

Voraussetzungen und Elemente für Validierungsarbeiten von praxisorientierten Humusbilanzmodellen

– vergleichbar den Richtlinien für Ringversuche des VDLUFA –

- **Unter praxisorientierten Humusbilanzmodellen** werden Verfahren angesehen, die lediglich auf einem in der Praxis **üblichen Merkmalskatalog** angewiesen sind und ohne weitere Merkmale und Kenntnisse von Nutzern der Praxis (z.B. Landwirt, Berater, Behörden, Studenten, etc.) ordnungsgemäß angewendet werden können.
- **Die Zielmerkmale des Verfahrens müssen eindeutig formuliert sein, damit entsprechende Kriterien für die Validierung festgelegt werden können:** Dient das Verfahren zur „Humusbilanzierung“ oder wird es vom Nutzer in dieser Funktion angesehen und verwendet, so müssen hierfür zutreffende Kriterien zur Validierung verwendet werden. Bei Methoden zur Humusbilanzierung sind es in erster Linie z.B. die C_{org} -Gehalte bzw. C_{org} -Mengen von Versuchsvarianten aus z.B. Dauerfeldversuchen. Andere Begleitmerkmale aus einschlägigen experimentellen Arbeiten können zusätzlich herangezogen werden (z.B. Erträge, N-Bilanz, etc.).
- **Zur Validierung sollten entsprechend aufbereitete Daten aus Exaktversuchen**, wie z.B. aus Dauerfeldversuchen Verwendung finden. Es ist auf eine hohe **Qualität und Genauigkeit** der erhobenen Daten zu achten. In der Regel erfüllen daher Daten aus entsprechenden Dauertestflächen der landwirtschaftlichen Praxis nicht die geforderten Ansprüche.
- Die **Vorgeschichte des Versuchsstandortes** ist genau zu erkunden. Liegen extreme Verhältnisse vor (z.B. Landnutzungsänderung von Grünland zu Ackerland etc.) so ist der Versuch für die Validierung u.U. ungeeignet.
- **Entsprechend der hohen standörtlichen und bewirtschaftungstypischen Variabilität ist auf eine entsprechende Quantität der Daten zu achten**, deren Umfang von der Nutzungsweise des zu prüfenden Modells abhängig ist. Bei einer vorgesehenen individuellen Nutzung für einen stark eingeschränkten Merkmalsbereich der Bewirtschaftung oder einer spezifischen regionalen Nutzung für nur einen Standort, kann **ein** Dauerversuch mit wenigen Varianten für die Validierung ausreichen. Für eine überregionale oder universelle Nutzung des Verfahrens müssen die Versuchsdaten möglichst umfassend **alle wesentlichen Standortaspekte** (Bodenarten, Textur, Niederschläge bzw. Wasserversorgung, Temperaturen etc.) und der **Bewirtschaftung** (Fruchtarten, Zwischenfrüchte, Fruchtfolgen, Ertragshöhe, organische Düngemittel u. Düngungshöhe, Intensität d. Bodenbearbeitung, etc.) abdecken. Ist z.B. Mitteleuropa (Deutschland) als potenzielles Einsatzgebiet des Modells vorgesehen, so ist erfahrungsgemäß hierfür eine Anzahl von 40 Versuchen mit ca. 400 Varianten in möglichst optimaler Verteilungsstruktur als ausreichend anzusehen.
- **Die Anzahl an Versuchen ist auch abhängig von der verwendeten Datenmenge je ausgewiesener Variante.** Werden z.B. für die C_{org} -Gehalte lediglich Anfangs- und Endwerte für den Beginn und das Ende des untersuchten Versuches verwendet, so muss zur Absicherung einer hohen statistischen Sicherheit die Versuchsanzahl bzw. Variantenanzahl höher sein, als bei Verwendung z.B. von jährlich erfassten Datensätzen. Zur Reduktion des Versuchsfehlers eignet sich auch die Bildung von Mittelwerten z.B. aus Daten von zwei bis vier Jahresscheiben, oder die Ausweisung des Versuchstrends z.B. für

C_{org} (insbesondere bei Vorlage weniger Untersuchungsdaten) mit Hilfe geeigneter statistischer Methoden (z.B. Regressionsanalyse).

- **Es ist zudem auf eine Mindestdauer der für die Prüfung heranzuziehenden Versuche von 10, besser von 15 Versuchsjahren zu achten**, damit die zeitliche Entwicklung (z.B. für C_{org}) statistisch abgesichert und fachlich eindeutig ausgewiesen werden kann.
- **Werden Anfangswerte z.B. für die C_{org} -Gehalte vorgegeben** (sog. Startwerte, die aus den (Anfangs)-Versuchsjahren generiert werden), so dürfen diese Werte nicht vom durchführenden Personal selber generiert werden, sondern sie müssen von neutraler Seite bereitgestellt werden.
- **Werden Startwerte verwendet, so können die Datenreihen der ersten 10 Jahre bzw. der ersten Versuchshälfte für die Validierungsauswertungen nicht verwendet werden**, da durch die Validierungsergebnisse dann eine zu genaue Modellübereinstimmung vorgetäuscht wird. Werden keine Anfangswerte vorgegeben, so können die Daten aus allen Versuchsjahren in die Validierung einfließen.
- **Die zeitliche Änderung der C_{org} -Gehalte vollzieht sich in Sättigungskurven**. Daher ist darauf zu achten, dass bei Einbeziehung von Jahrzehnte langen Dauerversuchen bei der Validierung von bestimmten (einfachen) Modellen bei der zeitlichen Fortschreibung eine **lineare Anrechnung der Bilanzergebnisse zu begrenzen** ist. Hierbei hat sich eine Begrenzung auf ungefähr 25 Jahre bewährt.
- **Für die Validierungsergebnisse sind folgende statistische Prüfwerte geeignet:**
 - **absolute bzw. relative Abweichung** zwischen experimentell ermittelten und mit den zu prüfenden Modellen berechneten Werten, hierbei kann auch die Richtung der Abweichung beurteilt werden;
 - **Korrelationskoeffizient** zwischen berechneten und experimentell ermittelten Werten, hierzu brauchen die Werte nicht auf eine gemeinsame Einheit (z.B. C_{org} -Gehalt) umgerechnet werden;
 - **statistische Streuung, Standardabweichung, Methodenstreuung;**
 - **graphische Darstellungen** z.B. durch Streudiagramm, Boxplot.
- **Für die Validierung müssen unabhängige Datensätze verwendet werden**, d.h. die Daten dürfen für die **Modellerstellung** nicht verwendet worden sein. Das trifft besonders zu, wenn Daten nur aus einer stark begrenzten Versuchsanzahl für die Modellierung verwendet worden ist. Im Bedarfsfall können die Daten zunächst geteilt werden, mit der einen Hälfte wird das Modell erstellt, mit der anderen Hälfte die Validierung vorgenommen.
- **Modelle zur Nutzung in der Praxis erfordern in der Regel die Bereitstellung eines Bewertungssystems für die Einstufung der berechneten Ergebnisse**. Dieses Bewertungssystem muss auch einer **experimentellen Prüfung und Validierung** unterzogen werden, wofür entsprechend der geforderten Zielstellungen, verschiedene **Merkmale** geeignet sind. So hat es sich bewährt, für das fünfstufige VDLUFA-Bewertungssystem zur Humusbilanzierung folgende Merkmale heranzuziehen: Ertragsreaktion der Fruchtarten (Ertragsoptimum), N-Schlag- bzw. -Flächenbilanz (brutto), C_{org} , N_t .
- **Ergebnisse einer Validierung werden nur als ordnungsgemäß anerkannt, wenn eine durchgängige Transparenz des Verfahrens und wenn eine Unabhängigkeit der durchführenden Personen gewährleistet werden kann**. Validierungsergebnisse, die

von Modellbauern selbst vorgenommen werden, können anerkannt werden, wenn eine Prüfung der Validierungsdurchführung von unabhängiger Seite erfolgt (z.B. zweite Arbeitsgruppe).

- Auch eine **Validierung entgegen der Intension des Modellbauers** kann auf Grund des Transparenz- und Informationsbedürfnisses des Nutzers geboten sein.
- **Die zu prüfenden Merkmale (z.B. C_{org}-Gehalte, Erträge, etc. von Versuchen) dürfen dem durchführenden Personal nicht bekannt sein.**
- **Von den zur Prüfung vorgesehenen Versuchen müssen in genau formuliertem Umfang Datenreihen als Eingabemerkmale zur Verfügung gestellt werden**, die zur ordnungsgemäßen Durchführung des Modells erforderlich sind, die aber vom Umfang nicht von dem Merkmalskatalog abweichen dürfen, die jeder Nutzer unter praktischen Verhältnissen auch bereitstellen muss, wenn er das Modell bedienen will.
- Einzelne Dauerversuche oder **Versuchsvarianten, die Merkmale enthalten, die in den zu prüfenden Modellen (noch) nicht implementiert sind** (z.B. selten vorkommende Fruchtarten oder organische Düngemittel, etc.) **können für die Validierung nicht verwendet werden.**
- Die **Datensätze über die Versuche müsse so anonymisiert werden**, dass das durchführende Validierungspersonal keine weiteren Erkenntnisse über die spezifischen Versuche erlangen können.
- **Arbeiten zur Datenbereitstellung, Anonymisierung und Auswertung der Validierungsergebnisse müssen in personeller Trennung von den Arbeiten zur Erstellung der Validierungsergebnisse vorgenommen werden** (z.B. in mindestens zwei Arbeitsgruppen).
- Die **Ergebnisse müssen in transparenter Weise in einem Bericht dokumentiert werden** und können in einer Tabelle nach folgendem Beispiel ausgegeben werden (Tab. 1).

Diese aufgeführten Bedingungen sollten in mehr oder weniger vollständiger Form erfüllt werden. Das Thesenpapier gibt allerdings keinen feststehenden Rahmen vor, sondern es ist als **Diskussionsgrundlage** für die ordnungsgemäße Durchführung einer Validierung von Humusbilanzmodellen z.B. im Rahmen des VDLUFA gedacht.

Tabelle 1: Vorschlag zur vergleichenden Darstellung von Validierungsergebnissen zur Genauigkeit von Humusbilanz-Methoden

Methode	Korrelation ¹⁾ zw. Modell- Ergebnis und Feldversuch		Mittlere Abweichung (0 ≈ 100 % Bedarfsdeckung mit organ. Substanz für genaue Reproduktion)			Ertrags- sicherung in VG ²⁾ C	Korrelation zw. Humus- u. N-Saldo ³⁾	N-Saldo bei Übergang zw. VG A/B bzw. D/E	Standort- anpassung	Einschätzung Methoden- genauigkeit
	C _{org} (r)	N _t (r)	C _{org} (%)	Humus- saldo (kg HÄQ- C/ha) ⁴⁾	Einschät- zung Repro- duktion					
						(opt = optimaler, max = maximaler Ertrag)	(r)	(kg N/ha)	(Bewertung)	(Bewertung)

1) Korrelation = Bewertung Korrelationskoeffizient (r): ≤ 0,25: keine bis schwache Beziehung; 0,25-0,50: schwache bis mäßige Beziehung; 0,50-0,75: deutliche bis gute Beziehung; ≥ 0,75: sehr hohe und enge Beziehung;

2) VG = VDLUFA-Versorgungsgruppen; 3) N-Saldo = entsprechend Brutto-N-Saldo (z.B. nach PARCOM-Richtlinie)

4) HÄQ-C = Humusäquivalente