicht in die vorherige Anpassung eingegangen sind, wurde die Genauigkeit der Berechnung getestet, indem die Vorhersage des neuronalen Netzes mit den tatsächlichen Ertragsdaten verglichen wurde. Diese Auswahl von Trainings- und Testdatensätzen wurde insgesamt zehn Mal wiederholt, um bessere statistische Aussagen über die Genauigkeit der Vorhersage treffen zu können.

5 Ergebnisse

Bei den Berechnungen zeigt sich, dass die Temperaturen des Erntetages und der zurückliegenden drei Tage ausreichen, um eine gute Beschreibung der Ertragsdaten zu erhalten. Eine Erweiterung des Ansatzes auf weitere Vortagestemperaturen lieferte keine signifikante Verbesserung der Vorhersage. Darüber hinaus wurde festgestellt, dass sowohl nur die Kronentemperaturen als auch nur die Temperaturen in 5 cm Bodentiefe für eine Vorhersage genutzt werden können. Eine Unterscheidung welche dieser Temperaturen bessere Ergebnisse liefert und ob eine Kombination dieser Temperaturen (wie erwartet) weitere Verfeinerungen zulässt, konnte im Rahmen der zu Verfügung stehenden Daten noch nicht definitiv getroffen werden.

Die zuvor beschriebenen Ansätze (1) und (2) liefern in etwa gleiche Resultate. D.h. sowohl die Vorhersage der Ertragsdaten nur aus den Temperaturwerten und der Tages-Temperatursumme als auch die Vorhersage aus den Temperaturwerten und dem kumulierten Ertrag führen zu guten Ergebnissen.

In Abbildung 3 ist der tatsächliche Ertragsverlauf (Zweitagesmittel) einiger Sorten als durchgezogene Linie gemeinsam mit den vorhergesagten Ertragswerten (Datenpunkte als Kreise) dargestellt. Die wiederholte zufällige Auswahl von Trainings- und Testdatensätze bewirkt, dass die einzelnen Tageserträge unterschiedlich oft berechnet werden. Die geringe Streuung dieser Datenpunkte zeigt die Unabhängigkeit der Modellierung von den gewählten Trainingsdatensätzen. Der Verlauf der Erträge der einzelnen Sorten wird jeweils richtig wiedergegeben, insbesondere das unterschiedliche Verhalten von frühen und späten Sorten zu

Erntebeginn sowie die Reaktion dieser Sorten auf die relativ großen Temperatursenkungen in der Mitte und am Ende der Ernteperiode.

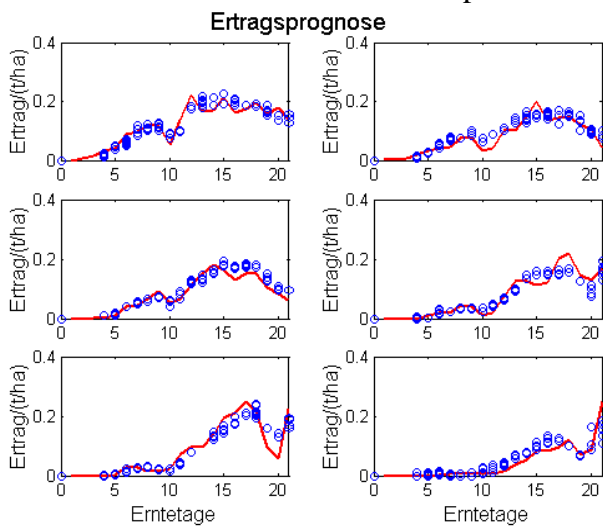


Abb. 3: Ertragsprognose

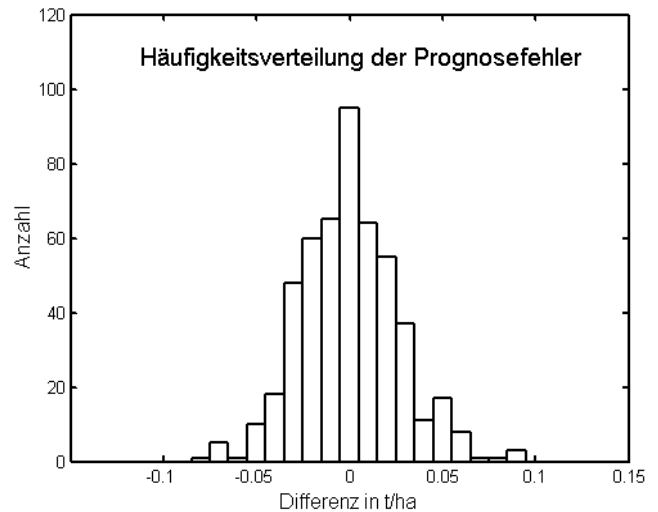


Abb.4: Fehlerverteilung

Die Abbildung 4 zeigt die Häufigkeit der aufgetretenen Abweichungen zwischen tatsächlicher Ernte und Vorhersage aufsummiert über alle Sorten. Die Standardabweichung dieser Kurve beträgt 0,026 t/ha. D.h. in 67% aller Fälle konnte der Ertrag innerhalb dieser Fehlergrenzen vorhergesagt werden. Nur in 6% aller Fälle lag der Fehler über 0,05 t/ha.

6 Folgerungen

Neuronale Netze stellen einen erfolgsversprechenden Ansatz zur Modellierung und Vorhersage von Ertragsdaten von Bleichspargeln dar. Mit den in der Folgezeit verfügbaren Daten aus den nächsten Erntejahren des Versuchsfeldes in Ingelheim sowie mit Daten von ausgewählten Erzeugerbetrieben, in denen ebenfalls Temperaturmessungen durchgeführt werden (und in denen Ertragsdaten mit wesentlich größerer statistischen Genauigkeit anfallen) können die beschriebenen Ansätze weiter verbessert und verfeinert werden. Sobald laufende Forschungsarbeiten zur Vorhersage von Bodentemperaturverläufen aus Daten der Wettervorhersage erfolgsversprechende Berechnungen liefern, können diese leicht in den hier beschriebenen Ansatz eingearbeitet werden, der dann ein wichtiges Vorhersagemodell für Spargelerträge bildet. In die weiteren Berechnungen werden auch anbaubeschleunigende und anbauverzögernde Maßnahmen (Folieneinsatz) einbezogen, die in diesem Ansatz allein über die Veränderung der Temperaturprofile berücksichtigt werden können. Darüber hinaus wurden erste Untersuchungen mit Hilfe von Neuronalen Netzen unternommen, um aus gemessenen Temperaturverläufen quantitative Aussagen über das Auftreten von Qualitätsmängeln wie hohlen Stangen, Berostung und Aufblühen abzuleiten.

7 Literatur

- DEMUTH, H. (1998): Neural Network Toolbox – for use with MathLab, The Math Works. Inc. Version 3.0
- FORESEE, F.D. (1997): “Gauss-Newton approximation to Bayesian regularization” Proceedings of the 1997 International Joint Conference on Neural Networks, pages 1930-1935
- CAUDILL, M. (1989): Neural Networks Primer, San Francisco CA: Miller Freeman Publications
- PASCHOLD, P.-J., HERMANN, G., ARTELT (2002): Zur Eignung neuer Spargelsorten. Gemüse 38, 2, 36-39