

Werner Haufe

Kleinwanzlebener Saatzucht AG, Einbeck

ZUCHTBUCHFOHRUNG MIT HILFE DER EDV

Die Datenorganisation für die Züchtung der KWS gliedert sich in

- Elitespeicher - Buchführung
- Basissaatgut - Abgabeprotokolle
- Abstammungsdatei
- Individual - Selektions-Daten
- Saatrübenbuch (Reproduktionen)
- Feldnotizbuch (Stammzuchten, Stecklinge, Versuchsfelder)
- Jahresversuchsdatei
- Div. temporäre Datensammlungen

Davon werden z.Z. über die EDV bearbeitet

- die Verrechnung der Routineversuche über das *geschlossene Verrechnungssystem (ZGES)* und die Jahresversuchsdateien
- die Individual-Selektions-Daten und div. Datensammlungen über die *offene Daten- und Methodenbank (ZOFF)*
- die gleichen Daten können auch mit einem *Paket selbständiger Programme (ZDIV)* ausgewertet werden.

1. Geschlossenes Versuchssystem (für Routine-Versuche)

Wird abgehandelt über das selbst entwickelte Datenbanksystem (für FORTRAN zugreifbar) der "Jahresversuchsdateien", gliedert in

- Datei für Urdaten
- Datei für Ergebnisse

für 3 Versionen (jeweils auf einer Wechselplatte), und zwar für die Fruchtarten

- Zucker- und Futterrüben
- Getreide
- Mais, Kartoffeln und Sonstige.

Über die 3 parallelen Datei-Versionen laufen 5 verschiedene Programm-Versionen

- Zucker- und Futterrüben
- Getreide
- Körnermais
- Silomais
- Kartoffeln

(sonstige Fruchtarten werden über eine der 5 Programm-Versionen bearbeitet, die am geeignetsten ist).

Die Kennzeichnungs-, Sortierungs- und Merkmalsvariablen haben feste Plätze. Die Sätze eines Versuches werden durch eine Kennung zusammengefaßt.

Die über die Jahresversuchsdateien abzuwickelnden Aktivitäten lassen sich wie folgt auflgliedern:

- Verrechnung der Einzelversuche
- Zusammenfassende Verrechnungen mit weiteren Auswertungen
- Verrechnung von mehrfaktoriellen Einzelversuchen
- Zusammenfassende Verrechnung der mehrfaktoriellen Versuche
- Spezielle Verrechnungen.

Nachfolgend dazu einige Hinweise.

Verrechnung der *Einzelversuche* (= einfaktorielle Versuche mit Wiederholungen) (Block, quadratische Gitter, Rechteckgitter, lateinische Quadrate, Wechselgitter)

Teil 1 = Einlesen der Urdaten, Formal- und Vollzähligkeitsprüfung, Plausibilitätsprüfungen (Errechnung der Erwartungswerte); Archivierung

Teil 2 = Abruf der Daten, Satzaufbau für Hilfsdatei, Boniturverrechnung, Zählwertverrechnung, Ko-Varianzanalyse, lineare Ergänzung, Ersatz fehlender Werte, Standardisierung der Merkmale, dazu Registrierungssteps zur automatischen Erfassung der Datenstruktur und Erstellen einer Entscheidungstabelle, schließlich Varianzanalyse, Druck der Ergebnisse einschl. Print-plots, Archivierung der Ergebnisse.

Zusammenfassende Verrechnung von Einzelversuchen

Einfache zusammenfassende Verrechnung (Orte oder Jahre oder "Versuche")
Abruf der Ergebnisse der Einzelversuche, Verrechnung und Druck mit Printplots, weitere Auswertungen und Zusammenstellungen (absolute Werte, relative Werte, Rangzahlen, Tableau der Abweichungen mit β_i und d_{ij}^2 - alles pro Merkmal)

Zusammenfassende Verrechnung über Orte und Jahre

- wie oben; vorweg automatische Erstellung der Parameter für alle möglichen Kombinationen der wiederholten einfachen zusammenfassenden Verrechnung.

Verrechnung von mehrfaktoriellen Einzelversuchen
(max. 5 Faktoren + Wiederholung)

Abruf der Urdaten, Ausgabe der Varianztabelle und aller Kombinationsmittelwerte mit Relativzahlen sowie aller möglichen Zweiwegetafeln, Zusammenstellung aller Kombinationen und Druck gewünschter Merkmale nebeneinander.

Zusammenfassende Verrechnung mehrfaktorieller Versuche
wie mehrfaktorielle Einzelversuche.

Diverse Verrechnungen und Auswertungen

Selektionstabellierungen, Gruppierung von Daten.

Verrechnung von Wechselgittern (Split-plot-Anlage mit Gitterauswertung der Kleinparzellen).

Kreuzungsanalyse. - Subsystem, mit dem eine neue logische Struktur über die in der Jahresversuchsdatei vorhandene physische Struktur gestülpt wird. Weitere Verrechnung in etwa wie mehrfaktorielle Versuche.

Sonderverrechnung Schosserverversuche.

Prediction. - Errechnung der Leistungen von Dreibegehybriden und Doppelhybriden aus Einzelkreuzungen - mit und ohne Selektion.

Erstellung von Planungsunterlagen für die Versuche (Drilllisten, Wägelisten, Bonitierungslisten).

2. Offenes System einer Daten- und Methodenbank (ZOFF)

(für spezielle statistische Bedürfnisse und für die Auswertung der Daten der Individualelektion)

Für die Verwaltung des offenen Systems haben wir einen Prototyp, mit dem Daten von beliebigen Datenträgern in beliebigen Formaten eingelesen, selektiert, transformiert und auf eine Schnittstellen-Datei mit festen Formaten ausgegeben werden. Auf diese Schnittstellen-Datei greifen die Auswertungsprogramme zu; die Datei verfällt nach Beendigung des Jobs.

Grundausstattung math.-stat. Auswertungsprogramme

- Regressionsanalysen
- Häufigkeitsanalysen
- Varianzanalysen
- Diskriminanzanalysen
- div. Gruppierungen
- Rangordnungen.

Alle Programme sind auch für mehrfach gegliedertes Datenmaterial eingerichtet.

In Zusammenarbeit mit der Gesellschaft für Strahlenforschung München wurde inzwischen ein neues Verwaltungssystem für eine offene Daten- und Methodenbank erstellt. - Es ist noch nicht ganz ausgetestet, die Auswertungsprogramme sind noch nicht angepaßt, es fehlen auch noch Verwaltungshilfsprogramme.

3. Sammlung diverser selbständiger Programme (ZDIV)

Es handelt sich um anspruchsvollere ad-hoc-Programme oder sonstige selbständige Programme für verschiedene Anwendungen: Selektionsroutinen, Zufallsverteilungen, überlagerte Gleich- und Zufallsverteilungen, Beurteilung von Kurvenverläufen, Prognoseverrechnungen, Erstellen von Versuchsplänen etc.

4. Integrierte Zuchtdatei (IZDB)

Wir arbeiten ferner an einem EDV-System, das alle wichtigen Aktivitäten eines Zuchtbetriebes einschließt, also die anfangs vorgestellte Gesamtorganisation abdeckt. Das Grobkonzept dieses Systems (= Integrierte Zuchtdatei = IZDB) ist abgeschlossen und die Einführung des 1. Teilprojektes ist für Anfang 1981 vorgesehen.

Da die Entwicklung eines eigenen Dateiverwaltungssystems zu aufwendig ist, haben wir uns für das IBM-Produkt DL/I-Entry entschieden.

An einem Beispiel soll die Übertragung der Prinzipien für die Integrierte Zuchtdatei erläutert werden.

Abstammungsdatenbank (Abb. 1)

Sie dient zur Verfolgung der Herkunft (= Eltern) und der Verwendung (Kinder, Nachkommen), aber auch zur Beurteilung.

Die Datenbanksätze bestehen aus dem *Rootsegment* mit der Identnummer; dazu Verweis auf Entstehungsvorgang, Fruchtartenschlüssel, Klassifizierung des Materials, Verweis auf Sorte, Verweis auf Urabstammung etc.

An abhängigen Segmenten sind vorgesehen

- *Abstammungssegmente* mit Ident der Eltern (mehrere !), Art der Elternschaft, ggf. Anteil

ABSTAMMUNGS - DATENBANK

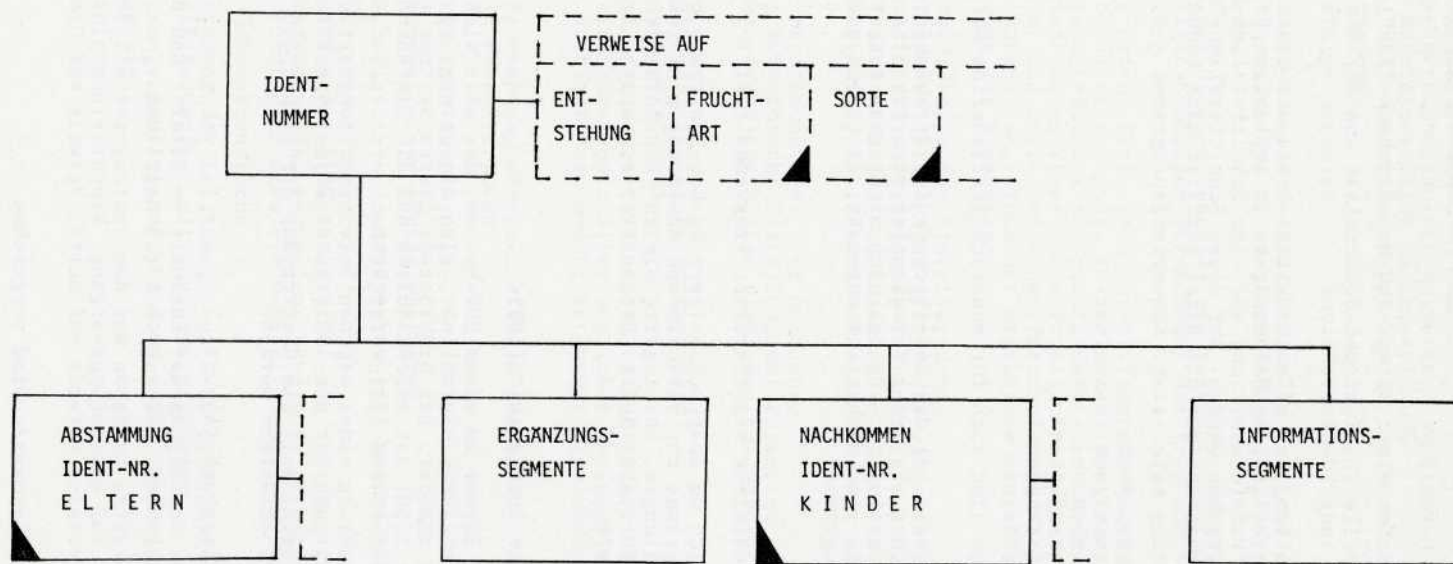


Abbildung 1

- *Ergänzungssegmente*
 - a) zur Urabstammung (= Teilelemente einer Kreuzungsformel)
 - b) Pedigree-Nr. (nach Generationen gegliederte Arbeitsnummern)
- *Nachkommen-Segmente* mit Ident der Kinder, Verweis auf Entstehungsvorgang, Art der Elternschaft, ggf. Anteilsziffern
- *Informations-Segmente* mit komprimierten Daten und Beurteilungen.

Obersicht über die gesamte Integrierte Zuchtdatei

Aus Kapazitäts- und Zweckmäßigkeitsgründen haben wir uns entschlossen, die Informationen nicht in *einer* alles umfassenden Datenbank zu speichern, sondern entsprechend ihren Funktionen in 6 einzelnen Datenbanken, die durch Schlüssel verknüpft sind.

Im einzelnen sind vorgesehen

- *Abstammungsdatenbank* zur Verfolgung von Herkunft und Verwendung sowie zur allgemeinen Beurteilung.
- *Lagerverwaltungs-Datenbank* für die Verwaltung des Saatgutes im Elite-Speicher.
- *Verwendungs-Datenbank*. Diese verwaltet die Verwendungen aus der Lagerverwaltungs-Datenbank, sammelt Informationen, handelt weitere Verwendungen im Sinne von Selektionen, gibt Verweise auf die Abstammungs-Datenbank, Die Lagerverwaltungs-Datenbank und die Versuchsbeschreibungs-Datenbank.
- *Versuchsbeschreibungs-Datenbank*. Sie dient zur Erledigung von Verwaltungsaufgaben wie Erstellen von Drill-, Wäge- und Bonitierungslisten, Feldnotizbüchern und Saatrübenbüchern. Sie gibt den Anstoß für die Bereichsfestlegungen in den Jahresversuchs-Dateien und die Vermittlung bei der Datenerfassung. Außerdem liefert sie die Abstammungsdaten beim Druck der Ergebnisse.
- *Sorten-Datenbank*. Damit werden Informationsbedürfnisse über eigene und fremde Sorten abgedeckt. Recherchen in Richtung Basissaatgut und Verkaufsmischungen werden möglich.
- *Urabstammungs-Datenbank*. Sie dient der besseren Gliederung des gesamten Materials, der Überwachung des Gen-Pools etc.

Die eingangs erwähnte Jahresversuchsdatei ist zwar mit der wichtigste Teil des Gesamtsystems, wird aber nicht durch DL/I-Entry verwaltet. Die Verknüpfung erfolgt durch Schlüssel im Rahmen der Versuchsbeschreibungs-Datenbank.

Hier ein Überblick über das Zusammenspiel der 6 Datenbanken (Abb. 2). Mit dem System "Integrierte Zuchtdatei" soll u.a. erreicht werden

- die Wiederverwendung (fehlerfrei) von einmal gespeicherten Daten
- die Umsetzung der Planung der Züchter in Arbeitsanweisungen
- die zwangsweise Erstellung der Zuchtbücher aus den Abläufen
- die Reduktion des Erfassungsaufwandes für die Versuchsverrechnung
- die Entlastung der wissenschaftlichen Mitarbeiter von Verwaltungstätigkeiten
- die einheitliche Speicherung von Informationen
- die bessere Auswertbarkeit der Informationen (direktes Hilfsmittel für die Selektion, generelle Kontrolle, langfristige Analysen).

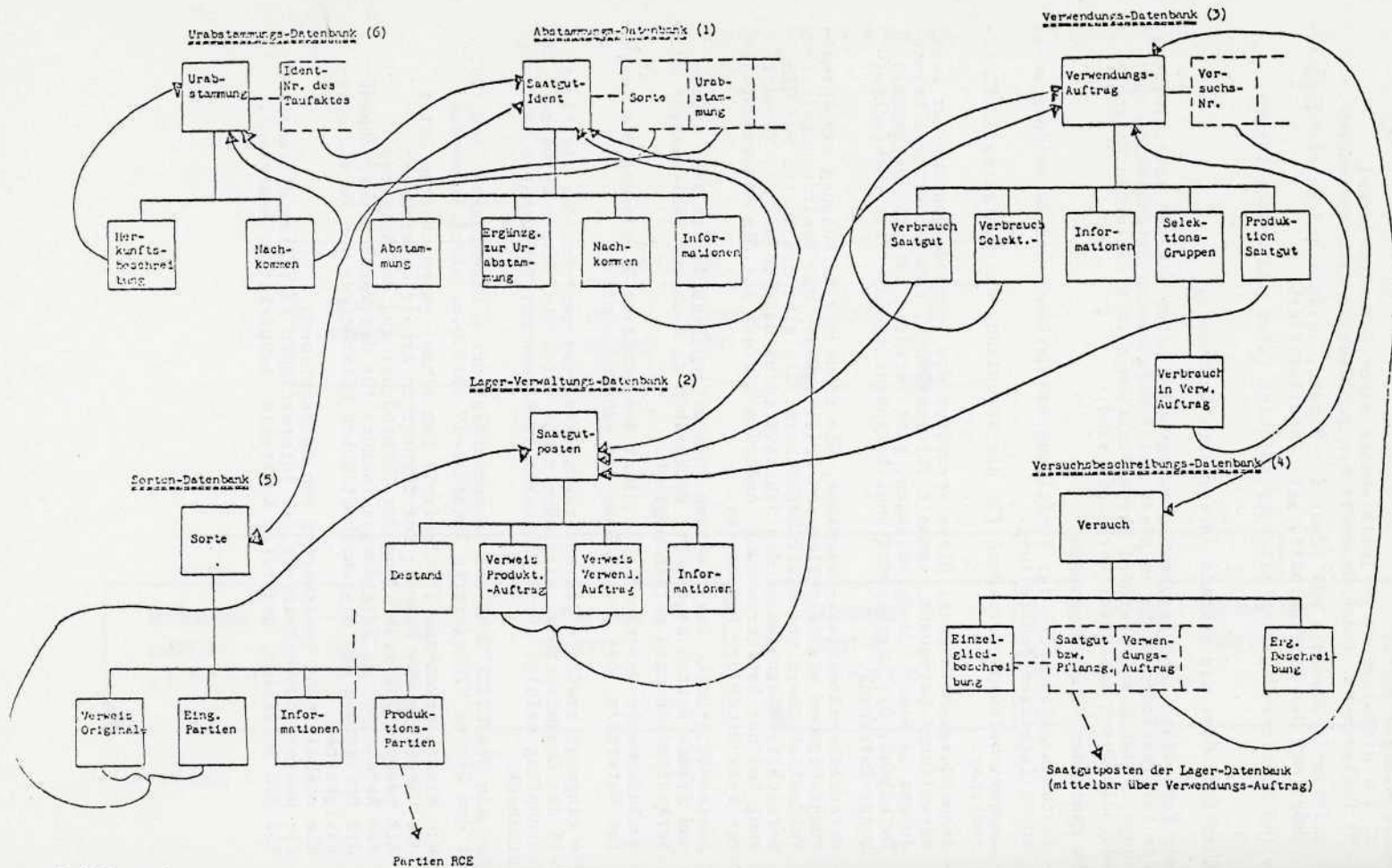


Abbildung 2

LITERATUR

HAUFE, W., 1969: Planung, Verrechnung und Auswertung von Feldversuchen für die Zuckerrübenzüchtung mit Hilfe von elektronischen Datenverarbeitungsanlagen. Vortrag Verb. D. Landw. Unters. und Forschungsanstalten, Lübeck; ferner Zucker, 22, 40-46.

HAUFE, W., 1973: Erfahrungsbericht über die Einführung der EDV bei der KWS. Vortrag Jahrestagung VDLUFA, Regensburg, 17.-22.9.1973; ferner Sonderheft 30/1, Landw. Forschung, 258-262.

HAUFE, W. und H. GEIDEL, 1974: Datenbankprobleme bei der Auswertung von Versuchsserien. EDV in Medizin und Biologie, 5, 13-20.

HAUFE, W., 1979: Datenorganisation in einem Zuchtbetrieb - die Erfassung, Auswertung und Archivierung von Versuchs-, Beobachtungs- und Produktionsdaten im Rahmen von Datenbanken -. Vortrag 30. Saatzuchtleitertagung, Gumpenstein, 27.-29.11.1979, ferner Bericht über diese Arbeitstagung.

DISKUSSIONSBEITRÄGE

BRUNDKE: Ist das System der KWS als Software-Paket übertragbar und damit verwendbar für andere Saatzuchtbetriebe?

HAUFE: Im Prinzip ja, in der Praxis wird es aber erhebliche Schwierigkeiten geben.

BAIER: Zur Wirtschaftlichkeit von Großrechenanlagen in der Pflanzenzüchtung. Nach der Investition entwickelt eine solche Rechenanlage oft ein kostspieliges Eigenleben.

HAUFE: Die Wirtschaftlichkeit sollte gegeben sein.

BAIER: In Zukunft wird die Hardware billiger, aber die Software teurer; damit sind die billigeren Kleinrechner in Zukunft auch nicht preisgünstiger.