



Ein temporales Data Warehouse für die Futtermitteldatenbank



Annelies Bracher
Agroscope Liebefeld-
Posieux (ALP)

in Zusammenarbeit mit:
Database Technology Research Group,
Universität Zürich
Prof. M. Böhlen, Dr. A. Taliun, F. Cafagna

29.03.11 – SVT-Tagung – Inforama Rütli Zollikofen



Hintergrund

- die schweiz. online Futtermitteldatenbank (FMDB) wurde 2007 aufgeschaltet
- Futterdaten von > 600 Futtermittel
- Nährstoffe und Nährwerte entsprechen aufbereiteten Mittelwerten von Futterproben
- aktualisierte Werte **überschreiben** "alte" Daten

➤ Ausbau der FMDB bis auf Stufe Einzelprobe

- Ziele:
- optimierte Datenbankpflege
 - erhöhter Informationsgehalt, differenzierte Abfragen
 - Statistik, Datenhistory, Trendanalysen, Monitoring ...

 - Forschungsarbeiten auf dem Gebiet der Datenbanktechnologie
 - Operationen (OLAP) auf temporalen Daten
 - SNF Projekt eingereicht



Umsetzung als temporales Data Warehouse

Definition

A data warehouse is a subject-oriented, integrated, time-variant, nonvolatile collection of data in support of management's decision-making process (Immon 1996).

Ein Data-Warehouse ist ein physischer Datenbestand, der eine integrierte Sicht auf die zugrundeliegenden Datenquellen ermöglicht (Zeh 2003).

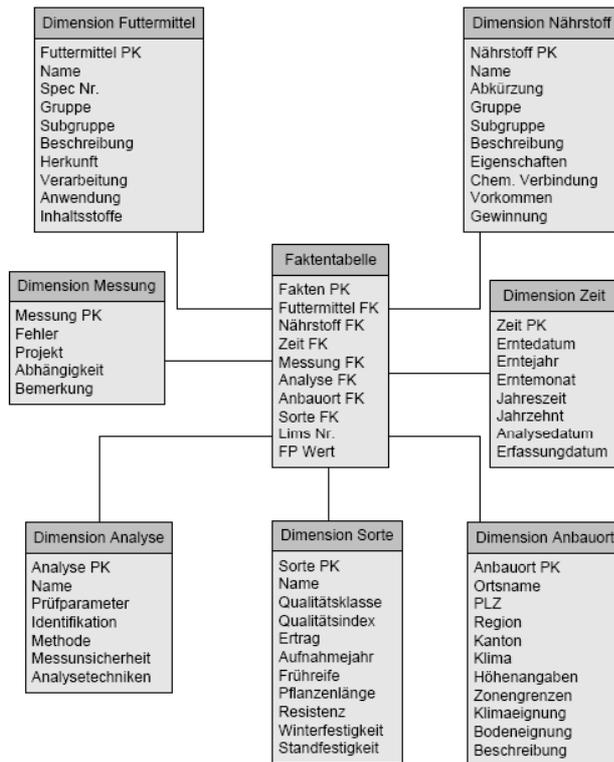
- multidimensionale Datenstruktur - biologisch, technisch, räumlich - die mit der Zeitdimension verknüpft ist
- Abfrage und Analyse **zeitabhängiger** Daten nach definierten Merkmalen



Vereinfachtes Datenmodell: Datenwürfel



Jede Achse beschreibt eine Merkmalsdimension, in die die Daten abgelegt sind. Die gewählte Achseneinheit wird als **Granularität** bezeichnet, welche die Aggregationsstufen bei Datenabfragen festlegt. Mehrdimensionale Datenstrukturen entsprechen einem Schneeflockenschema mit **Dimensionstabellen**.



Möglicher Prototyp für Data Warehouse

(Scheuner, 2010. Masterarbeit Uni Zürich)



Aggregation zeitabhängiger Daten: eine Herausforderung

Eigenheiten von Futterdaten:

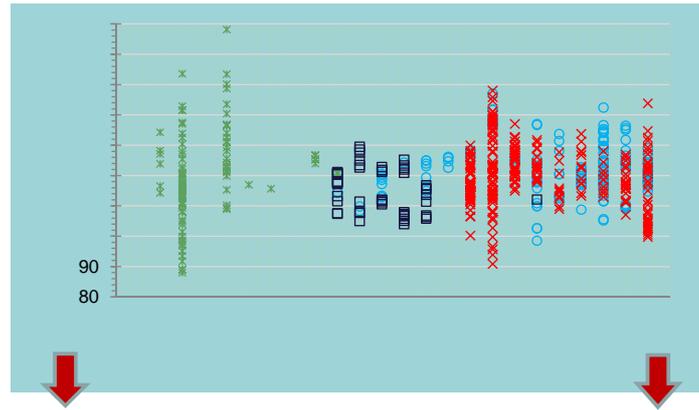
- multidimensional
- multigranular
- fuzzy
- Abhängigkeiten innerhalb Datensatz
- komplexe Zeitbezüge: Zeitpunkte – Zeitintervalle
- immer lückenhaft

Neue Datenbankoperationen

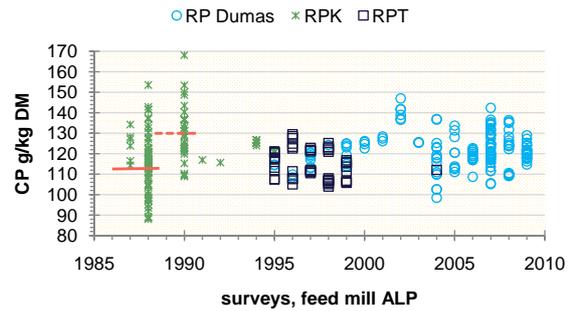
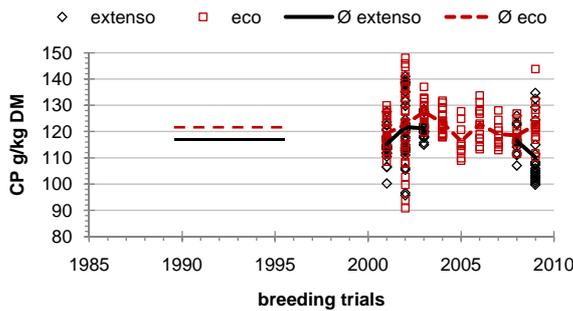
- **Intervallunifikation**: angleichen von Zeitintervallen, um nach Regeln der Algebra rechnen zu können
- **Zeitmodifiers**: **SEQ** (zu jedem Zeitpunkt), **NOW** (jetzt), **NSEQ** (unabhängig von der Zeit)
- parallele Zeitdimensionen: **transaction time**, **valid time**



Anwendungsbeispiele: RP in Gerste



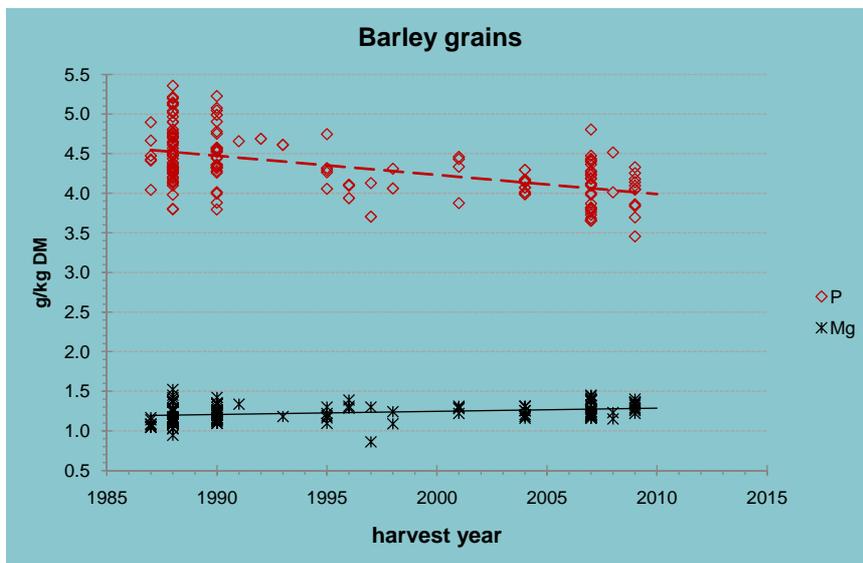
- Filterkriterien:
- Erntejahr
 - Datenquelle
 - Produktionsmethode
 - Analysenmethode



Zeittrend im Mineralstoffgehalt von Gerste



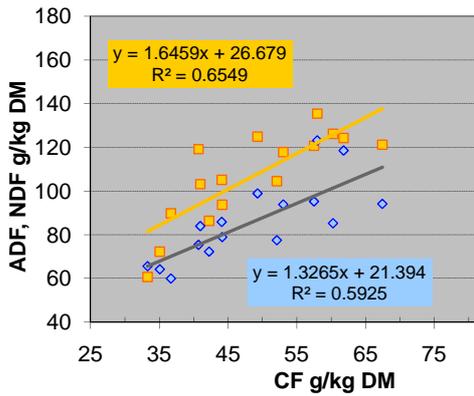
Relevanz des Zeitintervalls bei Mittelwertbildung





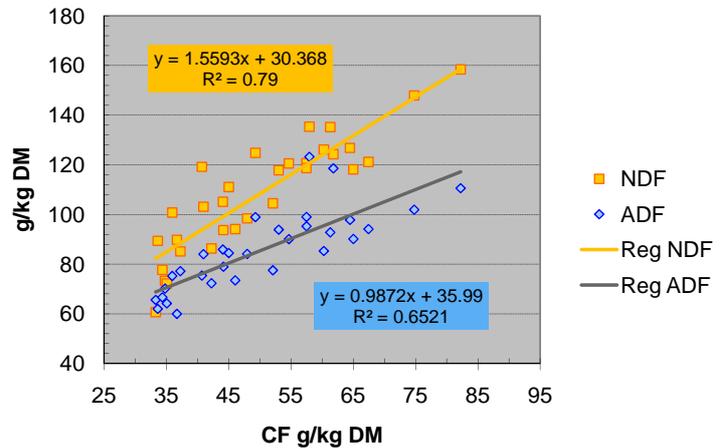
Zeitabhängige Korrelationen: Schätzung von ADF und NDF aus RF

Soybean meal update 2007



- transaction time
- valid time

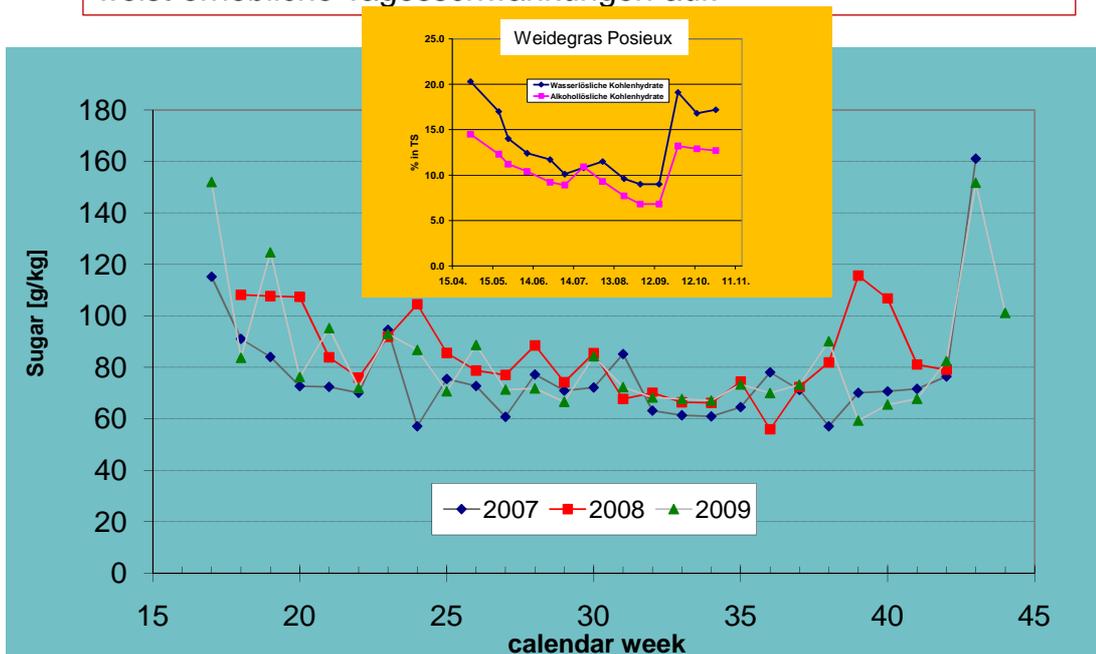
Soybean meal update 2010



Zeitgranularität: abhängig vom Futtertyp



Raufutter hat neben Jahreseinflüssen auch saisonale Muster und weist erhebliche Tagesschwankungen auf.

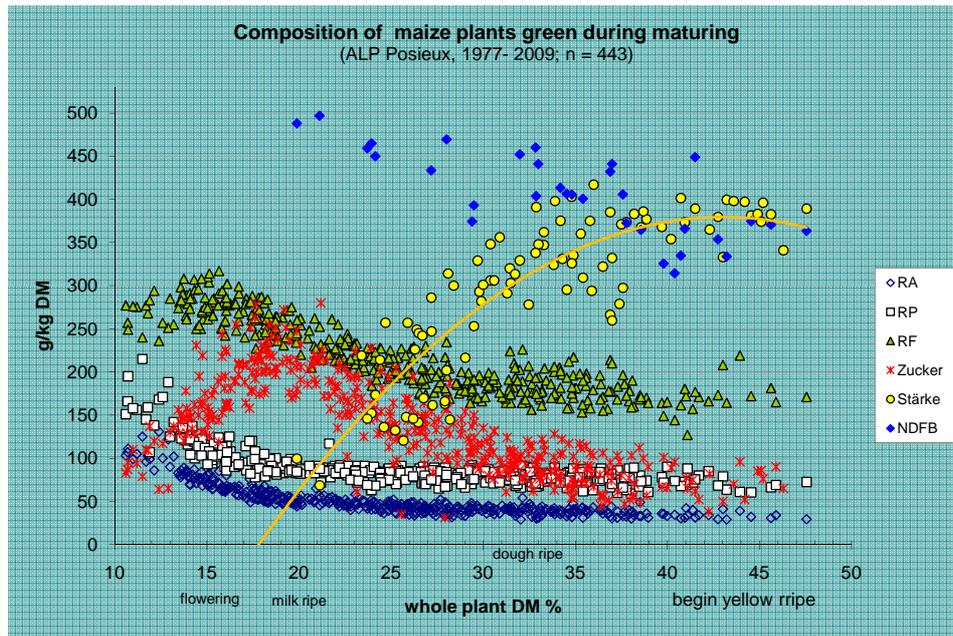




Physiologische Zeitaspekte: Reifungsprozesse



Verläufe sind dynamisch + fuzzy (fliessend)



Ein temporales Data Warehouse für die Futtermitteldatenbank SVT 29.3.2011
Annelies.Bracher@alp.admin.ch

11



Fazit

- ⇒ Die Weiterentwicklung der FMDB verbessert den Informationsgehalt von verfügbaren Futterdaten, die für praktische wie wissenschaftliche Fragestellungen als wertvolle Informationsquelle genutzt werden können.
- ⇒ Eine Erschliessung weiterer Datenquellen zur breiteren Abstützung der Datenbasis sinnvoll
- ⇒ Die Aktualisierung ist eine Daueraufgabe

Ein temporales Data Warehouse für die Futtermitteldatenbank SVT 29.3.2011
Annelies.Bracher@alp.admin.ch

12