



SAS[®] Forecast Server

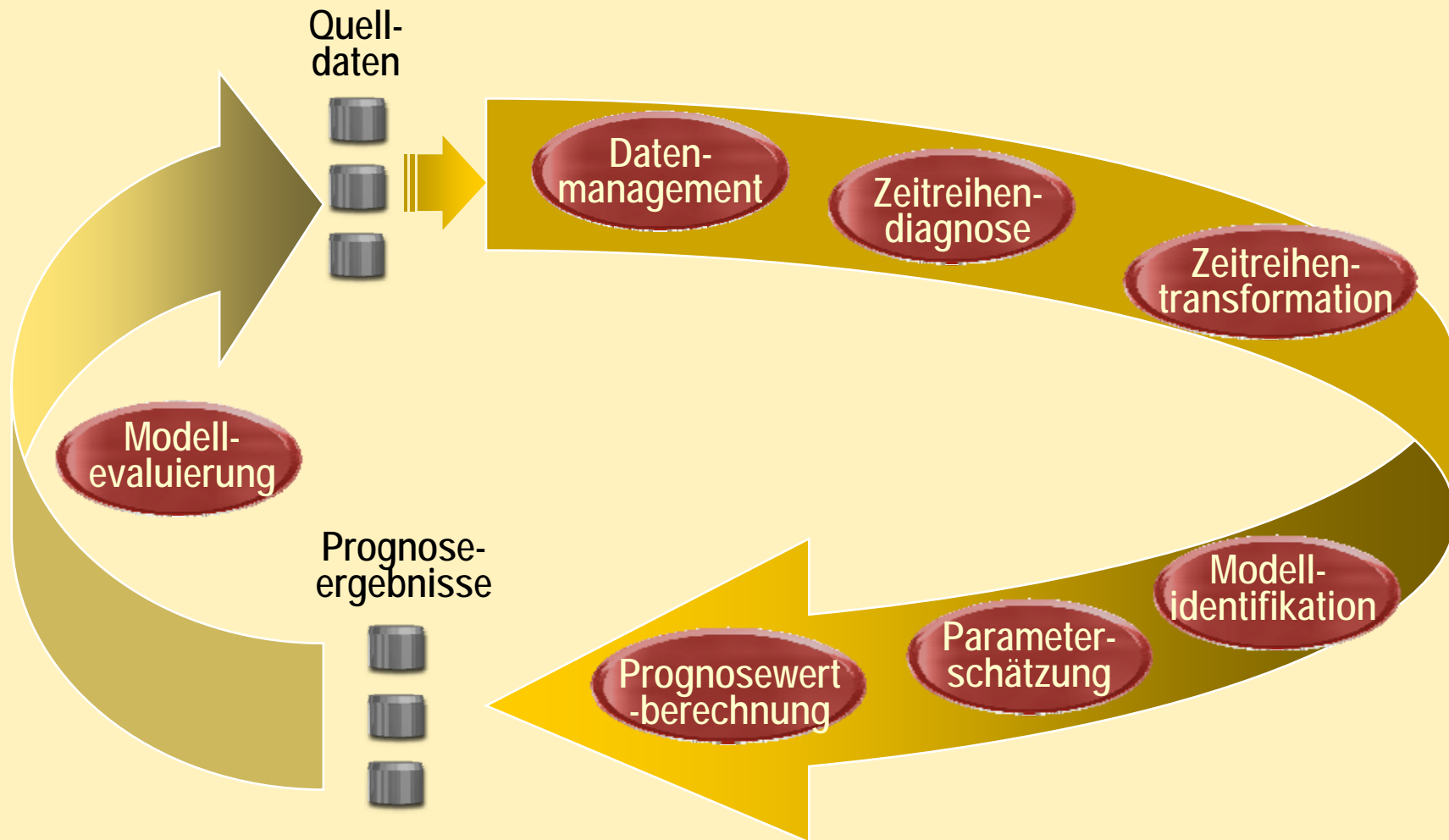
Eine neue Lösung für
automatisierte Prognosen
bei großen Datenmengen

Stefan Ahrens
Competence Center Analytic Intelligence
SAS Deutschland

Agenda

- Arbeitsschritte bei der Prognose
- Prognose mit SAS[®] – bisher
- Prognose mit dem neuen SAS[®] Forecast Server
- Zusammenfassung

Prognose als iterativer Prozess



Einzelne Arbeitsschritte (1)

- Datenmanagement
 - Datenimport und Konvertierung in Zeitreihen
 - Ersetzen fehlender Werte, und Ausreißerbehandlung
- Zeitreihendiagnostik
 - Visualisierung durch Zeitreihendiagramme
 - Autokorrelations-/Kreuzkorrelationsmuster, Signifikanztests
 - Saisonbereinigungsverfahren
- Zeitreihentransformation
 - Differenzenbildung (lineare Filter)
 - Varianzstabilisierung (z.B. Log-Transformation)
- Modellidentifikation
 - Modellansatz (ARIMA, Exponentielles Glätten, Regression ...)
 - Parametrisierung (MA- und AR-Terme, Transferfunktion ...)

Einzelne Arbeitsschritte (2)

- Schätzung der Modellparameter
 - Startwerte und Glättungsparameter bei exponentiellem Glätten
 - AR und MA Parameter bei ARIMA-Ansatz
 - Regressionskoeffizienten bei Regressionsansatz ...
- Generierung der Prognosewerte
 - Einsetzen der Daten in die Modellgleichung
 - Verwenden von Fortschreibungsformeln
- Evaluierung der Modellgüte
 - Allgemeine Fehlermaße (MAPE, RMSE ...)
 - Parsimonität (AIC, BIC, SBC ...)
 - Common Sense

Bisherige Prognosewerkzeuge in SAS

The screenshot shows a SAS code editor window titled 'Prognose.sas' with the following code:

```

/* Exponentielles Glätten */
proc forecast data=sashelp.air interval=month
  method=expo trend=2 lead=12
  out=out outfull outest=est;
  id date;
  var air;
run;

/* ARIMA-Methodologie */
proc arima data=sashelp.air;
  identify var=air (1,12) nlag=15;
run;
  estimate q=(1) (12) noconstant method=uls;
run;
  forecast out=b lead=24 id=date interval=month noprint;
run;
quit;

```

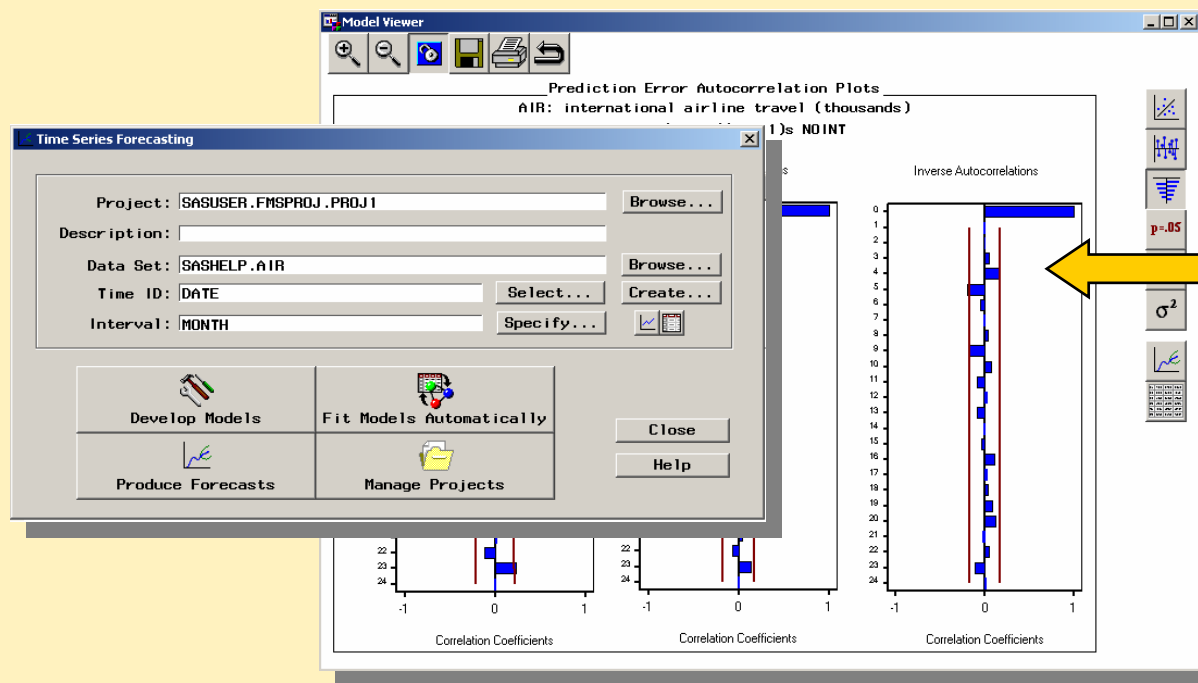
Overlaid on the right is the 'Autocorrelations' window, which displays a table of autocorrelation values for lags from -1 to 1. A yellow arrow points from the text 'Aufruf von Prozeduren im Programmierer' to the 'Std Error' column of the table.

Lag	Covariance	Correlation	Std Error
-1			0
0			0.087370
1			0.095389
2			0.096114
3			0.096856
4			0.097626
5			0.097916
6			0.097916
7			0.098450
8			0.100621
9			0.102300
10			0.102662
11			0.103982
12			0.104539
13			0.104539

Aufruf von Prozeduren im Programmierer

- SAS/ETS® (Proc FORECAST, Proc ARIMA ...)
- SAS® High Performance Forecasting (Proc HPF ...)
- Vollständige Kontrolle über alle Aspekte der Modellierung
- Erfordert Syntaxkenntnisse und gutes Methodenwissen

Bisherige Prognosewerkzeuge in SAS – Forts.



Interaktive Steuerung
über SAS[®] Time Series
Forecasting System

- Grafische Benutzeroberfläche als Code-Generator
- Integrierte Projektverwaltung
- Automatische Anpassung oder manuelle Modellentwicklung
- Diagnosewerkzeuge (ACF, PACF, White Noise Tests, Unit Root ...)

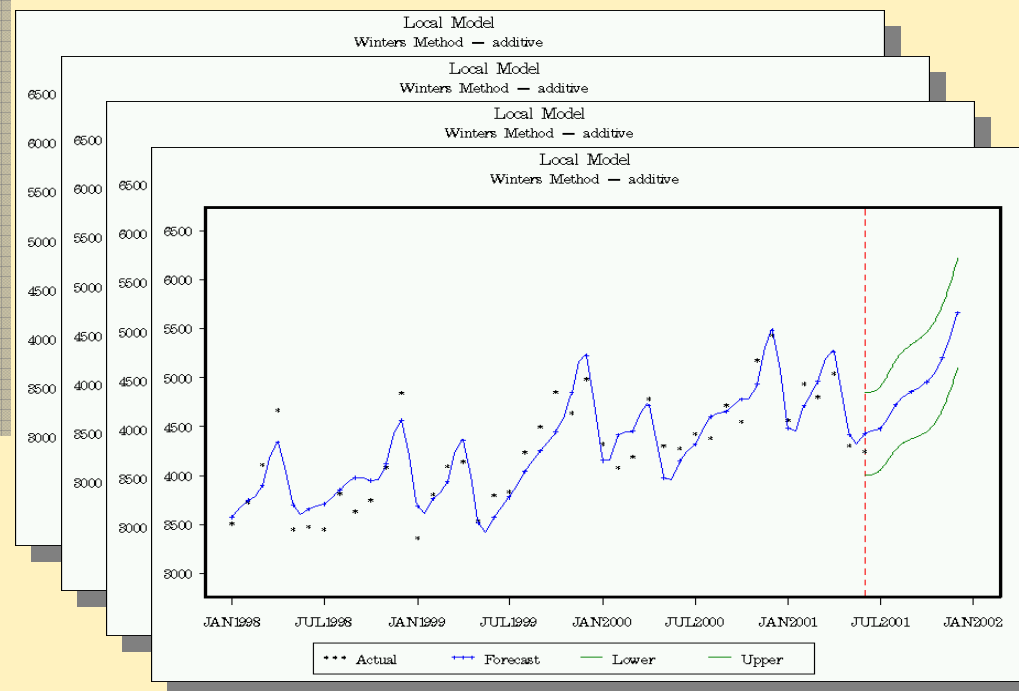
Herausforderungen in der Praxis

- Sicherstellung hinreichender Performance bei Prognosen für großen Datenmengen
- Minimieren des Zeitaufwands für Modellfindung bei Existenz vieler Zeitreihen
- Bequemes Einbinden externer Einflussfaktoren (Regressoren) in die Prognose
- Konsolidierung von Prognosewerten bei Existenz von Zeitreihenhierarchien
- Möglichkeit des manuellen Überschreibens der Prognosewerte
- Integration der Prognosen in operative und dispositive IT-Systeme

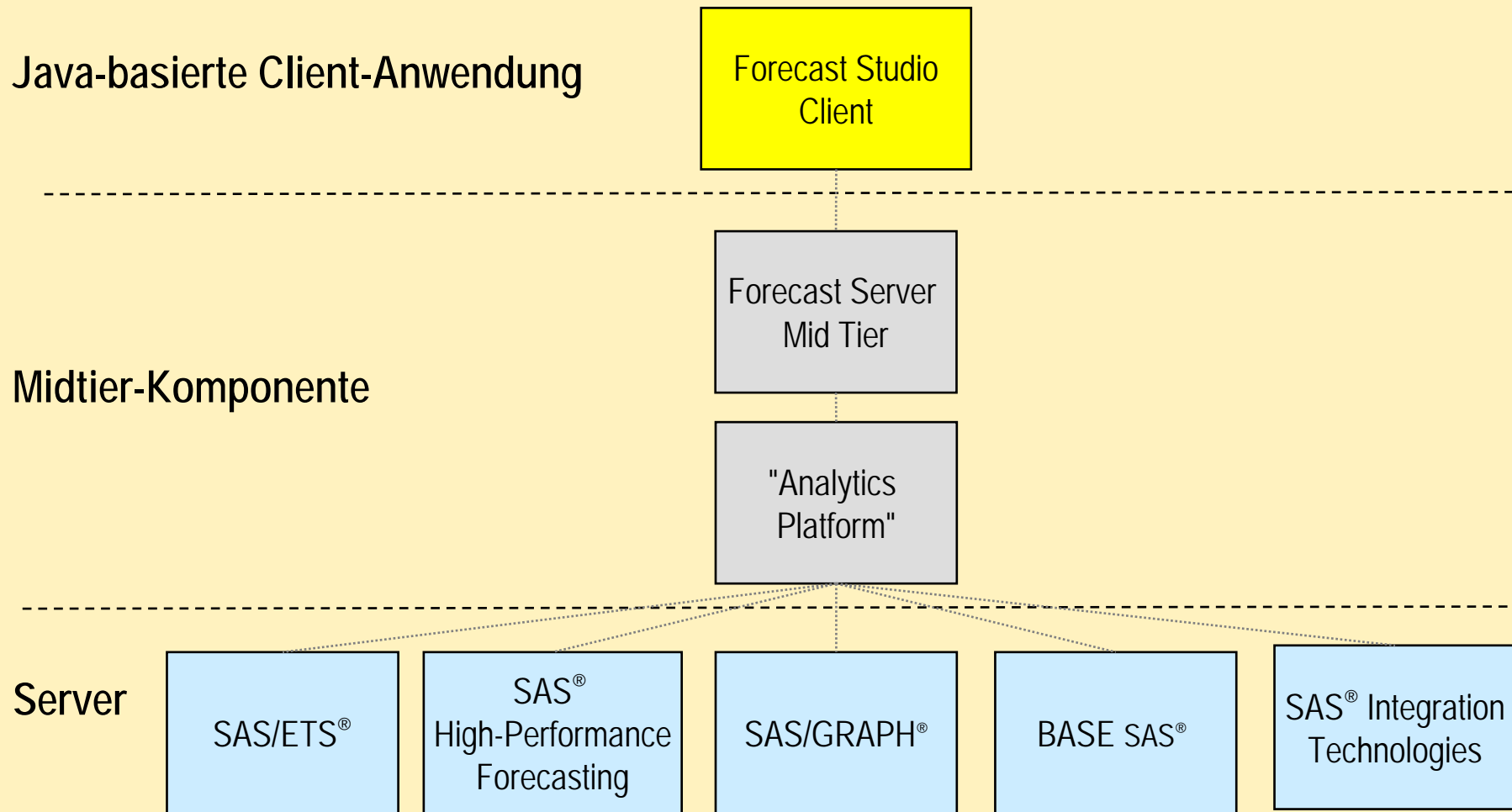
Was ist SAS[®] Forecast Server?

Eine neue leistungsfähige SAS Lösung zur **automatisierten** Durchführung von **qualitativ hochwertigen** Prognosen für große Datenmengen:

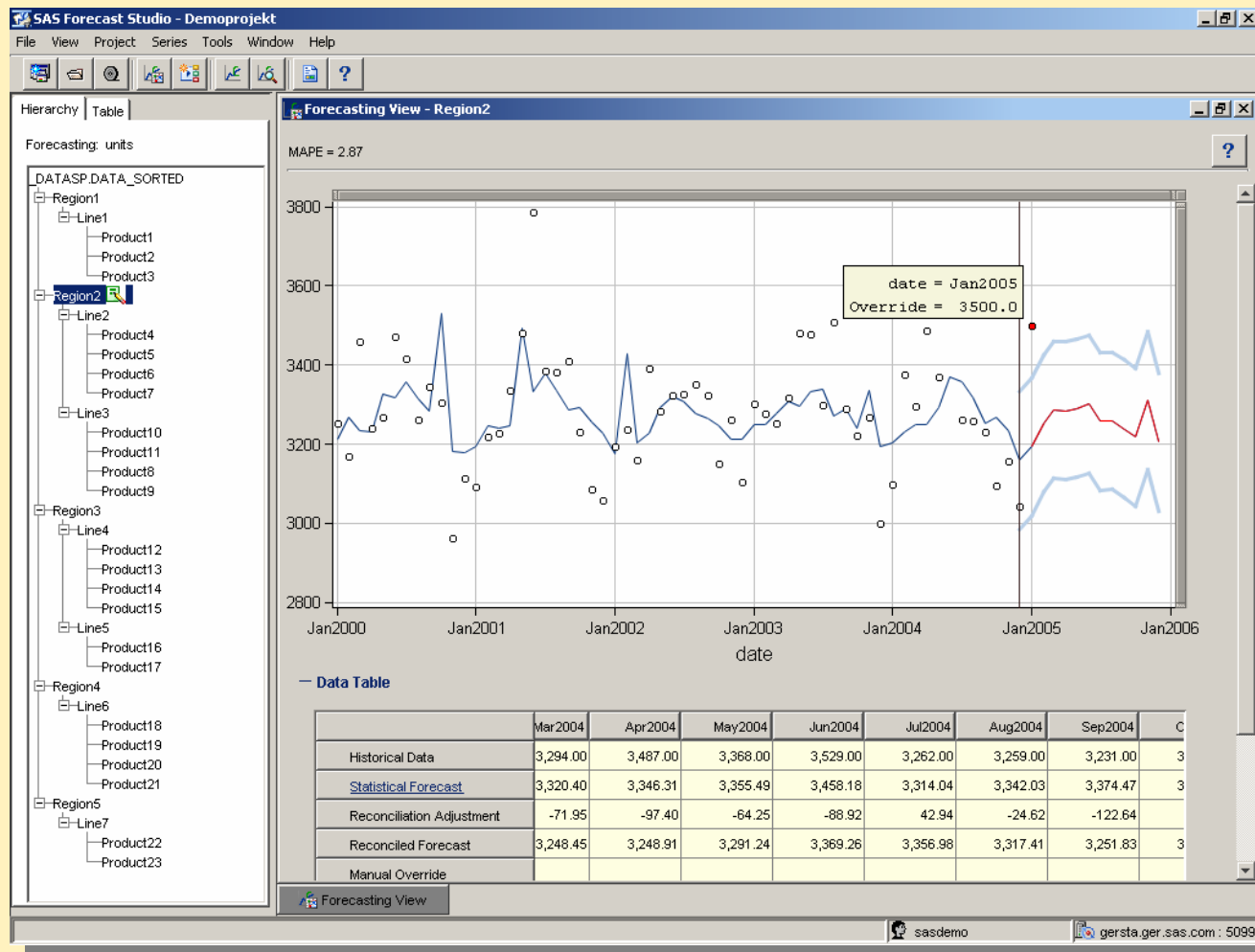
- Schnell und skalierbar
- Komfortabel bedienbar
- Flexibel einsetzbar
- Statistisch fundiert



Technologische Komponenten

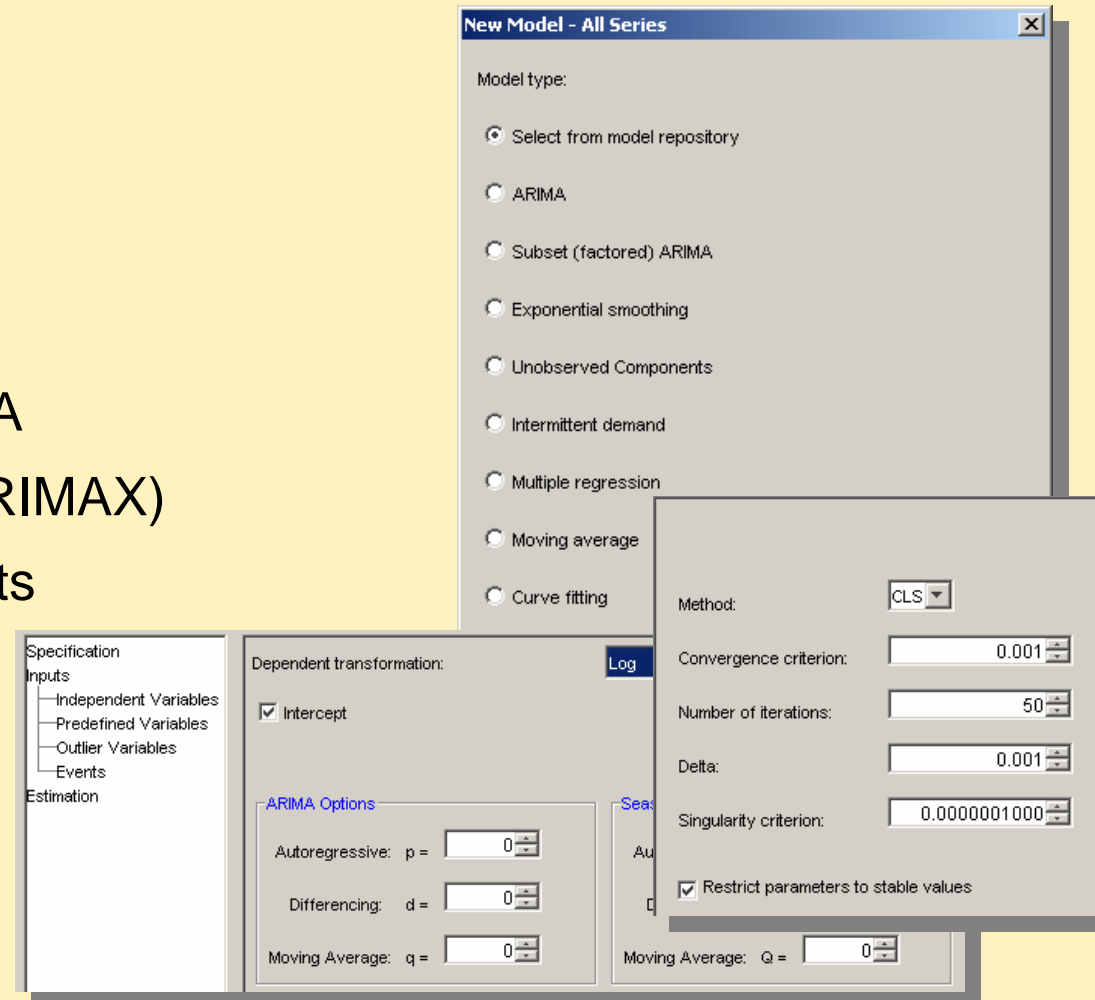


SAS[®] Forecast Studio als grafisches Frontend



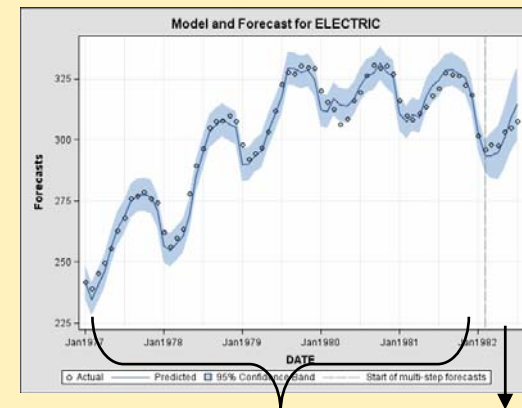
Große Auswahl statistischer Modelle

- Gleitende Durchschnitte
- Random Walk
- Exponentielles Glätten
- Multiple Regression
- Kurvenanpassung
- ARIMA, Factored ARIMA
- Transfer-Funktionen (ARIMAX)
- Unobserved Components
- Intermittent Demand
- ...



Automatisierte Möglichkeit der Modellsuche

- Automatische Gütemaß-Minimierung
 - MAPE, MSE, RMSE, R^2 , Adjusted R^2 ...
 - AIC, BIC, SBC ...
- Evaluierung via Holdout-Sample
- Ausreißerbereinigung über Signifikanztests
- Intermittency-Tests für sporadische Reihen
- Signifikanztests für Saison-Existenz
- Tests auf Stationarität (Unit Root-Tests)
- Transformation (Logarithmus, Quadratwurzel, Box-Cox ...)
- Vorläufige/endgültige Identifikation von AR und MA Ordnung

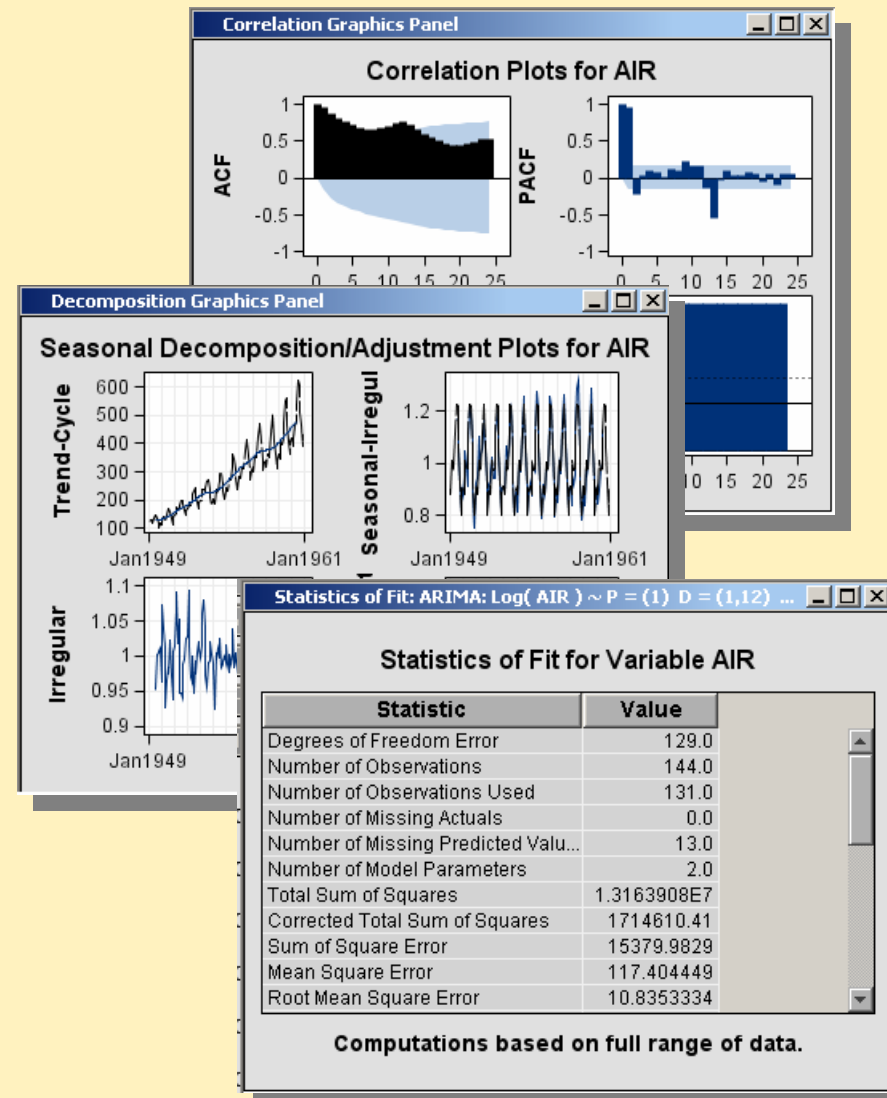


Anpassungsregion

Holdout-Sample

Diagnosewerkzeuge

- Zeitreihendekomposition
 - Saison-Komponente
 - Trendzyklus-Komponente
 - Irreguläre Komponente
- Korrelogramme
 - Autokorrelation (ACF)
 - Partielle Autokorrelation (PACF)
 - Inverse Autokorrelation (IACF)
 - Kreuzkorrelation (CCF)
- White Noise Tests
- Residuenverteilung
- Anpassungsstatistiken
- ...



Einbinden relevanter externer Einflussgrößen

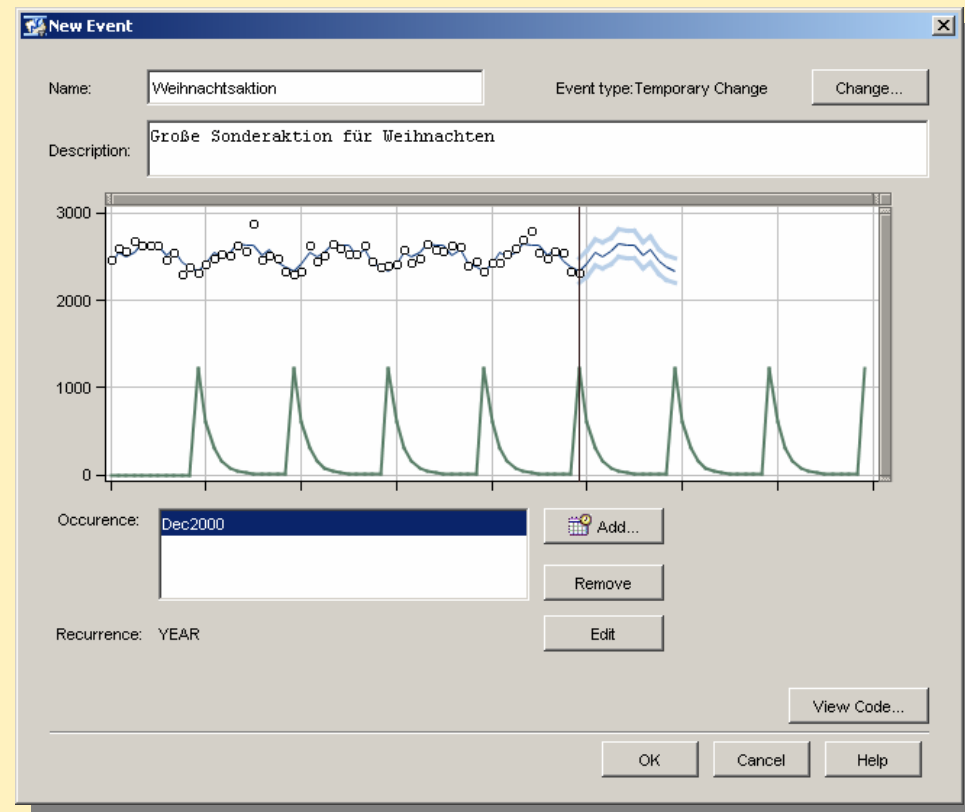
Modelle mit Regressoren und Transfer-Funktion verbessern
Prognosegenauigkeit durch flexible Berücksichtigung von ...

... Kalenderereignissen ...

- Aktionen, Werbe-Flights
- Sportereignisse
- Feiertage, Schulferien
- Gesetzesänderungen ...

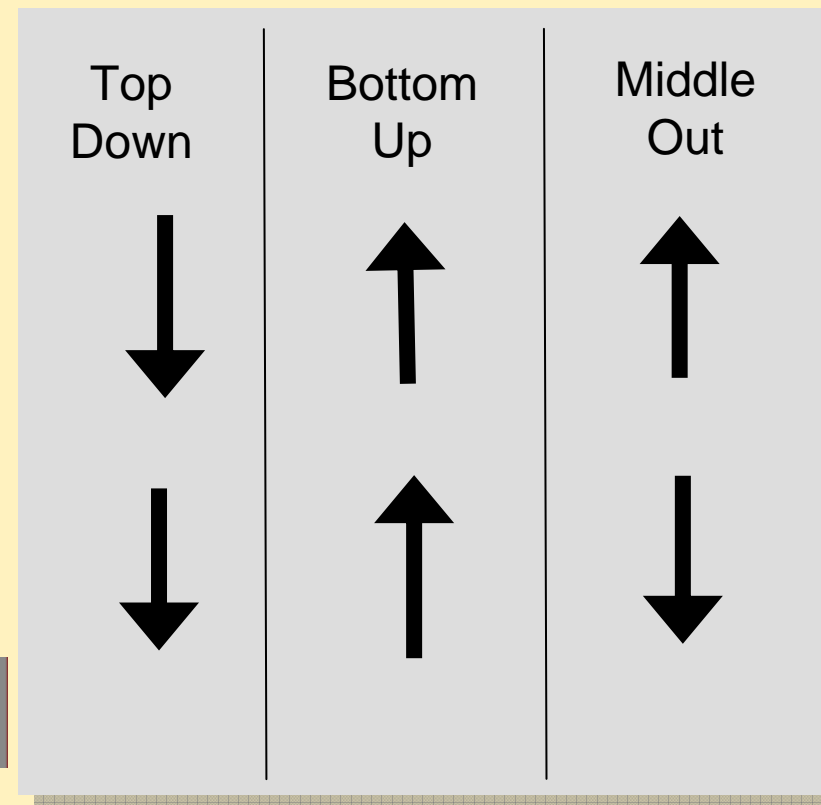
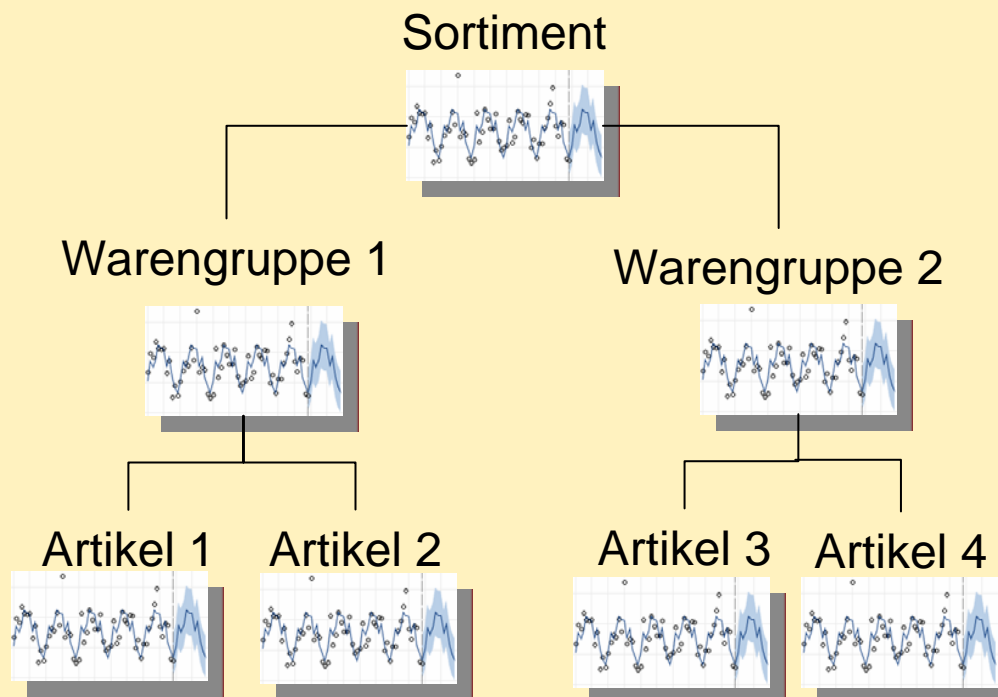
... und stetigen Variablen!

- Wetterdaten
- Preisindizes
- Werbeaufwendungen ...



Konsolidierung von Prognosehierarchien

Abgleich von Prognosewerten über mehrstufige Hierarchien:



Überschreiben von Prognosewerte

Komfortable Eingabemasken für manuelle Korrektur und Kommentare:

1. Select one or more cells in any one row

	Jan2005	Feb2005	Mar2005	Apr2005	May2005	Jun2005	Jul2005
Reconciled Forecast	443.70	455.63	460.46	450.40	468.64	451.69	447.73
Manual Override	480.00	500.00					
Final Forecast	443.70	455.63	460.46	450.40	468.64	451.69	447.73

2. Specify the changes to make to the selected values

Increase : %

Set to :

Set each selected cell to this value

Spread this value equally across selected cells

Spread this value proportionately across selected cells

3. Apply the changes

1 Jan2002 Jan2003 Jan2004 Jan2005 Jan2006
date

Weitere Highlights

- Projektverwaltung mit automatischer Prognoseaktualisierung
- Definieren von Filtern für Exception Reporting
- Stored Processes für Berichte oder Ergebnisexport
- Oberfläche als Code-Generator für Batch Jobs
- ...

The screenshot displays three overlapping SAS windows:

- Forecasting: units**: A window with a 'Show:' dropdown set to 'Mape' and a 'Refresh' button. Below is a table with columns: regionName, productLine, productName, MAPE /, and Exceptions.
- Reports**: A window showing a tree view with 'Samples' and 'Generate Shipment Forecasts'.
- SAS Code**: A window showing SAS macro code for forecasting.

regionName	productLine	productName	MAPE /	Exceptions
Region3	Line5	Product17	6.00	✘
Region4	Line6	Product18	6.06	✘
Region3	Line4	Product15	6.07	✘
Region4	Line6	Product21	6.10	✘
Region2	Line5	Product18	6.08	✘
Region2	Line6	Product17	6.08	✘
Region1	Line5	Product17	6.08	✘
Region4	Line6	Product17	6.08	✘
Region2	Line5	Product17	6.08	✘
Region3	Line5	Product17	6.08	✘
Region1	Line5	Product17	6.08	✘
Region2	Line5	Product17	6.08	✘

```

2346 %hpfrecon(disdata=_HPF2.finalfor, aggdata=_HPF1.finalfor, outagg=_HPF1)
2347   id=date, interval=MONTH,
2348   by=regionName productLine, direction=bu);
2349 *-----
2350 * reconcile overrides _hpf1 up to _hpf0
2351 *-----
2352 %hpfrecon(disdata=_HPF1.finalfor, aggdata=_HPF0.finalfor, outagg=_HPF1)
2353   id=date, interval=MONTH,
  
```

Fazit und Ausblick

- SAS[®] setzt mit dieser Lösung neue Maßstäbe
- Prognosen auch für große Datenmengen handhabbar
- Bietet dem Analytiker Entlastung von Routinearbeiten
- Ausblick für mögliche Weiterentwicklungen
 - Prognose bei kurzen Lebenszyklen
 - Prognose für neue Produkte
 - Zeitreihen-Clustering
 - What-If-Szenarien
 - ...





The Power to Know.