

## Ein statistisches Modell zur Schätzung des Rohproteingehaltes der Dauerwiesen Südtirols

Romano, G.<sup>1</sup>, Piepho, H.P.<sup>2</sup>, Schaumberger, A.<sup>3</sup>, Bodner, A.<sup>1</sup>, Florian, C.<sup>1</sup>, Figl, U.<sup>1</sup>  
& Peratoner, G.<sup>1</sup>

<sup>1</sup> Land- und Forstwirtschaftliches Versuchszentrum Laimburg,

<sup>2</sup> Universität Hohenheim, <sup>3</sup> HBLFA Raumberg-Gumpenstein

giovanni.peratoner@provinz.bz.it

### Einleitung und Problemstellung

Die Futterqualität kann am besten durch die laboranalytische Untersuchung von Futterproben bestimmt werden. Hinweise zur potenziellen Futterqualität von Wiesen können mittels Futterwerttabellen gewonnen werden, welche Schätzwerte für die verschiedenen Parameter der Futterqualität unter Berücksichtigung der Art der Futterkonservierung, des phänologischen Entwicklungsstadiums des Pflanzenbestandes sowie seiner botanischen Zusammensetzung im Sinne der Abundanz der funktionellen Gruppen Gräser, Leguminosen und Kräuter liefern (siehe z.B. DACCORD *et al.*, 2007; RESCH *et al.*, 2006). Für die Dauerwiesen Südtirols existieren bereits Futterwerttabellen für das Grünfutter beim ersten Aufwuchs (RESCH *et al.*, 2006). Das durchschnittliche Entwicklungsstadium des Bestandes ist ein Indikator für die Änderungen des Blatt/Halm-Verhältnisses sowie des Zellinhalt/Zellwände-Verhältnisses verschiedener Pflanzenteile im Laufe der phänologischen Entwicklung (BRUINENBERG *et al.*, 2002). Die botanische Zusammensetzung berücksichtigt dazu grundsätzliche Unterschiede zwischen Gräsern, Leguminosen und Kräutern im Verlauf der phänologischen Entwicklung.

Insbesondere stellt die Bestimmung des Entwicklungsstadiums des Pflanzenbestandes einen wichtigen Schritt dar. Da jede Pflanzenart einen artspezifischen Verlauf der phänologischen Entwicklung in Abhängigkeit der Umweltfaktoren aufweist (BOVOLENTA *et al.*, 2008), kann die Festlegung eines Entwicklungsstadiums, das für den gesamten Pflanzenbestand repräsentativ ist, schwierig werden. Auch die Verwendung von Referenzarten für die Bestimmung des Vegetationsstadiums, wie Knautgras und Goldhafer (RESCH *et al.*, 2006), kann problematisch werden, wenn der Hauptbestandbildner starke Unterschiede in der Entwicklung im Vergleich zur Referenzart aufweist (z.B. ein Wiesenfuchsschwanz-dominiertes Bestand, in dem Knautgras als Referenzart eingesetzt wird).

Ein möglicher ergänzender Ansatz zur genaueren Schätzung der Futterqualität besteht in der Mitbeziehung der Prognose von weiteren Variablen, welche den Wetterverlauf, die Topographie, die Bewirtschaftung, die Bodeneigenschaften sowie eine genauere Einstufung der botanischen Zusammensetzung (Wiesentyp) beschreiben. Ein erster Versuch in diese Richtung wurde mittels binärer logistischer Regression für den Rohproteingehalt der Südtiroler Dauerwiesen beim ersten Schnitt unternommen (PERATONER *et al.*, 2010). Damit konnte die Wahrscheinlichkeit der Überschreitung eines Grenzwertes mit relativ hoher Genauigkeit beschrieben werden. Im vorliegenden Beitrag wird die Weiterentwicklung dieses statistischen Modells beschrieben, um zu einer quantitativen Voraussage zu gelangen.

### Material und Methoden

Für die Entwicklung der statistischen Vorhersagemodelle wurde ein umfangreicher Datenbestand herangezogen, welcher den Verlauf der Futterqualität des ersten Aufwuchses von Dauerwiesen in Südtirol ab dem Entwicklungsstadium des Schossens beschreibt. Futterproben wurden durch wöchentliche sequentielle Beprobung für einen Zeitraum von 7 Wochen ab dem Weidestadium (als 15 cm Bestandeshöhe definiert und dem Schossen gleichgestellt) an 175 Umwelten (35 Standorte x 5 Jahre) von 2003 bis 2007, an 20 Umwelten (5 Standorte x 4 Jahre) von 2009 bis 2012 und an 7 Umwelten in den Jahren 2013 und 2014 gesammelt. Die Umwelten umfassten eine breite Palette

an klimatischen Bedingungen (666 m bis 1.593 m Meereshöhe) und Bewirtschaftungsweisen (2 bis 5 Schnitte/Jahr). Bei jedem Erntetermin wurden die Proben in vierfacher Wiederholung innerhalb einer Fläche von 50 bis 100 m<sup>2</sup> in Kleinquadraten (0,25 m<sup>2</sup> Fläche) bei einer Schnitthöhe von 5 cm mit Akkuscheren gewonnen (KASAL *et al.*, 2010). Der Rohproteingehalt wurde mit einem N-Analyser TruSpec<sup>TM</sup> (Leco Instrumente GmbH, D) nach Trocknung bei 60°C und der Vermahlung mit einer P 25 Labormühle (Fritsch GmbH, D) mit einem 0,5 mm Sieb bestimmt.

Bei der Beprobung wurde der Pflanzenbestand jeder Probe nach DACCORD *et al.* (2007) beurteilt. Beim 3. oder 4. Beprobungstermin wurden dazu jährlich die Ertragsanteile aller vorkommenden Pflanzenarten geschätzt. Jede Beprobungsfläche wurde mittels Clusteranalyse zu einem von 7 Wiesentypen zugeordnet (PERATONER *et al.*, 2010). Zu Beginn der Untersuchung wurden dazu die wichtigsten Angaben zur Bewirtschaftung (Schnitthäufigkeit, Art und Menge der ausgebrachten Wirtschaftsdünger, Vorhandensein einer Beregnungsanlage) erhoben und eine Bodenprobe analysiert. Meereshöhe, Hangneigung und Exposition wurden aus einem DGM mit 5 m Auflösung mit ArcGIS 10.0 ermittelt. Der Stickstoffinput wurde aufgrund der mittleren Gehalte an Gesamtstickstoff und Ammoniumstickstoff (unveröffentlichte Daten des Bodenlabors des Land- und Forstwirtschaftlichen Versuchszentrums Laimburg) sowohl im Bezug auf den ersten Schnitt (Gaben zwischen Herbst und dem ersten Schnitt) als auch für das gesamte Jahr berechnet.

Für die Entwicklung der statistischen Vorhersagemodelle wurden außerdem die folgenden meteorologischen und klimatischen Größen herangezogen:

1) **Temperatursumme:** Für 20 der 35 Versuchsstandorte wurden Temperaturdaten in 2 m Höhe von Wetterstationen in begrenzter Entfernung, mit ähnlichen Meereshöhe und Exposition sowie mit fehlendem Abschirmungseffekt von dazwischenstehenden Reliefs herangezogen. Für die restlichen Standorte wurden nach SCHAUMBERGER (2011) räumlich interpolierte Daten mit 250 m Auflösung verwendet. Die Basistemperatur von 0°C wurde nach ROMANO *et al.*, (2014) ermittelt.

2) **Summe der potenziellen Globalstrahlung,** welche mit dem Tool Solar Analyst von ArcGIS 10.0 mit einer Auflösung von 100 m ermittelt wurde.

3) **Die Summe (Sum $\Delta$ P) und der Mittelwert (Mean $\Delta$ P) der täglichen Abweichungen des Niederschlags vom langjährigen Tagesmittelwert (20 Jahre) von Referenzwetterstationen innerhalb homogener Niederschlagsbezirke.** Letztere ergaben sich als Kombination einer Clusteranalyse der monatlichen Niederschlagssummen von 2001 bis 2013 der Wetterstationen des Messnetzes der Autonomen Provinz Bozen mit den Wassereinzugsgebieten, welche auf GIS-Basis definiert wurden und nach dem Ergebnis der Clusteranalyse agglomeriert wurden.

1), 2) und 3) wurden auf das Zeitintervall zwischen einer Woche vor dem Datum des Schossens und dem Schnittzeitpunkt bezogen.

Die statistische Auswertung erfolgte mittels gemischter Modelle, welche die serielle Korrelation verschiedener Schnittzeitpunkte mit der Umwelt als Subjekt wiederholter Messungen berücksichtigten. Der Versuchsstandort, die Beprobungsfläche innerhalb des Standortes, das Jahr und die Wechselwirkung Jahr x Standort wurden als zufällige Effekte ins Modell eingeschlossen. Der Einfluss der Kovariaten wurde mittels polynomialer Regression beschrieben. Das statistische Modell wurde, ausgehend von einem Basismodell mit den Designeffekten und der Temperatursumme, schrittweise vorwärts (stepwise forward) entwickelt. Das Aikakesche Informationskriterium (AIC) wurde als Kriterium für die Aufnahme zusätzlicher Effekte verwendet. Effekte mit  $P > 0,1$  wurden im Laufe der Modellentwicklung entfernt, mit der Ausnahme von marginalen Effekten von Polynomen oder Haupteffekten bei Wechselwirkungen (NELDER, 2000). Das Modell wurde weiter optimiert, bis keine weitere Verbesserung von AIC zu erreichen war. Anschließend wurde das erhaltene Modell schrittweise vorwärts and rückwärts durch fünffache Kreuzvalidierung (HAWKINS *et al.*, 2003) weiterentwickelt, bis die quadrierte Korrelation zwischen beobachteten und vorhergesagten Werten seinen höchsten Wert erreichte. Die vorhergesagten Werte des fixen Teils des Modells wurden dafür verwendet. Die Komponenten der mittleren quadrierten Abweichung (Mean Squared Error) nach GAUCH *et al.* (2003) wurden zusätzlich für die Modellbeurteilung verwendet. Die Auswertung erfolgte mit wurzeltransformierten Daten, um die Normalverteilung der Residuen sowie die Varianzhomogenität zu gewährleisten. Vier unterschiedliche Modelle wurden in der Annahme entwickelt, dass nicht alle notwendigen Informationen bei ihrem praktischen Einsatz dem Nutzer effektiv zur Verfügung stehen oder geliefert werden können. So wurde ein Modell mit allen potenziell verfügbaren Variablen, eins nur mit den wahrscheinlich immer verfügbaren Variablen (Verzicht auf die Information zu Wiesentyp und Bodeneigenschaften) sowie zwei weitere Modelle, bei denen die

zusätzliche Verfügbarkeit der Information zu Wiesentyp (Verzicht auf die Bodeneigenschaften) oder zu den Bodeneigenschaften (Verzicht auf den Wiesentyp) angenommen wurde.

### Ergebnisse und Diskussion

Die Modellselektion wurde in 13 bis 16 Schritten in Abhängigkeit der Konstellation an verfügbaren Variablen abgeschlossen (Abb. 1a bis 1d). Die Änderung der AIC-Werte im Laufe der Modellentwicklung für Rohprotein zeigen, dass die Temperatursummen und der Pflanzenbestand die größte Verbesserung der Modellanpassung hervorrufen. Quantitativ weniger relevante aber dennoch signifikante Verbesserungen konnten durch die Aufnahme weiterer Variablen ins Modell erzielt werden. Alle meteorologischen und klimatischen Variablen wurden im Modell aufgenommen mit der Ausnahme von Mean $\Delta$ P bei den Modellen c) und d), die vollständig übereinstimmten. Im Gegensatz zum Pflanzenbestand spielte der Wiesentyp eine untergeordnete Rolle und wurde nur im Laufe der Optimierung durch Kreuzvalidierung des Modells a) aufgenommen.

Somit wird die Bedeutung der phänologischen Phase der Vegetation bestätigt, welche im Bezug auf die Temperatursummen und die botanische Zusammensetzung des Futters beschrieben werden kann (BOVOLENTA *et al.*, 2008).

- Temperatursumme (2 kennzeichnet das quadratische Glied des Polynoms)
- Pflanzenbestand

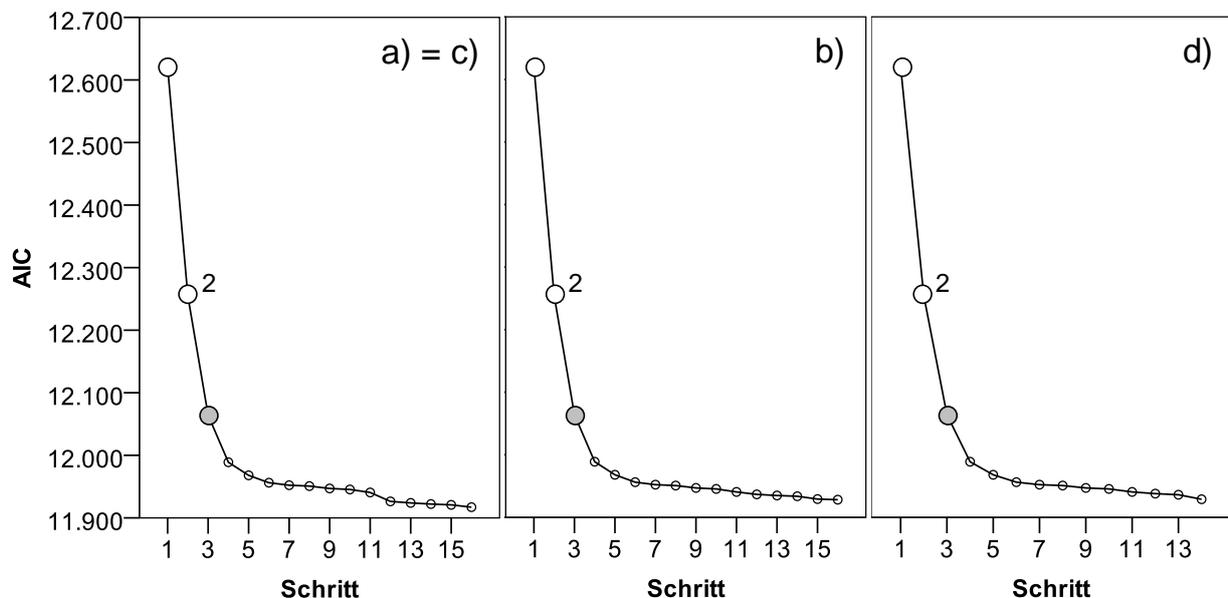


Abb. 1: Änderung der AIC-Werte bei der schrittweisen Modellselektion für Rohprotein. a) Verwendung aller verfügbaren Variablen, b) Ohne Verwendung der Bodeneigenschaften, c) Ohne Verwendung des Wiesentyps, d) Ohne Verwendung der Bodeneigenschaften und des Wiesentyps.

Tab. 1: Anpassung der entwickelten Modelle für Rohprotein. a) Verwendung aller verfügbaren Variablen, b) Ohne Verwendung der Bodeneigenschaften, c) Ohne Verwendung des Wiesentyps, d) Ohne Verwendung der Bodeneigenschaften und des Wiesentyps.

Parameter	Modell			
	a)	b)	c)	d)
R <sup>2</sup>	0,702	0,680	0,686	0,680
SB (Quadrierte Verzerrung)	0,002	0,003	0,002	0,003
NU (Uneinheitliche Neigung)	0,001	0,002	0,001	0,002
LC (Fehlende Korrelation)	1,116	1,198	1,178	1,198

Die Validierung des Modells wies ein Bestimmtheitsmaß zwischen 0,680 und 0,702 in Abhängigkeit der Anzahl der berücksichtigten Variablen auf (Tab. 1). Die Komponenten der mittleren quadrierten Abweichung zeigen bei allen Modellen sehr niedrige Werte für Translation und Rotation, welche von SB bzw. NU quantifiziert werden, während die fehlende Korrelation (LC) als vorwiegende Fehlerquelle zu betrachten ist (Tab. 1). Modell a), das unter Berücksichtigung aller verfügbaren Variablen entwickelt wurde, wies auch die niedrigste fehlende Korrelation auf. Die Unterschiede zwischen den Modellen a) bis d) sind allerdings insgesamt bescheiden.

### Schlussfolgerungen

Durch statistische Modelle kann der Rohproteingehalt von Dauerwiesen relativ genau vorausgesagt werden. Dabei haben Temperatursummen und Pflanzenbestand den größten Einfluss, aber die Verwendung anderer Variablen führt zu einer weiteren Verbesserung der Vorhersagegenauigkeit.

### Danksagung

Wir danken dem operationellen Programm "Regionale Wettbewerbsfähigkeit und Beschäftigung" EFRE 2007–2013 der Autonomen Provinz Bozen für die Finanzierung des Projektes webGRAS und dem Hydrographischen Amt der Autonomen Provinz Bozen für die Bereitstellung der meteorologischen Daten.

### Literatur

- BOVOLENTA, S., SPANGHERO, M., DOVIER, S., ORLANDI, D. & CLEMENTEL, F. (2008): Chemical composition and net energy content of alpine pasture species during the grazing season. *Animal Feed Science and Technology* 146, 178-191.
- BRUINENBERG, M.H., VALK, H., KOREVAAR, H. & STRUIK, P.C. (2002): Factors affecting digestibility of temperate forages from seminatural grasslands: a review. *Grass and Forage Science* 57, 292–301.
- DACCORD, R., WYSS, U., JEANGROS, B. & MEISSER, M. (2007): Bewertung von Wiesenfutter. Nährstoffgehalt für die Milch- und Fleischproduktion. AGFF Merkblatt 3. AGFF, Zürich.
- GAUCH, H.G., HWANG, J.T.G. & FICK, G.W. (2003): Model evaluation by comparison of model-based prediction and measured values. *Agronomy Journal* 95, 1442–1446.
- HAWKINS, D.M., BASAK, S.C. & MILLS, D. (2003): Assessing model fit by cross-validation. *Journal of Chemical Information and Computer Science* 43, 579–586.
- KASAL, A., STIMPFL, E. & PERATONER, G. (2010): A test of sampling methods for the investigation of forage quality in permanent meadows. *Grassland Science in Europe* 15, 542–544.
- NELDER, J.A. (2000): Functional marginality and response-surface fitting. *Journal of Applied Statistics* 27, 109–112.
- PERATONER, G., BODNER, A., STIMPFL, E., WERTH, E., SCHAUMBERGER, A. & KASAL, A. (2010): A simple model for the estimation of protein content of first-cut meadow forage. *Grassland Science in Europe* 15, 539-541.
- RESCH, R., GUGGENBERGER, T., GRUBER, L., RINGDORFER, F., BUCHGRABER, K., WIEDNER, G., KASAL, A. & WURM, K. (2006): Futterwerttabellen für das Grundfutter im Alpenraum. *Der fortschrittliche Landwirt* 84, 1–20.
- ROMANO, G., SCHAUMBERGER, A., PIEPHO, H.P., BODNER, A. & PERATONER, G. (2014): Optimal base temperature for computing growing degree-day sums to predict forage quality of mountain permanent meadow in South Tyrol. *Grassland Science in Europe* 19, 655-657.
- SCHAUMBERGER, A. (2011): Räumliche Modelle zur Vegetations- und Ertragsdynamik im Wirtschaftsgrünland. *Dissertation*. Universität von Graz.