

Statistische Verfahren WS 2018

Projekt 4 – Nahinfrarotspektroskopie I

Problemstellung:

Herleitung von Kalibrierungsfunktionen zur Schätzung des Bodenkohlenstoffgehalts mit Hilfe der Nahinfrarotspektroskopie

Datensatz: (NIR.csv, Dr. A. Don, Heinrich von Thünen-Institut Braunschweig)

Der Datensatz enthält Daten zum Bodennickstoffgehalt (N) sowie die dazugehörigen Nahinfrarotspektren, d.h. die gemessenen Reflexionswerte bei verschiedenen Wellenlängen aus dem Nahinfrarotbereich, für 534 Bodenproben. Diese Daten sollen zur Herleitung von Prognosefunktionen verwendet werden, die die Schätzung des nur aufwändig bestimmbarcn Bodenparameters N aus den viel leichter messbaren Nahinfrarotspektren gestattet.

Aufgaben:

- Bestimmen Sie mit Hilfe eines geeigneten Modellselektionsverfahrens Vorhersagemodelle für den Bodennickstoffgehalt N.

Simulationsaufgabe:

- Wählen Sie eines der im ersten Teil hergeleiteten Prognosemodelle als Grundlage zur Simulation von Pseudobeobachtungen. Entscheiden Sie sich für jeden simulierten Pseudo-Datensatz basierend auf Mallows Cp-Kriterium für ein bestes Modell. Vergleichen Sie den aus dem minimalen Cp-Wert geschätzten Vorhersagefehler mit dem tatsächlichen, theoretisch erwarteten Vorhersagefehler SPSE. Untersuchen Sie dabei den Einfluss des Stichprobenumfangs (indem Sie aus dem Originaldatensatz zufällig Spektren für die Designmatrix der Simulation auswählen).