

*IBM SPSS Forecasting 23*

**IBM**

**Hinweis**

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die Informationen unter „Bemerkungen“ auf Seite 67 gelesen werden.

**Produktinformation**

Diese Ausgabe bezieht sich auf Version 23, Release 0, Modifikation 0 von IBM SPSS Statistics und alle nachfolgenden Releases und Modifikationen, bis dieser Hinweis in einer Neuausgabe geändert wird.

Diese Veröffentlichung ist eine Übersetzung des Handbuchs

*IBM SPSS Forecasting 23,*

herausgegeben von International Business Machines Corporation, USA

© Copyright International Business Machines Corporation 2014

Informationen, die nur für bestimmte Länder Gültigkeit haben und für Deutschland, Österreich und die Schweiz nicht zutreffen, wurden in dieser Veröffentlichung im Originaltext übernommen.

Möglicherweise sind nicht alle in dieser Übersetzung aufgeführten Produkte in Deutschland angekündigt und verfügbar; vor Entscheidungen empfiehlt sich der Kontakt mit der zuständigen IBM Geschäftsstelle.

Änderung des Textes bleibt vorbehalten.

Herausgegeben von:

TSC Germany

Kst. 2877

Dezember 2014

---

# Inhaltsverzeichnis

## Kapitel 1. Einführung in Zeitreihen . . . . 1

Zeitreihendaten . . . . .	1
Transformieren von Daten . . . . .	2
Schätz- und Validierungsperioden . . . . .	2
Erstellen von Modellen und Vorhersagen . . . . .	3

## Kapitel 2. Zeitreihenmodellierung . . . . 5

Angeben von Optionen für den Expert Modeler . . . . .	7
Modellauswahl und Ereignisspezifikation. . . . .	7
Behandeln von Ausreißern mit Expert Modeler . . . . .	8
Exponentielles Glätten: Benutzerdefinierte Modelle . . . . .	8
Benutzerdefinierte ARIMA-Modelle. . . . .	9
Modellspezifikation für benutzerdefinierte ARI- MA-Modelle . . . . .	9
Transferfunktionen in benutzerdefinierten ARI- MA-Modellen . . . . .	10
Ausreißer in benutzerdefinierten ARIMA-Model- len . . . . .	11
Ausgabe . . . . .	12
Statistiken und Vorhersagetabellen. . . . .	12
Diagramme . . . . .	13
Einschränken der Ausgabe auf die am besten/ schlechtesten angepassten Modelle . . . . .	14
Speichern von Modellvorhersagen und Modellspezi- fikationen . . . . .	14
Optionen . . . . .	15
Zusätzliche Funktionen beim Befehl TSMODEL . . . . .	16

## Kapitel 3. Zuweisen von Zeitreihenmo- dellen . . . . . 17

Ausgabe . . . . .	19
Statistiken und Vorhersagetabellen. . . . .	19
Diagramme . . . . .	20
Einschränken der Ausgabe auf die am besten/ schlechtesten angepassten Modelle . . . . .	21
Speichern von Modellvorhersagen und Modellspezi- fikationen . . . . .	21
Optionen . . . . .	22
Zusätzliche Funktionen beim Befehl TSAPPLY. . . . .	23

## Kapitel 4. Saisonale Zerlegung . . . . 25

Saisonale Zerlegung: Speichern . . . . .	26
Zusätzliche Funktionen beim Befehl SEASON . . . . .	26

## Kapitel 5. Spektraldiagramme . . . . 27

Zusätzliche Funktionen beim Befehl SPECTRA . . . . .	29
--	----

## Kapitel 6. Temporale kausale Modelle 31

Erstellen eines temporalen kausalen Modells . . . . .	32
---	----

Zeitreihen für die Modellierung . . . . .	32
Auswählen von Dimensionswerten . . . . .	33
Beobachtungen . . . . .	34
Zeitintervall für die Analyse. . . . .	35
Aggregation und Verteilung . . . . .	35
Fehlende Werte . . . . .	36
Allgemeine Datenoptionen . . . . .	37
Allgemeine Erstellungsoptionen . . . . .	37
Anzuzeigende Zeitreihe . . . . .	37
Ausgabeoptionen . . . . .	39
Schätzperiode . . . . .	41
Vorhersage . . . . .	41
Speichern . . . . .	42
Interaktive Ausgabe . . . . .	42

## Kapitel 7. Anwenden temporaler kausaler Modelle . . . . . 45

Anwenden temporaler kausaler Modelle. . . . .	45
Vorhersage temporaler kausaler Modelle. . . . .	45
Verwenden von "Vorhersage temporaler kausaler Modelle" . . . . .	45
Modellparameter und Vorhersagen . . . . .	46
Allgemeine Optionen . . . . .	47
Anzuzeigende Zeitreihe . . . . .	47
Ausgabeoptionen . . . . .	48
Speichern . . . . .	50
Temporale kausale Modellszenarios . . . . .	51
Ausführen von "Temporale kausale Modellszena- rios". . . . .	51
Definieren der Szenarioperiode. . . . .	52
Hinzufügen von Szenarios und Szenariogruppen . . . . .	53
Optionen . . . . .	57

## Kapitel 8. Maße für die Anpassungsgü- te . . . . . 59

## Kapitel 9. Ausreißertypen. . . . . 61

## Kapitel 10. Übersicht über ACF-/PACF- Diagramme . . . . . 63

## Bemerkungen. . . . . 67

Marken. . . . .	68
-----------------	----

## Index . . . . . 71



---

## Kapitel 1. Einführung in Zeitreihen

Eine **Zeitreihe** ist eine Reihe von Beobachtungen, die gewonnen werden, indem eine einzelne Variable regelmäßig über einen Zeitraum beobachtet wird. In einer Zeitreihe mit Bestandsdaten beispielsweise, können die Beobachtungen tägliche Bestandsniveaus für mehrere Monate repräsentieren. Eine Zeitreihe, die den Marktanteil eines Produkts anzeigt, kann aus den wöchentlichen Marktanteilen über den Zeitraum einiger Jahre bestehen. Eine Zeitreihe aus den Gesamtverkaufszahlen kann aus einer Beobachtung pro Monat über den Zeitraum vieler Jahre bestehen. Was alle diese Beispiele gemeinsam haben, ist, dass eine Variable über einen bestimmten Zeitraum in regelmäßigen, bekannten Intervallen beobachtet wurde. Daher ist die Art der Daten für eine typische Zeitreihe eine einzelne Sequenz oder Liste von Beobachtungen, die in regelmäßigen Abständen durchgeführte Messungen repräsentieren.

Tabelle 1. Tägliche Bestandszeitreihe

Zeit	Woche	Tag	Bestandsniveau
$t_1$	E	Montag	160
$t_2$	E	Dienstag	135
$t_3$	E	Mittwoch	129
$t_4$	E	Donnerstag	122
$t_5$	E	Freitag	108
$t_6$	Z	Montag	150
		...	
$t_{60}$	12	Freitag	120

Eines der wichtigsten Ziele bei der Durchführung einer Zeitreihenanalyse ist der Versuch, die zukünftigen Werte der Zeitreihe vorherzusagen. Ein Modell der Zeitreihe, das die früheren Werte erklärte, kann möglicherweise auch vorhersagen, ob und wie stark die nächsten Werte zu- bzw. abnehmen. Die Fähigkeit, erfolgreich derartige Vorhersagen zu treffen, ist offensichtlich für jeden Betrieb und alle wissenschaftlichen Bereiche von großer Bedeutung.

---

## Zeitreihendaten

### Spaltenbasierte Daten

Jedes Zeitreihenfeld enthält die Daten für eine einzige Zeitreihe. Dies ist die traditionelle Struktur für Zeitreihendaten, wie sie von der Prozedur "Zeitreihenmodellierung", der Prozedur "Saisonale Zerlegung" und der Prozedur "Spektraldiagramme" verwendet wird. Um beispielsweise eine Zeitreihe im Dateneditor zu definieren, klicken Sie auf die Registerkarte **Variablenansicht** und geben Sie in eine beliebige leere Zeile einen Variablennamen ein. Jede Beobachtung in einer Zeitreihe entspricht einem Fall (einer Zeile im Dateneditor).

Wenn Sie eine Tabelle mit Zeitreihendaten öffnen, sollte jede Zeitreihe in einer Spalte der Tabelle angeordnet sein. Wenn in Ihrer Tabelle bereits Zeitreihen in Zeilen angeordnet sind, können Sie die Tabelle dennoch öffnen und mit der Funktion **Transponieren** im Menü **Daten** die Zeilen in Spalten umwandeln.

### Mehrdimensionale Daten

Bei mehrdimensionalen Daten enthält jedes Zeitreihenfeld die Daten für mehrere Zeitreihen. Innerhalb eines bestimmten Felds werden separate Zeitreihen dann durch ein Set von Werten für kategoriale Felder ermittelt, die als *Dimensionsfelder* bezeichnet werden.

Beispielsweise können Verkaufsdaten für verschiedene Regionen und Marken in einem einzigen Feld *sales* gespeichert werden, sodass die Dimensionen in diesem Fall *region* und *brand* wären. Jede Kombination aus *region* und *brand* kennzeichnet eine bestimmte Zeitreihe für *sales*. In der folgenden Tabelle z. B. definieren die Datensätze mit dem Wert 'north' für *region* und dem Wert 'brandX' für *brand* eine einzelne Zeitreihe.

Tabelle 2. Mehrdimensionale Daten

date	region	brand	sales
01/01/2014	north	brandX	82350
01/01/2014	north	brandY	86380
01/01/2014	south	brandX	91375
01/01/2014	south	brandY	70320
01/02/2014	north	brandX	83275
01/02/2014	north	brandY	85260
01/02/2014	south	brandX	94760
01/02/2014	south	brandY	69870

**Anmerkung:** Daten, die aus OLAP-Würfeln importiert werden, z. B. aus IBM® Cognos TM1, werden als mehrdimensionale Daten dargestellt.

---

## Transformieren von Daten

Eine Reihe der im Core-System zur Verfügung stehenden Prozeduren für Datentransformationen können bei der Analyse von Zeitreihen nützlich sein. Diese Transformationen gelten nur für spaltenbasierte Daten, wobei jedes Zeitreihenfeld die Daten für eine einzige Zeitreihe enthält.

- Die Prozedur "Datum definieren" (im Menü "Daten") erzeugt Datumvariablen zum Erstellen von Periodizität und zum Unterscheiden zwischen historischen Perioden, Validierungsperioden und Vorhersageperioden. Forecasting wurde für die Arbeit mit den durch die Prozedur "Datum definieren" erstellten Daten entwickelt.
- Die Prozedur "Zeitreihen erstellen" (im Menü "Transformieren") erstellt neue Zeitreihenvariablen als Funktionen bestehender Zeitreihenvariablen. Dazu gehören Funktionen, die benachbarte Beobachtungen zum Glätten, zur Durchschnittsberechnung und zur Differenzierung verwenden.
- Die Prozedur "Fehlende Werte ersetzen" (im Menü "Transformieren") ersetzt system- und benutzerdefiniert fehlende Werte durch Schätzwerte auf der Grundlage einer von mehreren möglichen Methoden. Fehlende Daten am Anfang oder Ende einer Zeitreihe stellen kein größeres Problem dar. Sie verkürzen nur die brauchbare Länge der Zeitreihe. Lücken im Inneren einer Zeitreihe (*eingebettete* fehlende Daten) können ein viel schwerwiegenderes Problem darstellen.

Ausführliche Informationen zu Datentransformationen für Zeitreihen finden Sie im Benutzerhandbuch zum Core-System.

---

## Schätz- und Validierungsperioden

Häufig ist es sinnvoll, die Zeitreihe in eine *Schätzperiode* bzw. *historische* Periode und eine *Validierungsperiode* aufzuteilen. Sie entwickeln ein Modell auf der Grundlage der Beobachtungen in der Schätzperiode (historischen Periode) und testen dann in der Validierungsperiode, wie gut es funktioniert. Wenn Sie das Modell zum Vorhersagen von bekannten Punkten (die Punkte in der Validierungsperiode) verwenden, erhalten Sie einen Eindruck von der Güte der Vorhersagen des Modells.

Die Fälle in der Validierungsperiode werden als Holdout-Fälle bezeichnet, da Sie aus dem Erstellungsprozess des Modell ausgeschlossen (und lediglich geprüft) werden. Wenn Sie feststellen, dass das Modell hinreichende Vorhersagen liefert, können Sie die Schätzperiode um die Holdout-Fälle erweitern und anschließend das endgültige Modell erstellen.

---

## Erstellen von Modellen und Vorhersagen

Das Zusatzmodul "Forecasting" bietet zwei Prozeduren zum Erstellen von Modellen zum Treffen von Vorhersagen.

- Die Prozedur in Kapitel 2, „Zeitreihenmodellierung“, auf Seite 5 erstellt Modelle für Zeitreihen und ermöglicht Vorhersagen. Sie enthält den Expert Modeler, der automatisch das beste Modell für jede der Zeitreihen bestimmt. Erfahrene Analysten, die ein größeres Maß an Kontrolle wünschen, finden im Expert Modeler außerdem Tools zum Erstellen von benutzerdefinierten Modellen.
- Die Prozedur in Kapitel 3, „Zuweisen von Zeitreihenmodellen“, auf Seite 17 wendet vorhandene, mit der Prozedur "Zeitreihenmodellierung" erstellte Zeitreihenmodelle auf das aktive Dataset an. Hierbei können Sie Vorhersagen für Zeitreihen berechnen, für neue oder überarbeitete Daten verfügbar sind, ohne die Modelle erneut zu erstellen. Wenn Sie Grund zu der Annahme haben, dass sich ein Modell verändert hat, können Sie es mit der Prozedur "Zeitreihenmodellierung" erneut erstellen.



---

## Kapitel 2. Zeitreihenmodellierung

Die Prozedur "Zeitreihenmodellierung" schätzt Modelle für die exponentielle Glättung, die univariate ARIMA (Autoregressive Integrated Moving Average) und die multivariate ARIMA (oder Transferfunktionsmodelle) für Zeitreihen und erstellt Vorhersagen. Die Prozedur enthält den Expert Modeler, der für eine oder mehrere abhängige Variablenreihen automatisch das jeweils am besten angepasste Modell für die ARIMA oder die exponentielle Glättung zu ermitteln und zu schätzen versucht, sodass das geeignete Modell nicht mehr nach dem Prinzip von Versuch und Irrtum ermittelt werden muss. Sie können auch ein benutzerdefiniertes ARIMA-Modell oder ein Modell der exponentiellen Glättung angeben.

**Beispiel:** Als Produktmanager ist es Ihre Aufgabe, für 100 verschiedene Produkte die Verkaufsstückzahlen und Verkaufserlöse für den Folgemonat vorherzusagen. Sie verfügen jedoch nur über geringe oder gar keine Erfahrungen in der Zeitreihenmodellierung. Die historischen Verkaufsstückzahlen für alle 100 Produkten sind in einem einzigen Excel-Arbeitsblatt gespeichert. Nachdem Sie das Arbeitsblatt in IBM SPSS Statistics geöffnet haben, fordern Sie im Expert Modeler Vorhersagen für den nächsten Monat an. Der Expert Modeler ermittelt für jedes Produkt das beste Modell für die Verkaufsstückzahlen und erstellt anhand dessen die Vorhersagen. Der Expert Modeler kann mehrere Eingabereihen verarbeiten, sodass Sie die Prozedur nur einmal ausführen müssen, um Vorhersagen für sämtliche Produkte zu erhalten. Wenn Sie die Vorhersagen im aktiven Dataset speichern, können Sie die Ergebnisse problemlos zurück nach Excel exportieren.

**Statistiken.** Maße für die Anpassungsgüte:  $R$ -Quadrat für stationären Teil,  $R$ -Quadrat ( $R^2$ ), Wurzel der mittleren Fehlerquadratsumme (RMSE), mittlerer absoluter Fehler (MAE), mittlerer absoluter Fehler in Prozent (MAPE), maximaler absoluter Fehler (MaxAE), maximaler absoluter Fehler in Prozent (MaxAPE), normalisiertes bayessches Informationskriterium (BIC). Residuen: Autokorrelationsfunktion, partielle Autokorrelationsfunktion, Ljung-Box- $Q$ . Für ARIMA-Modelle: ARIMA-Ordnungen für abhängige Variablen, Transferfunktionsordnungen für unabhängige Variablen und Ausreißerschätzungen. Außerdem Glättungsparameterschätzungen für Modelle mit exponentiellem Glätten.

**Diagramme.** Diagramme für alle Modelle: Histogramme von  $R$ -Quadrat für stationären Teil,  $R$ -Quadrat ( $R^2$ ), Wurzel der mittleren Fehlerquadratsumme (RMSE), mittlerer absoluter Fehler (MAE), mittlerer absoluter Fehler in Prozent (MAPE), maximaler absoluter Fehler (MaxAE), maximaler absoluter Fehler in Prozent (MaxAPE), normalisiertes bayessches Informationskriterium (BIC), Boxplots der Residuen-Autokorrelationen und partiellen Autokorrelationen. Ergebnisse für einzelne Modelle: Vorhersagewerte, Anpassungswerte, beobachtete Werte, untere und obere Konfidenzgrenzen, Residuen-Autokorrelationen und partielle Autokorrelationen.

Erläuterungen der Daten für die Prozedur "Zeitreihenmodellierung"

**Daten.** Die abhängige Variable und etwaige unabhängige Variablen müssen numerisch sein.

**Annahmen.** Die abhängige Variable und alle unabhängigen Variablen werden als Zeitreihen behandelt, d. h. jeder Fall repräsentiert einen Zeitpunkt, und die nachfolgenden Fälle liegen jeweils ein konstantes Zeitintervall auseinander.

- **Stationarität.** Bei benutzerdefinierten ARIMA-Modellen muss die zu modellierende Zeitreihe stationär sein. Die effektivste Methode zur Transformation einer nicht stationären Zeitreihe in eine stationäre besteht in einer Differenztransformation, die über das Dialogfeld Zeitreihe erstellen aufgerufen werden kann.
- **Vorhersagen.** Wenn Vorhersagen anhand von Modellen mit unabhängigen Variablen (Prädiktorvariablen) erstellt werden sollen, muss das aktive Dataset für alle Fälle innerhalb der Vorhersageperiode Werte dieser Variablen enthalten. Darüber hinaus dürfen die unabhängigen Variablen in der Schätzperiode keine fehlenden Werte enthalten.

## Datum definieren

Es ist zwar nicht zwingend erforderlich, aber es empfiehlt sich, im Dialogfeld "Datum definieren" das dem ersten Fall zugewiesene Datum und die Intervalle zwischen den nachfolgenden Fällen anzugeben. Diese Angabe erfolgt vor der Verwendung der Zeitreihenmodellierung und ergibt ein Set von Variablen, die das den einzelnen Fällen zugewiesene Datum beschriften. Dabei wird außerdem eine angenommene Periodizität der Daten festgelegt, beispielsweise eine Periodizität von 12, wenn das Zeitintervall zwischen aufeinander folgenden Fällen einen Monat beträgt. Diese Periodizität ist erforderlich, wenn Sie saisonale Modelle erstellen möchten. Wenn Sie keine saisonalen Modelle erstellen möchten und keine Datumsbeschriftungen in der Ausgabe benötigen, müssen Sie das Dialogfeld "Datum definieren" nicht aufrufen. Die den einzelnen Fällen zugewiesene Beschriftung enthält dann einfach die Fallnummer.

So verwenden Sie die Zeitreihenmodellierung:

1. Wählen Sie in den Menüs Folgendes aus:  
**Analysieren > Vorhersage > Traditionelle Modelle erstellen...**
2. Wählen Sie auf der Registerkarte "Variablen" mindestens eine abhängige Variable für die Modellierung aus.
3. Wählen Sie in der Dropdown-Liste "Methode" eine Modellierungsmethode aus. Übernehmen Sie für automatische Modellierung die Standardmethode des **Expert Modeler**. Der Expert Modeler ermittelt dann für jede abhängige Variable das am besten angepasste Modell.  
So erstellen Sie Vorhersagen:
4. Klicken Sie auf die Registerkarte **Optionen**.
5. Geben Sie die Vorhersageperiode an. Es wird ein Diagramm mit Vorhersagen und beobachteten Werten erstellt.

Die folgenden Optionen sind verfügbar:

- Wählen Sie mindestens eine unabhängige Variable aus. Unabhängige Variablen werden weitgehend wie Prädiktoren in der Regressionsanalyse behandelt, sind jedoch optional. Sie können in ARIMA-Modelle, nicht jedoch in Modelle mit exponentiellem Glätten eingeschlossen werden. Wenn Sie **Expert Modeler** als Modellierungsmethode angeben und unabhängige Variablen einschließen, werden nur ARIMA-Modelle berücksichtigt.
- Klicken Sie zum Angeben der Modellierungsdetails auf **Kriterien**.
- Speichern Sie Vorhersagen, Konfidenzintervalle und Restrauschen.
- Speichern Sie die geschätzten Modelle im XML-Format. Gespeicherte Modelle können neuen oder korrigierten Daten zugewiesen werden, um aktualisierte Vorhersagen ohne Neuerstellen der Modelle zu erhalten.
- Erstellen Sie Auswertungsstatistiken für alle geschätzten Modelle.
- Geben Sie Transferfunktionen für unabhängige Variablen in ARIMA-Modellen an.
- Aktivieren Sie die automatische Erkennung von Ausreißern.
- Definieren Sie spezifische Zeitpunkte zum Erkennen von Ausreißern für benutzerdefinierte ARIMA-Modelle.

## Modellierungsmethode

Die folgenden Modellierungsmethoden sind verfügbar:

**Expert Modeler.** Der Expert Modeler ermittelt automatisch das jeweils am besten angepasste Modell für die einzelnen abhängigen Zeitreihen. Wenn unabhängige Variablen (Prädiktoren) angegeben sind, wählt der Expert Modeler für den Einschluss in ARIMA-Modelle diejenigen aus, die eine statistisch signifikante Beziehung mit der abhängigen Zeitreihe aufweisen. Modellvariablen werden gegebenenfalls durch Differenzierung und/oder Quadratwurzeltransformation bzw. Transformation mit natürlichem Logarithmus transformiert. Der Expert Modeler berücksichtigt in der Standardeinstellung sowohl Modelle mit ex-

ponentiellem Glätten als auch ARIMA-Modelle. Sie können den Expert Modeler jedoch auch auf die abschließliche Suche nach ARIMA-Modellen bzw. nach Modellen mit exponentiellem Glätten einschränken. Sie können auch die automatische Erkennung von Ausreißern festlegen.

**Exponentielles Glätten.** Mit dieser Option können Sie ein benutzerdefiniertes Modell mit exponentiellem Glätten angeben. Dabei können Sie zwischen verschiedenen Modellen mit exponentiellem Glätten wählen, die sich hinsichtlich der Behandlung von Trends und Saisonalität unterscheiden.

**ARIMA.** Mit dieser Option können Sie ein ARIMA-Modell angeben. Hierfür müssen autoregressive Ordnungen, Ordnungen für gleitenden Durchschnitt und der Grad der Differenzierung angegeben werden. Sie können unabhängige Variablen (Prädiktoren) einschließen und für einige oder alle Transferfunktionen definieren. Außerdem können Sie die automatische Erkennung von Ausreißern oder eine bestimmte Gruppe von Ausreißern angeben.

Schätz- und Vorhersageperioden

**Schätzperiode.** Die Schätzperiode gibt das Set von Fällen an, anhand dessen das Modell bestimmt wird. In der Standardeinstellung umfasst die Schätzperiode sämtliche Fälle im aktiven Dataset. Sie können die Schätzperiode festlegen, indem Sie im Dialogfeld Fälle auswählen die Option **Nach Zeit- oder Fallbereich** auswählen. Die von der Prozedur verwendete Schätzperiode kann je nach den verfügbaren Daten unterschiedliche abhängige Variablen verwenden und daher vom angezeigten Wert abweichen. Die tatsächliche Schätzperiode für eine bestimmte abhängige Variable entspricht der Zeitspanne, die verbleibt, nachdem die direkt aufeinander folgenden fehlenden Werte von der Variablen entfernt wurden (am Anfang oder am Ende der angegebenen Schätzperiode).

**Vorhersageperiode.** Die Vorhersageperiode beginnt mit dem ersten Fall nach der Schätzperiode und reicht in der Standardeinstellung bis zum letzten Fall im aktiven Dataset. Das Ende der Vorhersageperiode kann auf der Registerkarte Optionen festgelegt werden.

---

## Angeben von Optionen für den Expert Modeler

Der Expert Modeler bietet Optionen, mit denen das Set möglicher Modelle eingeschränkt, die Behandlung von Ausreißern angegeben und Ereignisvariablen angegeben werden können.

### Modellauswahl und Ereignisspezifikation

Auf der Registerkarte "Modell" können Sie die Modelltypen, die vom Expert Modeler berücksichtigt werden sollen, und Ereignisvariablen angeben.

**Modelltyp.** Die folgenden Optionen sind verfügbar:

- **Alle Modelle.** Der Expert Modeler berücksichtigt sowohl ARIMA-Modelle als auch Modelle mit exponentiellem Glätten.
- **Nur Modelle mit exponentiellem Glätten.** Der Expert Modeler berücksichtigt nur Modelle mit exponentiellem Glätten.
- **Nur ARIMA-Modelle.** Der Expert Modeler berücksichtigt nur ARIMA-Modelle.

**Expert Modeler zieht saisonale Modelle in Betracht.** Diese Option ist nur verfügbar, wenn für das aktive Dataset eine Periodizität definiert wurde. Wenn diese Option aktiviert wird, berücksichtigt der Expert Modeler sowohl saisonale als auch nicht saisonale Modelle. Wenn diese Option inaktiviert ist, berücksichtigt der Expert Modeler nur nicht saisonale Modelle.

**Aktuelle Periodizität.** Gibt die gegenwärtig für das aktive Dataset definierte Periodizität an (sofern vorhanden). Die aktuelle Periodizität wird als Ganzzahl angegeben, z. B. 12 für eine jährliche Periodizität, wobei jeder Fall einen Monat darstellt. Wenn keine Periodizität angegeben wurde, wird der Wert *Keine* angezeigt. Für saisonales Modelle muss eine Periodizität angegeben werden. Die Periodizität wird im Dialogfeld "Datum definieren" angegeben.

**Ereignisse.** Wählen Sie unabhängige Variablen aus, die als Ereignisvariablen behandelt werden sollen. Bei Ereignisvariablen geben Fälle mit einem Wert von 1 die Zeitpunkte an, zu denen erwartet wird, dass die abhängigen Zeitreihen von dem Ereignis beeinflusst werden. Andere Werte als 1 geben an, dass sie nicht beeinflusst werden.

## Behandeln von Ausreißern mit Expert Modeler

Auf der Registerkarte "Ausreißer" können Sie die automatische Erkennung von Ausreißern und die Typen von Ausreißern angeben, die erkannt werden sollen.

**Ausreißer automatisch erkennen.** In der Standardeinstellung wird keine automatische Erkennung von Ausreißern durchgeführt. Aktivieren Sie diese Option, um eine automatische Erkennung von Ausreißern durchzuführen, und wählen Sie anschließend mindestens einen der folgenden Ausreißertypen aus:

- Additiv
- Niveauverschiebung
- Neuerung
- Kurzlebig
- Saisonal additiv
- Lokaler Trend
- Additive Gruppe

---

## Exponentielles Glätten: Benutzerdefinierte Modelle

**Modelltyp.** Modelle mit exponentiellem Glätten<sup>1</sup> werden als saisonal oder nicht saisonal klassifiziert. Saisonale Modelle sind nur verfügbar, wenn für das aktive Dataset eine Periodizität definiert wurde (siehe im Folgenden unter "Aktuelle Periodizität").

- *Einfach.* Dieses Modell eignet sich für Zeitreihen, bei denen weder ein Trend noch eine Saisonalität vorliegt. Sein einziger Glättungsparameter betrifft das Niveau. Einfaches exponentielles Glätten weist eine sehr große Ähnlichkeit auf mit einem ARIMA-Modell mit Autoregression der Ordnung null, Differenzbildung der Ordnung 1, gleitenden Durchschnitten der Ordnung 1 und fehlender Konstante.
- *Linearer Trend nach Holt.* Dieses Modell eignet sich für Zeitreihen, die einen linearen Trend, aber keine Saisonalität aufweisen. Seine Glättungsparameter betreffen Niveau und Trend und es wird angenommen, dass die Werte dieser beiden Elemente unabhängig voneinander sind. Das Holt-Modell ist allgemeiner als das Brown-Modell, aber es kann bei langen Zeitreihen mehr Rechenzeit erfordern. Exponentielles Glätten mit Holt-Modell weist eine sehr große Ähnlichkeit auf mit einem ARIMA-Modell mit Autoregression der Ordnung 0, Differenzbildung der Ordnung 2 und gleitenden Durchschnitten der Ordnung 2.
- *Linearer Trend nach Brown.* Dieses Modell eignet sich für Zeitreihen, die einen linearen Trend, aber keine Saisonalität aufweisen. Seine Glättungsparameter betreffen Niveau und Trend und es wird angenommen, dass diese Faktoren gleich sind. Das Brown-Modell ist daher ein Spezialfall des Holt-Modells. Exponentielles Glätten mit dem Brown-Modell weist sehr große Ähnlichkeit auf mit einem ARIMA-Modell mit Autoregression der Ordnung 0, Differenzbildung der Ordnung 2 und gleitenden Durchschnitten der Ordnung 2; dabei ist bei den gleitenden Durchschnitten der Koeffizient für den Abstand 2 gleich dem Quadrat der Hälfte des Koeffizienten für den Abstand 1.
- *Gedämpfter Trend.* Dieses Modell eignet sich für Zeitreihen mit einem sich abschwächenden linearen Trend, aber ohne Saisonalität. Seine Glättungsparameter betreffen Niveau, Trend und die Dämpfung des Trends. Exponentielles Glätten mit gedämpftem Trend weist eine sehr große Ähnlichkeit auf mit einem ARIMA-Modell mit Autoregression der Ordnung 1, Differenzbildung der Ordnung 1 und gleitenden Durchschnitten der Ordnung 2.
- *Einfach saisonal.* Dieses Modell eignet sich für Zeitreihen ohne Trend, aber mit einem saisonalen Effekt, der über die Zeit gleich bleibt. Seine Glättungsparameter betreffen Niveau und Saison. Einfaches saiso-

---

1. Gardner, E. S. 1985. Exponential smoothing: The state of the art. *Journal of Forecasting*, 4, 1-28.

nales exponentielles Glätten weist sehr große Ähnlichkeit auf mit einem ARIMA-Modell mit Autoregression der Ordnung 0, Differenzenbildung der Ordnung 1, saisonaler Differenzenbildung der Ordnung 1 und gleitenden Durchschnitten der Ordnung 1,  $p$  und  $p+1$ , wobei  $p$  die Anzahl der Perioden im Saisonintervall ist (bei monatlichen Daten ist  $p=12$ ).

- *Winters additiv*. Dieses Modell eignet sich für Zeitreihen mit einem linearen Trend und einem saisonalen Effekt, der nicht vom Niveau der Zeitreihe abhängt. Seine Glättungsparameter betreffen Niveau, Trend und Saison. Exponentielles Glätten mit dem additiven Winters-Modell weist sehr große Ähnlichkeit auf mit einem ARIMA-Modell mit Autoregression der Ordnung 0, Differenzenbildung der Ordnung 1, saisonaler Differenzenbildung der Ordnung 1 und gleitenden Durchschnitten der Ordnung  $p+1$ , wobei  $p$  die Anzahl der Zeitpunkte in einem Saisonintervall ist (für monatliche Daten ist  $p=12$ ).
- *Winters multiplikativ*. Dieses Modell eignet sich für Zeitreihen mit einem linearen Trend und einem saisonalen Effekt, der vom Niveau der Zeitreihe abhängt. Seine Glättungsparameter betreffen Niveau, Trend und Saison. Exponentielles Glätten mit dem multiplikativen Winters-Modell weist keine Ähnlichkeit zu irgendeinem ARIMA-Modell auf.

**Aktuelle Periodizität.** Gibt die gegenwärtig für das aktive Dataset definierte Periodizität an (sofern vorhanden). Die aktuelle Periodizität wird als Ganzzahl angegeben, z. B. 12 für eine jährliche Periodizität, wobei jeder Fall einen Monat darstellt. Wenn keine Periodizität angegeben wurde, wird der Wert *Keine* angezeigt. Für saisonale Modelle muss eine Periodizität angegeben werden. Die Periodizität wird im Dialogfeld "Datum definieren" angegeben.

**Transformation abhängiger Variablen.** Sie können angeben, dass für alle abhängigen Variablen vor der Modellierung eine Transformation durchgeführt wird.

- **Keine.** Es wird keine Transformation durchgeführt.
- **Quadratwurzel.** Quadratwurzeltransformation.
- **Natürlicher Logarithmus.** Transformation mit natürlichem Logarithmus.

---

## Benutzerdefinierte ARIMA-Modelle

Mit der Zeitreihenmodellierung können benutzerdefinierte nicht saisonale oder saisonale univariate ARIMA-Modelle (auch als Box-Jenkins-Modelle<sup>2</sup> bekannt) mit oder ohne feste Gruppe von Prädiktorvariablen erstellt werden. (ARIMA steht für Autoregressive Integrated Moving Average (autoregressiver integrierter gleitender Durchschnitt).) Sie können für einige oder alle Prädiktoren Transferfunktionen definieren, die automatische Erkennung von Ausreißern festlegen oder eine bestimmte Gruppe von Ausreißern angeben.

- Alle auf der Registerkarte "Variablen" angegebenen unabhängigen Variablen (Prädiktoren) werden explizit in das Modell eingeschlossen. Im Gegensatz dazu werden im Expert Modeler unabhängige Variablen nur eingeschlossen, wenn sie eine statistisch signifikante Beziehung mit der abhängigen Variablen aufweisen.

## Modellspezifikation für benutzerdefinierte ARIMA-Modelle

Auf der Registerkarte "Modelle" können Sie die Struktur eines benutzerdefinierten ARIMA-Modells festlegen.

**ARIMA-Ordnungen.** Geben Sie Werte für die verschiedenen ARIMA-Komponenten des Modells in die entsprechenden Zellen des Strukturrasters ein. Alle Werte müssen nicht negative Ganzzahlen sein. Bei autoregressiven Komponenten und Komponenten des gleitenden Durchschnitts stellt der Wert die höchste Ordnung dar. Alle positiven niedrigeren Ordnungen werden in das Modell eingeschlossen. Wenn Sie beispielsweise 2 angeben, enthält das Modell die Ordnungen 2 und 1. Die Zellen in der Spalte "Saisonal" sind nur verfügbar, wenn für das aktive Dataset eine Periodizität definiert wurde (siehe im Folgenden unter "Aktuelle Periodizität").

---

2. Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. 1994. *Time series analysis: Forecasting and control*, 3rd ed. Englewood Cliffs, N.J.: Prentice Hall.

- **Autoregressiv (p).** Die Anzahl autoregressiver Ordnungen im Modell. Autoregressive Ordnungen geben die zurückliegenden Werte der Zeitreihe an, die für die Vorhersage der aktuellen Werte verwendet werden. Eine autoregressive Ordnung von 2 gibt beispielsweise an, dass die Werte der Zeitreihe, die zwei Zeitperioden zurückliegt, für die Vorhersage der aktuellen Werte verwendet wird.
- **Differenz (d).** Gibt die Ordnung der Differenzierung an, die vor dem Schätzen der Modelle auf die Zeitreihe angewendet wurde. Differenzierung ist erforderlich, wenn Trends vorhanden sind. (Zeitreihen mit Trends sind normalerweise nicht stationär, und bei der ARIMA-Modellierung wird Stationarität angenommen.) Mithilfe der Differenzierung werden die Effekte der Trends entfernt. Die Ordnung der Differenzierung entspricht dem Grad des Trends der Zeitreihe: Differenzierung erster Ordnung erklärt lineare Trends, Differenzierung zweiter Ordnung erklärt quadratische Trends usw.
- **Gleitender Durchschnitt (q).** Die Anzahl von Ordnungen des gleitenden Durchschnitts im Modell. Ordnungen des gleitenden Durchschnitts geben an, wie Abweichungen vom Mittelwert der Zeitreihe für zurückliegende Werte zum Vorhersagen der aktuellen Werte verwendet werden. Moving-Average-Ordnungen von 1 und 2 geben beispielsweise an, dass beim Vorhersagen der aktuellen Werte der Zeitreihe Abweichungen vom Mittelwert der Zeitreihe von den beiden letzten Zeitperioden berücksichtigt werden sollen.

**Saisonale Ordnungen.** Saisonale autoregressive Komponenten, Komponenten des gleitenden Durchschnitts und Differenzierungskomponenten entsprechen im Prinzip ihren nicht saisonalen Gegenstücken. Bei saisonalen Ordnungen werden die Werte der aktuellen Zeitreihe jedoch von Werten zurückliegender Zeitreihen beeinflusst, die um eine oder mehrere saisonalen Perioden getrennt sind. Bei monatlichen Daten (saisonale Periode von 12) beispielsweise bedeutet eine saisonale Ordnung von 1, dass der Wert der aktuellen Zeitreihe durch den Zeitreihenwert beeinflusst wird, der 12 Perioden vor dem aktuellen liegt. Eine saisonale Ordnung von 1 entspricht bei monatlichen Daten einer nicht saisonalen Ordnung von 12.

**Aktuelle Periodizität.** Gibt die gegenwärtig für das aktive Dataset definierte Periodizität an (sofern vorhanden). Die aktuelle Periodizität wird als Ganzzahl angegeben, z. B. 12 für eine jährliche Periodizität, wobei jeder Fall einen Monat darstellt. Wenn keine Periodizität angegeben wurde, wird der Wert *Keine* angezeigt. Für saisonales Modelle muss eine Periodizität angegeben werden. Die Periodizität wird im Dialogfeld "Datum definieren" angegeben.

**Transformation abhängiger Variablen.** Sie können angeben, dass für alle abhängigen Variablen vor der Modellierung eine Transformation durchgeführt wird.

- **Keine.** Es wird keine Transformation durchgeführt.
- **Quadratwurzel.** Quadratwurzeltransformation.
- **Natürlicher Logarithmus.** Transformation mit natürlichem Logarithmus.

**Konstante in Modell einschließen.** Der Einschluss einer Konstanten ist das Standardverfahren, sofern Sie nicht sicher wissen, dass der Gesamtmittelwert der Zeitreihe 0 ist. Bei der Anwendung von Differenzierung empfiehlt es sich, die Konstante auszuschließen.

## Transferfunktionen in benutzerdefinierten ARIMA-Modellen

Auf der Registerkarte "Transferfunktion" (nur vorhanden, wenn unabhängige Variablen angegeben wurden) können Sie Transferfunktionen für bestimmte oder alle unabhängigen Variablen definieren, die auf der Registerkarte "Variablen" angegeben sind. Durch Transferfunktionen kann angegeben werden, wie zurückliegende Werte unabhängiger Variablen (Prädiktoren) zum Vorhersagen zukünftiger Werte der abhängigen Zeitreihe verwendet werden sollen.

**Transferfunktionsordnungen.** Geben Sie Werte für die verschiedenen Komponenten der Transferfunktion in die entsprechenden Zellen des Strukturrasters ein. Alle Werte müssen nicht negative Ganzzahlen sein. Bei Zähler- und Nennerkomponenten stellt der Wert die höchste Ordnung dar. Alle positiven niedrigeren Ordnungen werden in das Modell eingeschlossen. Darüber hinaus wird die Ordnung 0 bei Zählerkomponenten immer eingeschlossen. Wenn Sie beispielsweise 2 als Zähler angeben, enthält das Modell die Ordnungen 2, 1 und 0. Wenn Sie 3 als Nenner angeben, enthält das Modell die Ordnungen 3, 2 und 1. Die

Zellen in der Spalte "Saisonal" sind nur verfügbar, wenn für das aktive Dataset eine Periodizität definiert wurde (siehe im Folgenden unter "Aktuelle Periodizität").

- **Zähler.** Der Zählerterm der Transferfunktion. Dieser gibt an, welche zurückliegenden Werte aus der ausgewählten unabhängigen Zeitreihe (Prädiktoren) zum Vorhersagen der aktuellen Werte der abhängigen Zeitreihe verwendet werden. Ein Zählerterm von 1 gibt beispielsweise an, dass der Wert einer unabhängigen Zeitreihe, die eine Periode zurückliegt, und der aktuelle Wert der unabhängigen Zeitreihe zum Vorhersagen des aktuellen Werts der einzelnen abhängigen Zeitreihen verwendet werden.
- **Nenner.** Der Nennerterm der Transferfunktion. Dieser gibt an, wie Abweichungen vom Mittelwert der Zeitreihe für zurückliegende Werte der ausgewählten unabhängigen Zeitreihe (Prädiktoren) zum Vorhersagen der aktuellen Werte der abhängigen Zeitreihe verwendet werden. Ein Nennerterm von 1 gibt beispielsweise an, dass beim Vorhersagen der aktuellen Werte für die einzelnen abhängigen Zeitreihen Abweichungen vom Mittelwert einer unabhängigen Zeitreihe berücksichtigt werden sollen, die eine Zeitperiode zurückliegt.
- **Differenz.** Gibt die Ordnung der Differenzierung an, die vor dem Schätzen der Modelle auf die ausgewählte unabhängige Zeitreihe (Prädiktoren) angewendet wurde. Wenn Trends vorhanden sind, ist die Differenzierung erforderlich, um die Effekte der Trends zu entfernen.

**Saisonale Ordnungen.** Saisonale Zähler-, Nenner- und Differenzierungskomponenten entsprechen im Prinzip ihren nicht saisonalen Gegenstücken. Bei saisonalen Ordnungen werden die Werte der aktuellen Zeitreihe jedoch von Werten zurückliegender Zeitreihen beeinflusst, die um eine oder mehrere saisonalen Perioden getrennt sind. Bei monatlichen Daten (saisonale Periode von 12) beispielsweise bedeutet eine saisonale Ordnung von 1, dass der Wert der aktuellen Zeitreihe durch den Zeitreihenwert beeinflusst wird, der 12 Perioden vor dem aktuellen liegt. Eine saisonale Ordnung von 1 entspricht bei monatlichen Daten einer nicht saisonalen Ordnung von 12.

**Aktuelle Periodizität.** Gibt die gegenwärtig für das aktive Dataset definierte Periodizität an (sofern vorhanden). Die aktuelle Periodizität wird als Ganzzahl angegeben, z. B. 12 für eine jährliche Periodizität, wobei jeder Fall einen Monat darstellt. Wenn keine Periodizität angegeben wurde, wird der Wert *Keine* angezeigt. Für saisonales Modelle muss eine Periodizität angegeben werden. Die Periodizität wird im Dialogfeld "Datum definieren" angegeben.

**Verzögerung.** Wenn eine Verzögerung festgelegt wird, verzögert sich der Einfluss der unabhängigen Variable um die Anzahl der angegebenen Intervalle. Bei einer Verzögerung mit dem Wert 5 beeinflusst der Wert der unabhängigen Variable zum Zeitpunkt  $t$  die Vorhersagen erst nach dem Ablauf von fünf Perioden ( $t + 5$ ).

**Transformation.** Die Angabe einer Transferfunktion für ein Set von unabhängigen Variablen enthält auch eine optionale Transformation, die für diese Variablen ausgeführt werden soll.

- **Keine.** Es wird keine Transformation durchgeführt.
- **Quadratwurzel.** Quadratwurzeltransformation.
- **Natürlicher Logarithmus.** Transformation mit natürlichem Logarithmus.

## Ausreißer in benutzerdefinierten ARIMA-Modellen

Auf der Registerkarte "Ausreißer" sind die folgenden Möglichkeiten für die Behandlung von Ausreißern verfügbar<sup>3</sup>: Ausreißer automatisch erkennen, bestimmte Punkte als Ausreißer angeben oder Ausreißer nicht erkennen bzw. modellieren.

**Keine Ausreißer erkennen oder modellieren.** In der Standardeinstellung werden Ausreißer weder erkannt noch modelliert. Aktivieren Sie diese Option, um die Erkennung und Modellierung von Ausreißern zu inaktivieren.

---

3. Pena, D., G. C. Tiao, and R. S. Tsay, eds. 2001. *A course in time series analysis*. New York: John Wiley and Sons.

**Ausreißer automatisch erkennen.** Aktivieren Sie diese Option, um eine automatische Erkennung von Ausreißern durchzuführen, und wählen Sie mindestens einen der folgenden Ausreißertypen aus:

- Additiv
- Niveauverschiebung
- Neuerung
- Kurzlebig
- Saisonal additiv
- Lokaler Trend
- Additive Gruppe

Weitere Informationen finden Sie im Thema Kapitel 9, „Ausreißertypen“, auf Seite 61.

**Bestimmte Zeitpunkte als Ausreißer modellieren.** Aktivieren Sie diese Option, um bestimmte Zeitpunkte als Ausreißer zu modellieren. Verwenden Sie für jeden Ausreißer eine eigene Zeile im Ausreißerdefinitionsraster. Geben Sie in alle Zellen einer bestimmten Zeile Werte ein.

- **Typ.** Der Ausreißertyp. Unterstützte Typen: additiv (Standardeinstellung), Verschiebung im Niveau, innovativ, transient, saisonal additiv und lokaler Trend.

*Hinweis 1:* Wenn für das aktive Dataset kein Datum angegeben wurde, wird im Ausreißerdefinitionsraster nur die Spalte *Beobachtung* angezeigt. Um einen Ausreißer festzulegen, geben Sie die Zeilennummer des entsprechenden Falles ein (wie im Dateneditor angezeigt).

*Hinweis 2:* Die Spalte *Zyklus* im Ausreißerdefinitionsraster (sofern vorhanden) bezieht sich auf die Werte der Variablen *CYCLE\_* im aktiven Dataset.

---

## Ausgabe

Als Ausgabe stehen sowohl Ergebnisse für einzelne Modelle als auch für alle Modelle errechnete Ergebnisse zur Verfügung. Die Ergebnisse für einzelne Modelle können über vom Benutzer angegebene Kriterien auf eine Gruppe von am besten/schlechtesten angepassten Modellen eingeschränkt werden.

## Statistiken und Vorhersagetabellen

Die Registerkarte "Statistik" bietet Optionen für die Anzeige von Tabellen mit den Modellierungsergebnissen.

**Anpassungsmaße, Ljung-Box-Statistik und Anzahl von Ausreißern nach Modell anzeigen.** Aktivieren Sie diese Option, um für jedes geschätzte Modell eine Tabelle mit den ausgewählten Anpassungsmaßen, Ljung-Box-Werten und der Anzahl von Ausreißern anzuzeigen.

**Anpassungsmaße.** Sie können eine oder mehrere der folgenden Optionen für die Aufnahme in die Tabelle auswählen, die die Anpassungsmaße für die einzelnen geschätzten Modelle enthält:

- *R*-Quadrat für stationären Teil
- *R*-Quadrat
- Wurzel der mittleren Fehlerquadratsumme
- Mittlerer absoluter Fehler in Prozent
- Mittlerer absoluter Fehler
- Maximaler absoluter Fehler in Prozent
- Maximaler absoluter Fehler
- Normalisiertes BIC

**Statistik für Modellvergleich.** Mit dieser Gruppe von Optionen wird die Anzeige der Tabellen gesteuert, die die für alle Modelle berechneten Statistiken enthalten. Für jede Option wird eine eigene Tabelle erstellt. Sie können eine oder mehrere der folgenden Optionen auswählen:

- **Güte der Anpassung.** Tabelle der Auswertungsstatistiken und Perzentile für  $R$ -Quadrat für stationären Teil,  $R$ -Quadrat, Wurzel der mittleren Fehlerquadratsumme, mittlerer absoluter Fehler in Prozent, mittlerer absoluter Fehler, maximaler absoluter Fehler in Prozent, maximaler absoluter Fehler und normalisiertes Bayes-Informationskriterium.
- **Residuen-Autokorrelationsfunktion (ACF).** Tabelle der Auswertungsstatistiken und Perzentile für Autokorrelationen der Residuen für alle geschätzten Modelle.
- **Part. Residuen-Autokorrelationsfunktion (PACF).** Tabelle der Auswertungsstatistiken und Perzentile für partielle Autokorrelationen der Residuen für alle geschätzten Modelle.

**Statistik nach einzelnen Modellen.** Mit dieser Gruppe von Optionen wird die Anzeige der Tabellen gesteuert, die detaillierte Informationen zu den einzelnen geschätzten Modellen enthalten. Für jede Option wird eine eigene Tabelle erstellt. Sie können eine oder mehrere der folgenden Optionen auswählen:

- **Parameterschätzungen.** Zeigt für jedes geschätzte Modell eine Tabelle der Parameterschätzungen an. Für Modelle mit exponentiellem Glätten und ARIMA-Modelle werden eigene Tabellen angezeigt. Wenn Ausreißer vorhanden sind, werden Parameterschätzungen für diese ebenfalls in einer eigenen Tabelle angezeigt.
- **Residuen-Autokorrelationsfunktion (ACF).** Zeigt für jedes geschätzte Modell eine Tabelle der Residuen-Autokorrelationen nach Intervall an. Die Tabelle enthält die Konfidenzintervalle für die Autokorrelationen.
- **Part. Residuen-Autokorrelationsfunktion (PACF).** Zeigt für jedes geschätzte Modell eine Tabelle der partiellen Residuen-Autokorrelationen nach Intervall an. Die Tabelle enthält die Konfidenzintervalle für die partiellen Autokorrelationen.

**Vorhersagen anzeigen.** Zeigt eine Tabelle der Modellvorhersagen und der Konfidenzintervalle für jedes geschätzte Modell an. Die Vorhersageperiode wird auf der Registerkarte "Optionen" festgelegt.

## Diagramme

Die Registerkarte "Diagramme" bietet Optionen für die Anzeige von Diagrammen der Modellierungsergebnisse.

Diagramme zum Vergleichen von Modellen

Mit dieser Gruppe von Optionen wird die Anzeige der Diagramme gesteuert, die die für alle Modelle berechneten Statistiken enthalten. Für jede Option wird ein eigenes Diagramm erstellt. Sie können eine oder mehrere der folgenden Optionen auswählen:

- $R$ -Quadrat für stationären Teil
- $R$ -Quadrat
- Wurzel der mittleren Fehlerquadratsumme
- Mittlerer absoluter Fehler in Prozent
- Mittlerer absoluter Fehler
- Maximaler absoluter Fehler in Prozent
- Maximaler absoluter Fehler
- Normalisiertes BIC
- Residuen-Autokorrelationsfunktion (ACF)
- Part. Residuen-Autokorrelationsfunktion (PACF)

Diagramme nach einzelnen Modellen

**Reihen.** Aktivieren Sie diese Option, um für die einzelnen geschätzten Modelle Diagramme der vorhergesagten Werte zu erhalten. Sie können eine oder mehrere der folgenden Optionen für die Aufnahme in das Diagramm auswählen:

- **Beobachtete Werte.** Die beobachteten Werte der abhängigen Zeitreihe.
- **Vorhersagen.** Die vom Modell vorhergesagten Werte für die Vorhersageperiode.
- **Anpassungswerte.** Die vom Modell vorhergesagten Werte für die Schätzperiode.
- **Konfidenzintervalle für Vorhersagen.** Die Konfidenzintervalle für die Vorhersageperiode.
- **Konfidenzintervalle für Anpassungswerte.** Die Konfidenzintervalle für die Schätzperiode.

**Residuen-Autokorrelationsfunktion (ACF).** Zeigt für jedes geschätzte Modell ein Diagramm der Residuen-Autokorrelationen an.

**Part. Residuen-Autokorrelationsfunktion (PACF).** Zeigt für jedes geschätzte Modell ein Diagramm der partiellen Residuen-Autokorrelationen an.

## Einschränken der Ausgabe auf die am besten/schlechtesten angepassten Modelle

Die Registerkarte "Ausgabefilter" enthält Optionen, mit denen sowohl Tabellen- als auch Diagrammausgabe auf ein Subset von geschätzten Modellen beschränkt werden kann. Sie können die Ausgabe mithilfe von Anpassungskriterien auf die am besten und/oder am schlechtesten angepassten Modelle beschränken. In der Standardeinstellung sind alle geschätzten Modelle in der Ausgabe enthalten.

**Am besten angepasste Modelle.** Aktivieren Sie diese Option, um die am besten angepassten Modelle in die Ausgabe einzuschließen. Wählen Sie ein Maß für Anpassungsgüte und geben Sie die Anzahl von Modellen an, die eingeschlossen werden sollen. Die Auswahl dieser Option hindert Sie nicht daran, auch die am schlechtesten angepassten Modelle auszuwählen. Die Ausgabe würde in diesem Fall sowohl die am schlechtesten als auch die am besten angepassten Modelle enthalten.

- **Feste Anzahl von Modellen.** Gibt an, dass Ergebnisse für die  $n$  am besten angepassten Modelle angezeigt werden. Wenn die Anzahl die Anzahl der geschätzten Modelle überschreitet, werden alle Modelle angezeigt.
- **Prozentsatz der Gesamtzahl von Modellen.** Gibt an, dass Ergebnisse für Modelle angezeigt werden, deren Werte für die Anpassungsgüte für alle geschätzten Modelle in den obersten  $n$  Prozent liegen.

**Am schlechtesten angepasste Modelle.** Aktivieren Sie diese Option, um die am schlechtesten angepassten Modelle in die Ausgabe einzuschließen. Wählen Sie ein Maß für Anpassungsgüte und geben Sie die Anzahl von Modellen an, die eingeschlossen werden sollen. Die Auswahl dieser Option hindert Sie nicht daran, auch die am besten angepassten Modelle auszuwählen. Die Ausgabe würde in diesem Fall sowohl die am besten als auch die am schlechtesten angepassten Modelle enthalten.

- **Feste Anzahl von Modellen.** Gibt an, dass Ergebnisse für die  $n$  am schlechtesten angepassten Modelle angezeigt werden. Wenn die Anzahl die Anzahl der geschätzten Modelle überschreitet, werden alle Modelle angezeigt.
- **Prozentsatz der Gesamtzahl von Modellen.** Gibt an, dass Ergebnisse für Modelle angezeigt werden, deren Werte für die Anpassungsgüte für alle geschätzten Modelle in den untersten  $n$  Prozent liegen.

**Maß für Anpassungsgüte.** Wählen Sie das Maß für die Anpassungsgüte, anhand dessen die Modelle gefiltert werden sollen. Der Standardwert ist  $R$ -Quadrat für den stationären Teil.

---

## Speichern von Modellvorhersagen und Modellspezifikationen

Auf der Registerkarte "Speichern" können Sie festlegen, dass Modellvorhersagen als neue Variablen im aktiven Dataset gespeichert werden sollen. Außerdem können die Modellspezifikationen in einer externen Datei im XML-Format gespeichert werden.

**Variablen speichern.** Sie können Modellvorhersagen, Konfidenzintervalle und Residuen als neue Variablen im aktiven Dataset speichern. Jede abhängige Zeitreihe erzeugt eine eigene Gruppe neuer Variablen, und jede neue Variable enthält Werte für Schätz- und Vorhersageperioden. Wenn die Vorhersageperiode die Länge der abhängigen Zeitreihe überschreitet, werden neue Fälle hinzugefügt. Wenn die neuen Variablen gespeichert werden sollen, aktivieren Sie jeweils das zugehörige Kontrollkästchen "Speichern". In der Standardeinstellung werden keine neuen Variablen gespeichert.

- **Vorhergesagte Werte.** Die vom Modell vorhergesagten Werte.
- **Untere Konfidenzgrenzen.** Die unteren Konfidenzgrenzen für die vorhergesagten Werte.
- **Obere Konfidenzgrenzen.** Die oberen Konfidenzgrenzen für die vorhergesagten Werte.
- **Restrauschen.** Die Modellresiduen. Bei der Transformation von abhängigen Variablen (z. B. mit natürlichem Logarithmus) sind dies die Residuen für die transformierte Reihe.
- **Präfix für Variablennamen.** Geben Sie Präfixe für die Namen der neuen Variablen an oder übernehmen Sie die Standardpräfixe. Variablennamen bestehen aus dem Präfix, dem Namen der zugewiesenen abhängigen Variablen und einer Modell-ID. Beim Auftreten von Namenskonflikten wird der Variablenname gegebenenfalls erweitert. Das Präfix muss den Regeln für gültige Variablennamen entsprechen.

**Modelldatei exportieren.** Modellspezifikationen für alle geschätzten Modelle werden in die angegebene XML-Datei exportiert. Mit gespeicherten Modellen können aktualisierte Vorhersagen erstellt werden.

- **XML-Datei.** Modellspezifikationen werden in einer XML-Datei gespeichert, die zusammen mit IBM SPSS-Anwendungen verwendet werden kann.
- **PMML-Datei.** Modellspezifikationen werden in einer PMML-kompatiblen XML-Datei gespeichert, die zusammen mit PMML-kompatiblen Anwendungen, einschließlich IBM SPSS-Anwendungen, verwendet werden kann.

---

## Optionen

Auf der Registerkarte "Optionen" können Sie die Vorhersageperiode, die Behandlung fehlender Werte und die Breite des Konfidenzintervalls festlegen, ein benutzerdefiniertes Präfix für Modell-IDs angeben und die für die Autokorrelationen angezeigte Anzahl von Lags festlegen.

**Vorhersageperiode.** Die Vorhersageperiode beginnt immer mit dem ersten Fall nach dem Ende der Schätzperiode (d. h. das Set von Fällen, das das Modell bestimmt) und endet entweder mit dem letzten Fall im aktiven Dataset oder an einem vom Benutzer festgelegten Datum. In der Standardeinstellung endet die Schätzperiode mit dem letzten Fall im aktiven Dataset. Das Ende kann jedoch im Dialogfeld "Fälle auswählen" durch Auswahl der Option **Nach Zeit- oder Fallbereich** geändert werden.

- **Erster Fall nach der Schätzperiode bis letzter Fall im aktiven Dataset.** Wählen Sie diese Option aus, wenn das Ende der Schätzperiode vor dem letzten Fall im aktiven Dataset liegt und die Vorhersage auch den letzten Fall erfassen soll. Über diese Option werden üblicherweise Vorhersagen für eine Prüfperiode erstellt, sodass die Modellvorhersagen mit einem Subset der tatsächlichen Werte verglichen werden können.
- **Erster Fall nach der Schätzperiode bis zum angegebenen Datum.** Wählen Sie diese Option aus, um das Ende der Vorhersageperiode explizit festzulegen. Über diese Option werden üblicherweise Vorhersagen erstellt, die über das Ende der tatsächlichen Zeitreihe hinausreichen. Geben Sie in alle Zellen des Datumsrasters Werte ein.

Wenn für das aktive Dataset kein Datum angegeben wurde, wird im Datumsraster nur die Spalte *Beobachtung* angezeigt. Um das Ende der Vorhersageperiode festzulegen, geben Sie die Zeilennummer des entsprechenden Falles ein (wie im Dateneditor angezeigt).

Die Spalte *Zyklus* im Datumsraster (sofern vorhanden) bezieht sich auf die Werte der Variablen *CYCLE\_* im aktiven Dataset.

**Benutzerdefiniert fehlende Werte.** Mit diesen Optionen wird die Behandlung benutzerdefinierter fehlender Werte gesteuert.

- **Als ungültig behandeln.** Benutzerdefiniert fehlende Werte werden wie systemdefiniert fehlende Werte behandelt.
- **Als gültig behandeln.** Benutzerdefiniert fehlende Werte werden als gültige Werte behandelt.

**Richtlinie für fehlende Werte.** Die folgenden Regeln gelten bei der Modellierungsprozedur für die Behandlung fehlender Werte. Sie gelten auch für systemdefiniert fehlende Wert und als ungültig behandelte benutzerdefiniert fehlende Werte:

- Fälle mit fehlenden Werten einer abhängigen Variablen, die innerhalb der Schätzperiode liegen, werden in das Modell eingeschlossen. Die genaue Behandlung des fehlenden Werts hängt von der Schätzmethode ab.
- Wenn eine unabhängige Variable innerhalb der Schätzperiode fehlende Werte aufweist, wird eine Warnung ausgegeben. Im Expert Modeler werden Modelle, die die unabhängige Variable enthalten, ohne die Variable geschätzt. Bei einer benutzerdefinierten ARIMA werden Modelle, die die unabhängige Variable enthalten, nicht geschätzt.
- Wenn unabhängige Variablen innerhalb der Vorhersageperiode fehlende Werte aufweisen, gibt die Prozedur eine Warnung aus und führt die Vorhersage auf der Grundlage der vorhandenen Werte aus.

**Konfidenzintervallbreite (%).** Konfidenzintervalle werden für die Modellvorhersagen und Residuen-Autokorrelationen berechnet. Es kann ein beliebiger positiver Wert unter 100 angegeben werden. In der Standardeinstellung wird ein Konfidenzintervall von 95 % verwendet.

**Präfix für Modell-IDs in Ausgabe.** Jede auf der Registerkarte "Variablen" angegebene abhängige Variable erzeugt ein eigenes geschätztes Modell. Jedes Modell weist einen eindeutigen Namen auf, der sich aus einem anpassbarem Präfix und einem ganzzahligen Suffix zusammensetzt. Sie können ein Präfix eingeben oder das vorgegebene Präfix unter *Model* übernehmen.

**In ACF- und PACF-Ausgabe angezeigte maximale Anzahl von Lags.** Sie können die Höchstanzahl von Intervallen festlegen, die in Tabellen und Diagrammen für Autokorrelationen und partielle Autokorrelationen angezeigt werden.

---

## Zusätzliche Funktionen beim Befehl TSMODEL

Sie können die Zeitreihenmodellierung an Ihre Bedürfnisse anpassen, indem Sie ihre Auswahl in ein Syntaxfenster einfügen und die resultierende Befehlssyntax für den Befehl TSMODEL bearbeiten. Mit der Befehlssyntaxsprache verfügen Sie über folgende Möglichkeiten:

- Angeben der saisonalen Periode für die Daten (mit dem Schlüsselwort SEASONLENGTH im Unterbefehl AUXILIARY). Dadurch wird die aktuelle Periodizität (sofern vorhanden) überschrieben, die bis dahin im aktiven Dataset gültig war.
- Angeben von nicht aufeinander folgenden Lags für Komponenten der benutzerdefinierten ARIMA und der Transferfunktion (mit den Unterbefehlen ARIMA und TRANSFERFUNCTION). Sie können beispielsweise ein benutzerdefiniertes ARIMA-Modell mit autoregressiven Lags der Ordnungen 1, 3 und 6 oder eine Transferfunktion mit Zählerlags der Ordnungen 2, 5 und 8 angeben.
- Angeben mehrerer Sets von Modellierungsspezifikationen (z. B. Modellierungsmethode, ARIMA-Ordnungen, unabhängige Variablen usw.) für eine Ausführung der Prozedur "Zeitreihenmodellierung" (mit dem Unterbefehl MODEL).

Vollständige Informationen zur Syntax finden Sie in der Befehlssyntaxreferenz.

---

## Kapitel 3. Zuweisen von Zeitreihenmodellen

Mit der Prozedur "Zeitreihenmodell zuweisen" werden vorhandene Zeitreihenmodelle aus einer externen Datei geladen und auf das aktive Dataset angewendet. Mit dieser Prozedur können Sie Vorhersagen für Zeitreihen berechnen, für die neue oder überarbeitete Daten verfügbar sind, ohne die Modelle erneut zu erstellen. Modelle werden mit der Prozedur Zeitreihenmodellierung erstellt.

**Beispiel:** Sie arbeiten als Bestandsmanager für einen Großhandel und sind für 5.000 Produkte zuständig. Mit dem Expert Modeler haben Sie Modelle erstellt, die den Absatz der einzelnen Produkte innerhalb der nächsten drei Monate vorhersagen. Das Data Warehouse wird jeden Monat mit den neuen Verkaufsdaten aktualisiert, mit denen Sie Ihre Vorhersagen monatlich aktualisieren möchten. Verwenden Sie hierfür die Prozedur "Zeitreihenmodell zuweisen", bei der Sie die ursprünglichen Modellen beibehalten können und lediglich die Modellparameter neu schätzen müssen, um die neuen Daten zu berücksichtigen.

**Statistiken.** Maße für die Anpassungsgüte:  $R$ -Quadrat für stationären Teil,  $R$ -Quadrat ( $R^2$ ), Wurzel der mittleren Fehlerquadratsumme (RMSE), mittlerer absoluter Fehler (MAE), mittlerer absoluter Fehler in Prozent (MAPE), maximaler absoluter Fehler (MaxAE), maximaler absoluter Fehler in Prozent (MaxAPE), normalisiertes bayessches Informationskriterium (BIC). Residuen: Autokorrelationsfunktion, partielle Autokorrelationsfunktion, Ljung-Box- $Q$ .

**Diagramme.** Diagramme für alle Modelle: Histogramme von  $R$ -Quadrat für stationären Teil,  $R$ -Quadrat ( $R^2$ ), Wurzel der mittleren Fehlerquadratsumme (RMSE), mittlerer absoluter Fehler (MAE), mittlerer absoluter Fehler in Prozent (MAPE), maximaler absoluter Fehler (MaxAE), maximaler absoluter Fehler in Prozent (MaxAPE), normalisiertes bayessches Informationskriterium (BIC), Boxplots der Residuen-Autokorrelationen und partiellen Autokorrelationen. Ergebnisse für einzelne Modelle: Vorhersagewerte, Anpassungswerte, beobachtete Werte, untere und obere Konfidenzgrenzen, Residuen-Autokorrelationen und partielle Autokorrelationen.

Erläuterungen der Daten für die Prozedur "Zeitreihenmodell zuweisen"

**Daten.** Die Variablen (abhängige und unabhängige), denen Modelle zugewiesen werden, müssen numerisch sein.

**Annahmen.** Modelle werden Variablen im aktiven Dataset zugewiesen, die die gleichen Namen wie die im Modell angegebenen Variablen tragen. Alle diese Variablen werden als Zeitreihen behandelt, d. h. jeder Fall repräsentiert einen Zeitpunkt, und die nachfolgenden Fälle liegen jeweils ein konstantes Zeitintervall auseinander.

- **Vorhersagen.** Wenn Vorhersagen anhand von Modellen mit unabhängigen Variablen (Prädiktorvariablen) erstellt werden sollen, muss das aktive Dataset für alle Fälle innerhalb der Vorhersageperiode Werte dieser Variablen enthalten. Wenn die Modellparameter erneut geschätzt werden, dürfen die unabhängigen Variablen in der Schätzperiode keine fehlenden Werte enthalten.

Datum definieren

Für die Prozedur "Zeitreihenmodell zuweisen" muss die Periodizität (sofern vorhanden) des aktiven Datensets mit der Periodizität der zuzuweisenden Modelle übereinstimmen. Wenn Sie bei der Vorhersage das gleiche Dataset (möglicherweise mit neuen oder korrigierten Daten) verwenden, mit dem auch das Modell erstellt wurde, ist diese Bedingung erfüllt. Wenn für das aktive Dataset keine Periodizität vorliegt, können Sie zum Dialogfeld Datum definieren wechseln und dort eine Periodizität erstellen. Wurden die Modelle hingegen ohne Angabe einer Periodizität erstellt, darf das aktive Dataset auch keine Periodizität aufweisen.

So weisen Sie Modelle zu:

1. Wählen Sie in den Menüs Folgendes aus:  
**Analysieren > Vorhersage > Traditionelle Modelle anwenden...**
2. Geben Sie die Dateispezifikation für eine Modelldatei ein oder klicken Sie auf **Durchsuchen** und wählen Sie eine Modelldatei aus (Modelldateien werden mit der Prozedur Zeitreihenmodellierung erstellt).

Die folgenden Optionen sind verfügbar:

- Schätzen Sie Modellparameter mithilfe der Daten im aktiven Dataset erneut. Vorhersagen werden anhand der erneut geschätzten Parameter erstellt.
- Speichern Sie Vorhersagen, Konfidenzintervalle und Restrauschen.
- Speichern Sie das erneut geschätzte Modell im XML-Format.

Modellparameter und Maße für die Anpassungsgüte

**Laden aus der Modelldatei.** Vorhersagen werden mithilfe der Modellparameter aus der Modelldatei erstellt, wobei die Parameter nicht erneut geschätzt werden. Maße für die Anpassungsgüte, die in der Ausgabe angezeigt und zum Filtern von Modellen (am besten/schlechtesten angepasste Modelle) verwendet werden, werden aus der Modelldatei abgerufen und entsprechen den Daten, die bei der Entwicklung (oder letzten Aktualisierung) des jeweiligen Modells verwendet wurden. Bei dieser Option werden bei Vorhersagen keine historischen Daten (für entweder abhängige oder unabhängige Variablen) im aktiven Dataset berücksichtigt. Sie müssen **Erneut aus den Daten schätzen** auswählen, wenn sich historische Daten auf die Vorhersage auswirken sollen. Darüber hinaus werden bei Vorhersagen Werte der abhängigen Zeitreihe in der Vorhersageperiode nicht berücksichtigt, Werte unabhängiger Variablen dagegen werden in der Vorhersageperiode berücksichtigt. Wenn Sie aktuellere Werte für die abhängige Zeitreihe haben und diese in die Vorhersagen einschließen möchten, müssen Sie eine erneute Schätzung vornehmen und die Schätzperiode entsprechend anpassen.

**Erneut aus den Daten schätzen.** Modellparameter werden mithilfe der Daten im aktiven Dataset erneut geschätzt. Eine erneute Schätzung der Modellparameter wirkt sich nicht auf die Modellstruktur aus. Ein ARIMA(1,0,1)-Modell bleibt beispielsweise unverändert, die autoregressiven Parameter und die Moving-Average-Parameter werden hingegen erneut geschätzt. Bei der erneuten Schätzung werden keine neuen Ausreißer erkannt. Ausreißer werden, sofern vorhanden, immer aus der Modelldatei abgerufen.

- **Schätzperiode.** Die Schätzperiode gibt das Set von Fällen an, anhand dessen die Modellparameter erneut geschätzt werden. In der Standardeinstellung umfasst die Schätzperiode sämtliche Fälle im aktiven Dataset. Sie können die Schätzperiode festlegen, indem Sie im Dialogfeld Fälle auswählen die Option **Nach Zeit- oder Fallbereich** auswählen. Die von der Prozedur verwendete Schätzperiode kann je nach den verfügbaren Daten unterschiedliche Modelle verwenden und daher vom angezeigten Wert abweichen. Die tatsächliche Schätzperiode für ein bestimmtes Modell entspricht der Zeitspanne, die verbleibt, nachdem die direkt aufeinander folgenden fehlenden Werte von der abhängigen Variablen des Modells entfernt wurden (am Anfang oder am Ende der angegebenen Schätzperiode).

Vorhersageperiode

Die Vorhersageperiode für die einzelnen Modelle beginnt immer mit dem ersten Fall nach dem Ende der Schätzperiode und endet entweder mit dem letzten Fall im aktiven Dataset oder an einem vom Benutzer festgelegten Datum. Wenn die Parameter nicht erneut geschätzt werden (Standardeinstellung), ist die Schätzperiode für die einzelnen Modelle das Set von Fällen, das bei der Entwicklung (oder letzten Aktualisierung) des jeweiligen Modells verwendet wurde.

- **Erster Fall nach der Schätzperiode bis letzter Fall im aktiven Dataset.** Wählen Sie diese Option aus, wenn das Ende der Schätzperiode vor dem letzten Fall im aktiven Dataset liegt und die Vorhersage auch den letzten Fall erfassen soll.
- **Erster Fall nach der Schätzperiode bis zum angegebenen Datum.** Wählen Sie diese Option aus, um das Ende der Vorhersageperiode explizit festzulegen. Geben Sie in alle Zellen des Datumsrasters Werte ein.

Wenn für das aktive Dataset kein Datum angegeben wurde, wird im Datumsraster nur die Spalte *Beobachtung* angezeigt. Um das Ende der Vorhersageperiode festzulegen, geben Sie die Zeilennummer des entsprechenden Falles ein (wie im Dateneditor angezeigt).

Die Spalte *Zyklus* im Datumsraster (sofern vorhanden) bezieht sich auf die Werte der Variablen *CYCLE\_* im aktiven Dataset.

---

## Ausgabe

Als Ausgabe stehen sowohl Ergebnisse für einzelne Modelle als auch Ergebnisse für alle Modelle zur Verfügung. Die Ergebnisse für einzelne Modelle können über vom Benutzer angegebene Kriterien auf eine Gruppe von am besten/schlechtesten angepassten Modellen eingeschränkt werden.

## Statistiken und Vorhersagetabellen

Die Registerkarte "Statistik" enthält Optionen, mit denen Tabellen für Statistiken zur Güte der Anpassung, Modellparameter, Autokorrelationsfunktionen und Vorhersagen angezeigt werden können. Sofern die Modellparameter nicht erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle"), stammen die angezeigten Werte für die Anpassungsmaße, Ljung-Box-Werte und Modellparameter aus der Modelldatei und entsprechen den Daten, die bei der Entwicklung (oder letzten Aktualisierung) des jeweiligen Modells verwendet wurden. Informationen zu Ausreißern werden immer aus der Modelldatei abgerufen.

**Anpassungsmaße, Ljung-Box-Statistik und Anzahl von Ausreißern nach Modell anzeigen.** Aktivieren Sie diese Option, um für jedes Modell eine Tabelle mit den ausgewählten Anpassungsmaße, Ljung-Box-Werten und der Anzahl von Ausreißern anzuzeigen.

**Anpassungsmaße.** Sie können eine oder mehrere der folgenden Optionen für die Aufnahme in die Tabelle auswählen, die die Anpassungsmaße für die einzelnen Modelle enthält:

- *R*-Quadrat für stationären Teil
- *R*-Quadrat
- Wurzel der mittleren Fehlerquadratsumme
- Mittlerer absoluter Fehler in Prozent
- Mittlerer absoluter Fehler
- Maximaler absoluter Fehler in Prozent
- Maximaler absoluter Fehler
- Normalisiertes BIC

**Statistik für Modellvergleich.** Mit dieser Gruppe von Optionen wird die Anzeige der Tabellen gesteuert, die die Statistiken für alle Modelle enthalten. Für jede Option wird eine eigene Tabelle erstellt. Sie können eine oder mehrere der folgenden Optionen auswählen:

- **Güte der Anpassung.** Tabelle der Auswertungsstatistiken und Perzentile für *R*-Quadrat für stationären Teil, *R*-Quadrat, Wurzel der mittleren Fehlerquadratsumme, mittlerer absoluter Fehler in Prozent, mittlerer absoluter Fehler, maximaler absoluter Fehler in Prozent, maximaler absoluter Fehler und normalisiertes Bayes-Informationskriterium.
- **Residuen-Autokorrelationsfunktion (ACF).** Tabelle der Auswertungsstatistiken und Perzentile für Autokorrelationen der Residuen für alle geschätzten Modelle. Diese Tabelle ist nur verfügbar, wenn die Modellparameter erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle").
- **Part. Residuen-Autokorrelationsfunktion (PACF).** Tabelle der Auswertungsstatistiken und Perzentile für partielle Autokorrelationen der Residuen für alle geschätzten Modelle. Diese Tabelle ist nur verfügbar, wenn die Modellparameter erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle").

**Statistik nach einzelnen Modellen.** Mit dieser Gruppe von Optionen wird die Anzeige der Tabellen gesteuert, die detaillierte Informationen zu den einzelnen Modellen enthalten. Für jede Option wird eine eigene Tabelle erstellt. Sie können eine oder mehrere der folgenden Optionen auswählen:

- **Parameterschätzungen.** Zeigt für jedes Modell eine Tabelle der Parameterschätzungen an. Für Modelle mit exponentiellem Glätten und ARIMA-Modelle werden eigene Tabellen angezeigt. Wenn Ausreißer vorhanden sind, werden Parameterschätzungen für diese ebenfalls in einer eigenen Tabelle angezeigt.
- **Residuen-Autokorrelationsfunktion (ACF).** Zeigt für jedes geschätzte Modell eine Tabelle der Residuen-Autokorrelationen nach Intervall an. Die Tabelle enthält die Konfidenzintervalle für die Autokorrelationen. Diese Tabelle ist nur verfügbar, wenn die Modellparameter erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle").
- **Part. Residuen-Autokorrelationsfunktion (PACF).** Zeigt für jedes geschätzte Modell eine Tabelle der partiellen Residuen-Autokorrelationen nach Intervall an. Die Tabelle enthält die Konfidenzintervalle für die partiellen Autokorrelationen. Diese Tabelle ist nur verfügbar, wenn die Modellparameter erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle").

**Vorhersagen anzeigen.** Zeigt eine Tabelle der Modellvorhersagen und der Konfidenzintervalle für jedes Modell an.

## Diagramme

Die Registerkarte "Diagramme" enthält Optionen, mit denen Diagramme für Statistiken zur Güte der Anpassung, Autokorrelationsfunktionen und Zeitreihenwerte (einschließlich Vorhersagen) angezeigt werden können.

Diagramme zum Vergleichen von Modellen

Mit dieser Gruppe von Optionen wird die Anzeige der Diagramme gesteuert, die die Statistiken für alle Modelle enthalten. Sofern die Modellparameter nicht erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle"), stammen die angezeigten Werte aus der Modelldatei und entsprechen den Daten, die bei der Entwicklung (oder letzten Aktualisierung) des jeweiligen Modells verwendet wurden. Autokorrelationsdiagramme sind zudem nur verfügbar, wenn die Modellparameter erneut geschätzt werden. Für jede Option wird ein eigenes Diagramm erstellt. Sie können eine oder mehrere der folgenden Optionen auswählen:

- $R$ -Quadrat für stationären Teil
- $R$ -Quadrat
- Wurzel der mittleren Fehlerquadratsumme
- Mittlerer absoluter Fehler in Prozent
- Mittlerer absoluter Fehler
- Maximaler absoluter Fehler in Prozent
- Maximaler absoluter Fehler
- Normalisiertes BIC
- Residuen-Autokorrelationsfunktion (ACF)
- Part. Residuen-Autokorrelationsfunktion (PACF)

Diagramme nach einzelnen Modellen

**Reihen.** Aktivieren Sie diese Option, um für die einzelnen Modelle Diagramme der vorhergesagten Werte zu erhalten. Beobachtete Werte, Anpassungswerte, Konfidenzintervalle für Anpassungswerte und Autokorrelationen sind nur verfügbar, wenn die Modellparameter erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle"). Sie können eine oder mehrere der folgenden Optionen für die Aufnahme in das Diagramm auswählen:

- **Beobachtete Werte.** Die beobachteten Werte der abhängigen Zeitreihe.
- **Vorhersagen.** Die vom Modell vorhergesagten Werte für die Vorhersageperiode.

- **Anpassungswerte.** Die vom Modell vorhergesagten Werte für die Schätzperiode.
- **Konfidenzintervalle für Vorhersagen.** Die Konfidenzintervalle für die Vorhersageperiode.
- **Konfidenzintervalle für Anpassungswerte.** Die Konfidenzintervalle für die Schätzperiode.

**Residuen-Autokorrelationsfunktion (ACF).** Zeigt für jedes geschätzte Modell ein Diagramm der Residuen-Autokorrelationen an.

**Part. Residuen-Autokorrelationsfunktion (PACF).** Zeigt für jedes geschätzte Modell ein Diagramm der partiellen Residuen-Autokorrelationen an.

## Einschränken der Ausgabe auf die am besten/schlechtesten angepassten Modelle

Die Registerkarte "Ausgabefilter" enthält Optionen, mit denen sowohl Tabellen- als auch Diagrammausgabe auf ein Subset von Modellen beschränkt werden kann. Sie können die Ausgabe mithilfe von Anpassungskriterien auf die am besten und/oder am schlechtesten angepassten Modelle beschränken. In der Standardeinstellung sind alle Modelle in der Ausgabe enthalten. Sofern die Modellparameter nicht erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle"), stammen die zum Filtern der Modelle verwendeten Werte der Anpassungsmaße aus der Modelldatei und entsprechen den Daten, die bei der Entwicklung (oder letzten Aktualisierung) des jeweiligen Modells verwendet wurden.

**Am besten angepasste Modelle.** Aktivieren Sie diese Option, um die am besten angepassten Modelle in die Ausgabe einzuschließen. Wählen Sie ein Maß für Anpassungsgüte und geben Sie die Anzahl von Modellen an, die eingeschlossen werden sollen. Die Auswahl dieser Option hindert Sie nicht daran, auch die am schlechtesten angepassten Modelle auszuwählen. Die Ausgabe würde in diesem Fall sowohl die am schlechtesten als auch die am besten angepassten Modelle enthalten.

- **Feste Anzahl von Modellen.** Gibt an, dass Ergebnisse für die  $n$  am besten angepassten Modelle angezeigt werden. Wenn die Anzahl die Gesamtanzahl von Modellen überschreitet, werden alle Modelle angezeigt.
- **Prozentsatz der Gesamtzahl von Modellen.** Gibt an, dass Ergebnisse für Modelle angezeigt werden, deren Werte für die Anpassungsgüte für alle Modelle in den obersten  $n$  Prozent liegen.

**Am schlechtesten angepasste Modelle.** Aktivieren Sie diese Option, um die am schlechtesten angepassten Modelle in die Ausgabe einzuschließen. Wählen Sie ein Maß für Anpassungsgüte und geben Sie die Anzahl von Modellen an, die eingeschlossen werden sollen. Die Auswahl dieser Option hindert Sie nicht daran, auch die am besten angepassten Modelle auszuwählen. Die Ausgabe würde in diesem Fall sowohl die am besten als auch die am schlechtesten angepassten Modelle enthalten.

- **Feste Anzahl von Modellen.** Gibt an, dass Ergebnisse für die  $n$  am schlechtesten angepassten Modelle angezeigt werden. Wenn die Anzahl die Gesamtanzahl von Modellen überschreitet, werden alle Modelle angezeigt.
- **Prozentsatz der Gesamtzahl von Modellen.** Gibt an, dass Ergebnisse für Modelle angezeigt werden, deren Werte für die Anpassungsgüte für alle Modelle in den untersten  $n$  Prozent liegen.

**Maß für Anpassungsgüte.** Wählen Sie das Maß für die Anpassungsgüte, anhand dessen die Modelle gefiltert werden sollen. Der Standardwert ist  $R$ -Quadrat.

---

## Speichern von Modellvorhersagen und Modellspezifikationen

Auf der Registerkarte "Speichern" können Sie festlegen, dass Modellvorhersagen als neue Variablen im aktiven Dataset gespeichert werden sollen. Außerdem können die Modellspezifikationen in einer externen Datei im XML-Format gespeichert werden.

**Variablen speichern.** Sie können Modellvorhersagen, Konfidenzintervalle und Residuen als neue Variablen im aktiven Dataset speichern. Jedes Modell erzeugt eine eigene Gruppe neuer Variablen. Wenn die Vorhersageperiode die Länge der abhängigen Zeitreihe überschreitet, die dem Modell zugewiesen ist,

werden neue Fälle hinzugefügt. Sofern die Modellparameter nicht erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle"), werden die vorhergesagten Werte und Konfidenzgrenzen nur für die Vorhersageperiode erstellt. Wenn die neuen Variablen gespeichert werden sollen, aktivieren Sie jeweils das zugehörige Kontrollkästchen "Speichern". In der Standardeinstellung werden keine neuen Variablen gespeichert.

- **Vorhergesagte Werte.** Die vom Modell vorhergesagten Werte.
- **Untere Konfidenzgrenzen.** Die unteren Konfidenzgrenzen für die vorhergesagten Werte.
- **Obere Konfidenzgrenzen.** Die oberen Konfidenzgrenzen für die vorhergesagten Werte.
- **Restrauschen.** Die Modellresiduen. Bei der Transformation von abhängigen Variablen (z. B. mit natürlichem Logarithmus) sind dies die Residuen für die transformierte Reihe. Diese Option ist nur verfügbar, wenn die Modellparameter erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle").
- **Präfix für Variablennamen.** Geben Sie Präfixe für die Namen der neuen Variablen an oder übernehmen Sie die Standardpräfixe. Variablennamen bestehen aus dem Präfix, dem Namen der zugewiesenen abhängigen Variablen und einer Modell-ID. Beim Auftreten von Namenskonflikten wird der Variablenname gegebenenfalls erweitert. Das Präfix muss den Regeln für gültige Variablennamen entsprechen.

**Modelldatei exportieren** Modellspezifikationen mit erneut geschätzten Parametern und Statistiken für Anpassungsgüte werden im XML-Format in die angegebene Datei exportiert. Diese Option ist nur verfügbar, wenn die Modellparameter erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle").

- **XML-Datei.** Modellspezifikationen werden in einer XML-Datei gespeichert, die zusammen mit IBM SPSS-Anwendungen verwendet werden kann.
- **PMML-Datei.** Modellspezifikationen werden in einer PMML-kompatiblen XML-Datei gespeichert, die zusammen mit PMML-kompatiblen Anwendungen, einschließlich IBM SPSS-Anwendungen, verwendet werden kann.

---

## Optionen

Auf der Registerkarte "Optionen" können Sie die Behandlung fehlender Werte, die Breite des Konfidenzintervalls und die für die Autokorrelationen angezeigte Anzahl von Lags festlegen.

**Benutzerdefiniert fehlende Werte.** Mit diesen Optionen wird die Behandlung benutzerdefinierter fehlender Werte gesteuert.

- **Als ungültig behandeln.** Benutzerdefiniert fehlende Werte werden wie systemdefiniert fehlende Werte behandelt.
- **Als gültig behandeln.** Benutzerdefiniert fehlende Werte werden als gültige Werte behandelt.

**Richtlinie für fehlende Werte.** Die folgenden Regeln gelten für die Behandlung fehlender Werte. Sie gelten auch für systemdefiniert fehlende Wert und als ungültig behandelte benutzerdefiniert fehlende Werte:

- Fälle mit fehlenden Werten einer abhängigen Variablen, die innerhalb der Schätzperiode liegen, werden in das Modell eingeschlossen. Die genaue Behandlung des fehlenden Werts hängt von der Schätzmethode ab.
- Bei ARIMA-Modellen wird eine Warnung ausgegeben, wenn ein Prädiktor innerhalb der Schätzperiode fehlende Werte aufweist. Modelle mit Prädiktoren werden nicht erneut geschätzt.
- Wenn unabhängige Variablen innerhalb der Vorhersageperiode fehlende Werte aufweisen, gibt die Prozedur eine Warnung aus und führt die Vorhersage auf der Grundlage der vorhandenen Werte aus.

**Konfidenzintervallbreite (%).** Konfidenzintervalle werden für die Modellvorhersagen und Residuen-Autokorrelationen berechnet. Es kann ein beliebiger positiver Wert unter 100 angegeben werden. In der Standardeinstellung wird ein Konfidenzintervall von 95 % verwendet.

In **ACF- und PACF-Ausgabe** angezeigte **maximale Anzahl von Lags**. Sie können die Höchstanzahl von Intervallen festlegen, die in Tabellen und Diagrammen für Autokorrelationen und partielle Autokorrelationen angezeigt werden. Diese Option ist nur verfügbar, wenn die Modellparameter erneut geschätzt werden (**Erneut aus den Daten schätzen** auf der Registerkarte "Modelle").

---

## Zusätzliche Funktionen beim Befehl TSAPPLY

Ihnen stehen zusätzliche Funktionen zur Verfügung, wenn Sie Ihre Auswahl in ein Syntaxfenster einfügen und die resultierende Befehlssyntax für den Befehl TSAPPLY bearbeiten. Mit der Befehlssyntaxsprache verfügen Sie über folgende Möglichkeiten:

- Festlegen, dass dem aktiven Dataset nur ein Subset der Modelle in einer Modelldatei zugewiesen werden (mit den Schlüsselwörtern DROP und KEEP im Unterbefehl MODEL).
- Den Daten Modelle aus zwei oder mehr Modelldateien zuweisen (mit dem Unterbefehl MODEL). So könnte eine Modelldatei beispielsweise Modelle für Zeitreihen enthalten, die Verkaufsstückzahlen repräsentieren, und eine andere Modelle für Zeitreihen, die für Verkaufserlöse stehen.

Vollständige Informationen zur Syntax finden Sie in der Befehlssyntaxreferenz.



---

## Kapitel 4. Saisonale Zerlegung

Die Prozedur "Saisonale Zerlegung" zerlegt eine Zeitreihe in eine saisonale Komponente, eine kombinierte Trend- und Zykluskomponente sowie eine "Fehler"-Komponente. Die Prozedur ist eine Umsetzung der Census-I-Methode, die auch als "Verhältnis zum gleitenden Durchschnitt" bekannt ist.

**Beispiel.** Ein Wissenschaftler interessiert sich für die Analyse monatlicher Messungen des Ozongehalts an einer bestimmten Wetterstation. Ziel ist es herauszufinden, ob bei den Daten ein Trend vorliegt. Um einen etwaigen wirklichen Trend aufzudecken, muss der Wissenschaftler zunächst die Schwankungen in den Messergebnissen berücksichtigen, die auf saisonale Effekte zurückzuführen sind. Mit der Prozedur "Saisonale Zerlegung" können etwaige systematische saisonale Schwankungen entfernt werden. Die Trendanalyse wird dann auf eine saisonbereinigte Zeitreihe angewendet.

**Statistiken.** Das Set der saisonalen Faktoren.

Erläuterungen der Daten für die saisonale Zerlegung

**Daten.** Die Variablen müssen numerisch sein.

**Annahmen.** Die Variablen dürfen keine eingebetteten fehlenden Daten enthalten. Es muss mindestens eine periodische Datumskomponente definiert sein.

Schätzen von saisonalen Faktoren

1. Wählen Sie in den Menüs Folgendes aus:

**Analysieren > Vorhersage > Saisonale Zerlegung...**

2. Wählen Sie eine oder mehrere Variablen aus der Liste der verfügbaren Variablen und verschieben Sie sie in die Liste "Variable(n)". Achten Sie darauf, dass die Liste nur numerische Variablen enthält.

**Modelltyp.** Die Prozedur "Saisonale Zerlegung" bietet zwei verschiedene Ansätze für die Modellierung der saisonalen Faktoren: multiplikativ oder kumulativ.

- *Multiplikativ.* Die saisonale Komponente ist ein Faktor, mit dem die saisonbereinigte Zeitreihe multipliziert wird, um die ursprüngliche Zeitreihe zu ergeben. Tatsächlich saisonale Komponenten, die proportional zum Gesamtniveau der Zeitreihe sind. Beobachtungen ohne saisonale Variation haben eine saisonale Komponente von 1.
- *Kumulativ.* Die saisonalen Anpassungen werden zu der saisonbereinigten Zeitreihe addiert, um die beobachteten Werte zu erhalten. Durch die Saisonbereinigung soll der saisonale Effekt aus einer Zeitreihe entfernt werden, sodass möglicherweise durch die saisonale Komponente verdeckte interessante Eigenschaften der Zeitreihe betrachtet werden können. Tatsächlich geht es um saisonale Komponenten, die nicht vom Niveau der Zeitreihe abhängen. Beobachtungen ohne saisonale Variation haben die saisonale Komponente 0.

**Gewichtung für gleitenden Durchschnitt.** Mit den Optionen im Gruppenfeld "Gewichtung für gleitenden Durchschnitt" können Sie festlegen, wie die Zeitreihen beim Berechnen der gleitenden Durchschnitte behandelt werden sollen. Diese Optionen sind nur verfügbar, wenn die Periodizität der Zeitreihen regelmäßig ist. Falls die Periodizität unregelmäßig ist, werden alle Punkte gleich gewichtet.

- *Alle Punkte gleich.* Gleitende Durchschnitte werden mit einer Spanne berechnet, die gleich der Periodizität ist, und so, dass alle Punkte die gleiche Gewichtung erhalten. Diese Methode wird immer dann verwendet, wenn die Periodizität ungerade ist.
- *Endpunkte gewichtet mit 0,5.* Gleitende Durchschnitte für Zeitreihen mit gerader Periodizität werden mit einer Spanne berechnet, die gleich der Periodizität plus 1 ist; dabei werden die Endpunkte der Spanne mit 0,5 gewichtet.

Die folgenden Optionen sind verfügbar:

- Klicken Sie auf **Speichern**, um anzugeben, wie neue Variablen gespeichert werden sollen.

---

## Saisonale Zerlegung: Speichern

**Variablen erstellen.** Hier können Sie auswählen, wie neue Variablen behandelt werden sollen.

- *Zur Datei hinzufügen.* Die neuen Zeitreihen, die durch die saisonale Zerlegung erstellt wurden, werden als reguläre Variablen im aktiven Dataset gespeichert. Variablennamen werden aus einem aus drei Buchstaben bestehenden Präfix, einem Unterstrich und einer Zahl gebildet.
- *Vorhandene ersetzen.* Die neuen Zeitreihen, die durch die saisonale Zerlegung erstellt wurden, werden als temporäre Variablen im aktiven Dataset gespeichert. Gleichzeitig werden alle eventuell vorhandenen temporären Variablen entfernt, die durch Vorhersagebefehle erzeugt wurden. Variablennamen werden aus einem aus drei Buchstaben bestehenden Präfix, einer Raute (#) und einer Zahl gebildet.
- *Nicht erstellen.* Die neuen Zeitreihen werden nicht zum aktiven Dataset hinzugefügt.

Benennung der neuen Variablen

Die Prozedur "Saisonale Zerlegung" erstellt vier neue Variablen (Zeitreihen) mit den folgenden dreibuchstabigen Präfixen für die einzelnen Reihen:

**SAF.** *Faktoren für die Saisonbereinigung (Seasonal Adjustment Factors).* Diese Werte geben die Auswirkung der einzelnen Perioden auf das Niveau der Zeitreihe an.

**SAS.** *Saisonbereinigte Zeitreihe (Seasonally Adjusted Series).* Dies sind die Werte, die sich nach dem Entfernen saisonaler Schwankungen einer Zeitreihe ergeben.

**STC.** *Geglättete Trendzykluskomponenten (Smoothed Trend-Cycle Components).* Diese Werte zeigen den Trend und das zyklische Verhalten in einer Zeitreihe an.

**ERR.** *Residuen- oder fehlerhafte Werte.* Die Werte, die nach dem Entfernen der saisonalen, Trend- und Zykluskomponenten verbleiben.

---

## Zusätzliche Funktionen beim Befehl SEASON

Die Befehlssyntax ermöglicht außerdem Folgendes:

- Angabe einer etwaigen Periodizität im Befehl SEASON, anstatt eine der Alternativen der Prozedur "Datum definieren" auszuwählen.

Vollständige Informationen zur Syntax finden Sie in der Befehlssyntaxreferenz.

---

## Kapitel 5. Spektraldiagramme

Die Prozedur "Spektraldiagramme" wird zur Identifizierung von periodischem Verhalten bei Zeitreihen verwendet. Anstatt die Schwankungen von einem Zeitpunkt zum nächsten zu analysieren, werden die Schwankungen der Zeitreihe als Ganzes in periodischen Komponenten unterschiedlicher Frequenzen analysiert. Glatte Zeitreihen weisen stärkere periodische Komponenten bei niedrigen Frequenzen auf; durch Zufallsschwankungen ("weißes Rauschen") wird die Komponentenstärke über alle Frequenzen verteilt.

Zeitreihen, die fehlende Daten enthalten, können mit dieser Prozedur nicht analysiert werden.

**Beispiel.** Die Neubauquote ist ein wichtiger Standardwert für die gesamtwirtschaftliche Lage. Die Daten für Baubeginne weisen typischerweise eine starke saisonale Komponente auf. Sind jedoch auch längere Zyklen in den Daten zu finden, deren sich die Analytiker bei der Auswertung der aktuellen Zahlen bewusst sein müssen?

**Statistiken.** Sinus- und Kosinustransformationen, Periodogrammwert und Spektraldichteschätzung für jede Frequenz- bzw. Periodenkomponente. Bei Auswahl von bivariater Analyse: Real- und Imaginärteile des Kreuzperiodogramms, Kospektraldichte, Quadraturpektrum, quadrierte Kohärenz und Phasenspektrum für jede Frequenz- bzw. Periodenkomponente.

**Diagramme.** Für univariate und bivariate Analysen: Periodogramm und Spektraldichte. Für bivariate Analysen: quadrierte Kohärenz, Quadraturpektrum, Kreuzamplitude, Kospektraldichte, Phasenspektrum und Gewinn.

Erläuterungen der Daten für Spektraldiagramme

**Daten.** Die Variablen müssen numerisch sein.

**Annahmen.** Die Variablen dürfen keine eingebetteten fehlenden Daten enthalten. Die zu analysierende Zeitreihe muss stationär sein und ein Mittelwert ungleich 0 muss aus der Zeitreihe subtrahiert werden.

- *Stationär.* Eine Bedingung, die von Zeitreihen eingehalten werden muss, an die ARIMA-Modelle angepasst werden sollen. Reine MA-Reihen sind stationär, aber AR- und ARMA-Reihen können auch nicht stationär sein. Eine stationäre Zeitreihe besitzt einen konstanten Mittelwert und eine konstante Varianz über die Zeit.

Berechnen einer Spektralanalyse

1. Wählen Sie in den Menüs Folgendes aus:

**Analyse > Zeitreihe > Spektralanalyse...**

2. Wählen Sie eine oder mehrere Variablen aus der Liste der verfügbaren Variablen und verschieben Sie sie in die Liste "Variable(n)". Achten Sie darauf, dass die Liste nur numerische Variablen enthält.
  3. Wählen Sie eine der Optionen im Gruppenfeld "Spektralfenster", um festzulegen, wie das Periodogramm geglättet werden soll, um eine Spektraldichteschätzung zu erhalten. Folgende Glättungsoptionen stehen zur Verfügung: "Tukey-Hamming", "Tukey", "Parzen", "Bartlett", "Daniell (Einheit)" und "Keine"
- *Tukey-Hamming.* Die Gewichtungen lauten:  $W_k = 0,54D_p(2 \pi f_k) + 0,23D_p(2 \pi f_k + \pi/p) + 0,23D_p(2 \pi f_k - \pi/p)$ , für  $k = 0, \dots, p$ . Dabei ist  $p$  der ganzzahlige Anteil der Hälfte der Spannweite und  $D_p$  ist der Dirichlet-Kern der Ordnung  $p$ .
  - *Tukey.* Die Gewichtungen lauten:  $W_k = 0,5D_p(2 \pi f_k) + 0,25D_p(2 \pi f_k + \pi/p) + 0,25D_p(2 \pi f_k - \pi/p)$ , für  $k = 0, \dots, p$ . Hierbei ist  $p$  der ganzzahlige Teil der Spannweitenhälfte und  $D_p$  ist der Dirichlet-Kern der Ordnung  $p$ .

- *Parzen*. Die Gewichtungen sind  $W_k = 1/p(2 + \cos(2\pi fk)) (F[p/2] (2\pi fk))^{**2}$ , für  $k = 0, \dots, p$ . Dabei ist  $p$  der ganzzahlige Teil der halben Spannweite und  $F[p/2]$  ist der Fejérsche Kern der Ordnung  $p/2$ .
- *Bartlett*. Die Form eines Spektralfensters, bei dem die Gewichtungen der oberen Hälfte des Fensters mit der Gleichung  $W_k = F_p (2\pi fk)$ , für  $k = 0, \dots, p$ , berechnet werden. Dabei bezeichnet  $p$  den ganzzahligen Anteil der halben Spannweite.  $F_p$  ist der fejérsche Kern der Ordnung  $p$ . Die untere Hälfte des Fensters ist symmetrisch zur oberen.
- *Daniell (Einheit)*. Die Form eines Spektralfensters, für das alle Gewichtungen gleich 1 sind.
- *Keine*. Keine Glättung. Wenn diese Option ausgewählt wird, entspricht die Schätzung der Spektraldichte dem Periodogramm.

*Spannweite*. Der Bereich aufeinanderfolgender Werte, für den die Glättung durchgeführt wird. In der Regel werden ungerade ganze Zahlen verwendet. Bei größeren Spannweiten ist die Glättung der Spektraldichtediagramme größer als bei kleineren Spannweiten.

*Variablen zentrieren*. Passt die Zeitreihe so an, dass sie vor der Berechnung des Spektrums den Mittelwert 0 hat, sodass der eventuell mit dem Mittelwert zusammenhängende große Term entfernt wird.

**Bivariate Analyse—Erste Variable mit jeder**. Wenn Sie zwei oder mehr Variablen ausgewählt haben, können Sie diese Option aktivieren, um eine bivariate Spektralanalyse berechnen zu lassen.

- Die erste Variable in der Liste "Variable(n)" wird als unabhängige Variable behandelt und alle übrigen Variablen gelten als abhängige Variablen.
- Jede auf die erste Zeitreihe folgende Zeitreihe wird unabhängig von anderen festgelegten Zeitreihen mit der ersten Zeitreihe analysiert. Außerdem werden univariate Analysen jeder Zeitreihe durchgeführt.

**Diagramm**. Periodogramm und Spektraldichte sind sowohl für univariate und bivariate Analysen verfügbar. Alle anderen Optionen stehen nur für bivariate Analysen zur Verfügung.

- *Periodogramm*. Ein ungeglättetes Diagramm der Spektralampplitude (dargestellt auf einer logarithmischen Skala) in Abhängigkeit von Frequenz oder Periode. Eine niederfrequente Streuung charakterisiert eine glatte Zeitreihe. Eine gleichmäßig über alle Frequenzen verteilte Streuung ist ein Zeichen für "weißes Rauschen".
- *Quadrierte Kohärenz*. Das Produkt der Zunahmewerte der beiden Zeitreihen.
- *Quadraturpektrum*. Der Imaginärteil des Kreuzperiodogramms; er ist ein Maß für die Korrelation der phasenverschobenen Frequenzkomponenten der beiden Zeitreihen. Die Komponenten sind um  $\pi/2$  phasenverschoben.
- *Kreuzamplitude*. Die Wurzel aus der Summe von quadrierter Kospektraldichte und quadriertem Quadraturpektrum.
- *Spektraldichte*. Ein Periodogramm, das geglättet wurde, um unregelmäßige Variationen zu entfernen.
- *Ko-Spektraldichte*. Der Realteil des Kreuzperiodogramms; er ist ein Maß für die Korrelation der phasengleichen Frequenzkomponenten der beiden Zeitreihen.
- *Phasenspektrum*. Ein Maß dafür, inwieweit jede Frequenzkomponente einer Zeitreihe der anderen Zeitreihe voraus- oder nachläuft.
- *Zunahme (Gewinn)*. Der Quotient aus Kreuzamplitude und Spektraldichte für eine der Zeitreihen. Jede der beiden Zeitreihen hat einen eigenen Zunahmewert.

*Nach Häufigkeit*. Alle Diagramme werden nach Häufigkeit erstellt. Der Häufigkeitsbereich reicht von 0 (der konstante Term oder Mittelwertterm) bis 0,5 (der Term für einen Zyklus von zwei Beobachtungen).

*Nach Periode*. Alle Diagramme werden nach der Periode erstellt. Der Periodenbereich reicht von 2 (der Term für einen Zyklus von zwei Beobachtungen) bis zur Anzahl der Beobachtungen (der konstante Term oder Mittelwertterm). Die Periode wird auf einer logarithmischen Skala dargestellt.

---

## Zusätzliche Funktionen beim Befehl SPECTRA

Die Befehlssyntax ermöglicht außerdem Folgendes:

- Die berechneten Spektralanalysevariablen können für eine spätere Verwendung im aktiven Dataset gespeichert werden.
- Für das Spektralfenster können benutzerdefinierte Gewichtungen festgelegt werden.
- Es können Diagramme sowohl nach Frequenz als auch nach Periode erstellt werden.
- Es kann eine vollständige Auflistung jedes im Diagramm angezeigten Werts ausgegeben werden.

Vollständige Informationen zur Syntax finden Sie in der Befehlssyntaxreferenz.



---

## Kapitel 6. Temporale kausale Modelle

Mithilfe der temporalen kausalen Modellierung lassen sich kausale Schlüsselbeziehungen in Zeitreihendaten erkennen. Bei der temporalen kausalen Modellierung geben Sie ein Set von Zielzeitreihen und ein Set von möglichen Eingaben für diese Ziele an. Daraufhin erstellt die Prozedur ein autoregressives Zeitreihenmodell für jedes der Ziele und schließt nur solche Eingaben ein, die eine kausale Beziehung zum Ziel haben. Diese Methode unterscheidet sich von der traditionellen Zeitreihenmodellierung, bei der Sie die Prädiktoren für eine Zielzeitreihe explizit angeben müssen. Da bei der temporalen kausalen Modellierung üblicherweise Modelle für mehrere zusammengehörige Zeitreihen erstellt werden, wird das Ergebnis als *Modellsystem* bezeichnet.

Im Zusammenhang mit der temporalen kausalen Modellierung bezieht sich der Begriff *kausal* auf die Granger-Kausalität. Nach dem Granger-Kausalitätsprinzip führt eine Zeitreihe X zu einer Zeitreihe Y, wenn eine Regression für Y in Bezug auf vorangehende Werte sowohl von X als auch von Y ein besseres Modell für Y hervorbringt, als wenn nur die vorangehenden Werte von Y einer Regression unterzogen worden wären.

### Beispiele

Entscheidungssträger in Unternehmen können die temporale kausale Modellierung verwenden, um innerhalb einer großen Menge zeitbasierter Metriken zur Beschreibung des Unternehmens kausale Beziehungen zu erkennen. Die Analyse macht möglicherweise einige steuerbare Eingaben sichtbar, die die größten Auswirkungen auf wesentliche Leistungsindikatoren haben.

Manager großer IT-Systeme können mithilfe der temporalen kausalen Modellierung Anomalien innerhalb einer großen Menge in Wechselbeziehung zueinander stehender operativer Metriken erkennen. Das kausale Modell ermöglicht dann über die Erkennung dieser Anomalien hinaus die Aufspürung ihrer wahrscheinlichsten Ursachen.

### Feldanforderungen

Es muss mindestens ein Ziel vorhanden sein. Standardmäßig werden Felder mit der vordefinierten Rolle "Keine" nicht verwendet.

### Datenstruktur

Die temporale kausale Modellierung unterstützt zwei Datenstrukturtypen.

#### Spaltenbasierte Daten

Bei spaltenbasierten Daten enthält jedes Zeitreihenfeld die Daten für eine einzelne Zeitreihe. Dies ist die traditionelle Struktur für Zeitreihendaten, wie sie von der Zeitreihenmodellierung verwendet wird.

#### Mehrdimensionale Daten

Bei mehrdimensionalen Daten enthält jedes Zeitreihenfeld die Daten für mehrere Zeitreihen. Innerhalb eines bestimmten Felds werden separate Zeitreihen dann durch ein Set von Werten für kategoriale Felder ermittelt, die als *Dimensionsfelder* bezeichnet werden. Beispielsweise können Verkaufsdaten für zwei verschiedene Vertriebskanäle ('retail' und 'web') in einem einzigen Feld *sales* gespeichert werden. Ein Dimensionsfeld mit der Benennung *channel* und den Werten 'retail' und 'web' gibt die Datensätze an, die jedem der beiden Verkaufskanäle zugeordnet sind.

---

## Erstellen eines temporalen kausalen Modells

Für diese Funktion ist die Option "Statistics Forecasting" erforderlich.

Wählen Sie in den Menüs Folgendes aus:

### Analysieren > Vorhersage > Temporale kausale Modelle erstellen...

1. Wenn die Beobachtungen durch ein Datums-/Zeitfeld definiert werden, geben Sie das entsprechende Feld an.
2. Handelt es sich um mehrdimensionale Daten, geben Sie die Dimensionsfelder an, die die Zeitreihen kennzeichnen.
  - Die angegebene Reihenfolge der Dimensionsfelder definiert die Reihenfolge, in der die Felder in allen nachfolgenden Dialogfeldern und Ausgaben angezeigt werden. Mit den Auf- und Abwärtspfeilen können Sie die Reihenfolge der Dimensionsfelder ändern.
  - Sie können Subsets von Dimensionswerten angeben, um die Analyse auf bestimmte Werte für die Dimensionsfelder zu begrenzen. Bei Dimensionen für Region und Marke können Sie die Analyse z. B. auf eine bestimmte Region begrenzen. Dimensionssubsets gelten für alle in der Analyse verwendeten Metrikfelder.
  - Sie können die Analyse auch anpassen, indem Sie Dimensionswerte nach Rollen auf der Metrikfeldenebene angeben. Wenn Sie beispielsweise eine Dimension für Vertriebskanäle (mit den Werten 'retail' und 'web') und in diesen Kanälen Metriken für Verkauf und Werbung haben, können Sie Internetwerbung als Eingabe sowohl für Einzelhandels- als auch für Internetverkäufe angeben. Standardmäßig ist dieser Anpassungstyp aktiviert und die Auswahl unterschiedlicher Werte ist auf eine bestimmte maximale Anzahl (standardmäßig 250) je Dimensionsfeld begrenzt.
3. Klicken Sie auf **Weiter**.

**Anmerkung:** Die Schritte 1, 2 und 3 entfallen, wenn das aktive Dataset eine Datumsspezifikation aufweist. Datumsspezifikationen werden über das Dialogfeld **Datum definieren** oder den Befehl DATE erstellt.

4. Klicken Sie auf **Felder**, um die Zeitreihe anzugeben, die in das Modell eingeschlossen werden soll, und um anzugeben, wie die Beobachtungen definiert werden. Mindestens ein Feld muss entweder als Ziel oder sowohl als Eingabe als auch als Ziel angegeben werden.
5. Klicken Sie auf **Datenspezifikationen**, um optionale Einstellungen anzugeben, z. B. das Zeitintervall für die Analyse, Einstellungen für Aggregation und Verteilung sowie für die Handhabung fehlender Werte.
6. Klicken Sie auf **Erstellungsoptionen**, um die Schätzperiode zu definieren, den Inhalt der Ausgabe festzulegen und Einstellungen für die Erstellung anzugeben, z. B. die maximale Anzahl Eingaben je Ziel.
7. Klicken Sie auf **Modelloptionen**, um Vorhersagen anzufordern, Vorhersagen zu speichern und das Modellsystem in eine externe Datei zu exportieren.
8. Klicken Sie auf **Ausführen**, um die Prozedur auszuführen.

---

## Zeitreihen für die Modellierung

Verwenden Sie auf der Registerkarte **Felder** die Einstellungen für **Zeitreihe**, um die Zeitreihe anzugeben, die in das Modellsystem eingeschlossen werden soll.

Bei spaltenbasierten Daten ist der Begriff *Zeitreihe* gleichbedeutend mit dem Begriff *Feld*. Bei mehrdimensionalen Daten werden Felder, die Zeitreihen enthalten, als *Metrikfelder* bezeichnet. Eine Zeitreihe wird bei mehrdimensionalen Daten durch ein Metrikfeld und je einen Wert für die einzelnen Dimensionsfelder definiert. Folgende Aspekte gelten sowohl für spaltenbasierte als auch für mehrdimensionale Daten.

- Zeitreihen, die als mögliche Eingaben oder als sowohl Ziel als auch Eingabe angegeben sind, können in das Modell der einzelnen Ziele eingeschlossen werden. Die Modelle der einzelnen Ziele beinhalten stets Lagwerte des Ziels selbst.
- Zeitreihen, die als erzwungene Eingaben angegeben sind, werden immer in das Modell eines Ziels eingeschlossen.
- Mindestens eine Zeitreihe muss entweder als Ziel oder sowohl als Ziel als auch als Eingabe angegeben werden.
- Wenn **Vordefinierte Rollen verwenden** ausgewählt ist, werden Felder mit der Rolle "Eingabe" als mögliche Eingaben festgelegt. Bei Feldern ohne vordefinierte Rolle erfolgt die Zuordnung zu den erzwungenen Eingaben.

## Mehrdimensionale Daten

Bei mehrdimensionalen Daten geben Sie Metrikfelder und zugeordnete Rollen in einem Raster an, wobei jede Zeile im Raster je eine Metrik und eine Rolle angibt. Standardmäßig umfasst das Modellsystem Zeitreihen für alle Dimensionsfelderkombinationen zu jeder Zeile im Raster. Wenn beispielsweise Dimensionen für *region* und *brand* vorhanden sind, gibt es bei Angabe der Metrik *sales* als Ziel zu jeder Kombination aus *region* und *brand* standardmäßig eine separate Zielzeitreihe für *sales*.

In den einzelnen Rasterzeilen können Sie das Set von Werten für jedes der Dimensionsfelder anpassen, indem Sie für die jeweilige Dimension auf die Schaltfläche mit Auslassungspunkten klicken. Daraufhin wird das Unterdialogfeld **Dimensionswerte auswählen** geöffnet. Sie können Rasterzeilen auch hinzufügen, löschen oder kopieren.

Die Spalte **Anzahl Zeitreihen** zeigt die Anzahl der Sets von Dimensionswerten an, die aktuell für die zugehörige Metrik angegeben sind. Der angezeigte Wert kann größer als die tatsächliche Anzahl Zeitreihen sein (eine Zeitreihe je Set). Diese Bedingung liegt vor, wenn einige der angegebenen Kombinationen von Dimensionswerten nicht den in der zugeordneten Metrik enthaltenen Zeitreihen entsprechen.

## Auswählen von Dimensionswerten

Bei mehrdimensionalen Daten können Sie die Analyse anpassen, indem Sie angeben, welche Dimensionswerte für ein bestimmtes Metrikfeld mit einer bestimmten Rolle gelten. Wenn z. B. *sales* ein Metrikfeld ist und *channel* eine Dimension mit den Werten 'retail' und 'web', können Sie festlegen, dass 'web' eine Eingabe und 'retail' ein Ziel ist. Sie können auch Dimensionssubsets angeben, die für alle in der Analyse verwendeten Metrikfelder gelten. Wenn beispielsweise *region* ein Dimensionsfeld zur Angabe der geografischen Region ist, können Sie die Analyse auf bestimmte Regionen begrenzen.

### Alle Werte

Gibt an, dass alle Werte des aktuellen Dimensionsfelds eingeschlossen sind. Dies ist die Standardoption.

### Ein- bzw. auszuschließende Werte auswählen

Mit dieser Option können Sie das Set von Werten für das aktuelle Dimensionsfeld angeben. Wenn für **Modus** die Option **Einschließen** ausgewählt ist, werden nur Werte eingeschlossen, die in der Liste **Ausgewählte Werte** aufgeführt sind. Wird stattdessen für **Modus** die Option **Ausschließen** ausgewählt, werden alle Werte eingeschlossen, die nicht in der Liste **Ausgewählte Werte** stehen.

Sie können in dem Set von Werten für die Auswahl Filter setzen. Werte, die mit der Filterbedingung übereinstimmen, sind auf der Registerkarte **Übereinstimmung** aufgeführt, während Werte, die mit dieser Bedingung nicht übereinstimmen, auf der Registerkarte **Keine Übereinstimmung** der Liste **Nicht ausgewählte Werte** angezeigt werden. Auf der Registerkarte **Alle** werden alle nicht ausgewählten Werte unabhängig von Filterbedingungen aufgelistet.

- Sie können Sterne (\*) verwenden, um beim Setzen von Filtern Platzhalterzeichen anzugeben.
- Wenn Sie den aktuellen Filter löschen wollen, geben Sie im Dialogfeld **Angezeigte Werte filtern** für den Suchbegriff einen leeren Wert an.

---

## Beobachtungen

Verwenden Sie auf der Registerkarte **Felder** die Einstellungen für **Beobachtungen**, um die Felder anzugeben, durch die die Beobachtungen definiert werden.

**Anmerkung:** Wenn das aktive Dataset eine Datumsspezifikation aufweist, werden die Beobachtungen durch die Datumsspezifikation definiert und können in der Prozedur der temporalen kausalen Modellierung nicht geändert werden. Datumsspezifikationen werden über das Dialogfeld **Datum definieren** oder den Befehl DATE erstellt.

### Beobachtungen, die durch Datum/Zeit definiert werden

Sie können angeben, dass die Beobachtungen durch ein Feld mit einem Datums-, Zeit- oder Datums-/Zeit-Format oder durch ein Zeichenfolgefild, das eine Datums-/Zeitangabe darstellt, definiert werden sollen. Zeichenfolgefelder können ein Datum im Format JJJJ-MM-TT, eine Zeit im Format HH:MM:SS oder eine Datums-/Zeitangabe im Format JJJJ-MM-TT darstellen. Führende Nullen können in der Zeichenfolgedarstellung übergangen werden. So wäre z. B. die Zeichenfolge 2014-9-01 äquivalent zu 2014-09-01.

Wählen Sie zusätzlich zu dem Feld, das die Beobachtungen definiert, das geeignete Zeitintervall aus, das die Beobachtungen beschreibt. Je nach angegebenem Zeitintervall können Sie auch andere Einstellungen angeben, beispielsweise das Intervall zwischen Beobachtungen (Inkrement) oder die Anzahl der Tage pro Woche. Folgende Aspekte gelten für das Zeitintervall:

- Verwenden Sie den Wert **Unregelmäßig**, wenn die Beobachtungen zeitlich unregelmäßig sind, z. B. in Bezug auf die Zeit, zu der ein Verkaufsauftrag verarbeitet wird. Wenn Sie die Option **Unregelmäßig** auswählen, müssen Sie in den Einstellungen für **Zeitintervall** auf der Registerkarte **Datenspezifikationen** das für die Analyse verwendete Zeitintervall angeben.
- Wenn die Beobachtungen ein Datum und eine Zeit darstellen und das Zeitintervall in Stunden, Minuten oder Sekunden angegeben ist, verwenden Sie **Stunden pro Tag**, **Minuten pro Tag** oder **Sekunden pro Tag**. Wenn die Beobachtungen eine Zeitdauer ohne Bezug zu einem Datum darstellen und das Zeitintervall in Stunden, Minuten oder Sekunden angegeben ist, verwenden Sie **Stunden (nicht periodisch)**, **Minuten (nicht periodisch)** oder **Sekunden (nicht periodisch)**.
- Je nach dem ausgewählten Zeitintervall kann die Prozedur fehlende Beobachtungen erkennen. Die Erkennung fehlender Beobachtungen ist erforderlich, da die Prozedur voraussetzt, dass alle Beobachtungen gleichmäßige zeitliche Intervalle aufweisen und keine Beobachtungen fehlen. Wenn das Zeitintervall z. B. **Tage** ist und unmittelbar auf das Datum 2014-10-27 das Datum 2014-10-29 folgt, fehlt die Beobachtung für das Datum 2014-10-28. Für fehlende Beobachtungen werden Werte imputiert. Einstellungen für die Handhabung fehlender Werte können auf der Registerkarte **Datenspezifikationen** vorgenommen werden.
- Aufgrund des angegebenen Zeitintervalls kann die Prozedur mehrere Beobachtungen in demselben Zeitintervall erkennen, die aggregiert werden müssen, und Beobachtungen an einer Intervallgrenze ausrichten, z. B. dem Monatsersten, um sicherzustellen, dass die Beobachtungen gleichmäßige Intervalle aufweisen. Wenn das Zeitintervall z. B. **Monate** ist, werden mehrere Daten desselben Monats aggregiert. Diese Art der Aggregation wird als *Gruppierung* bezeichnet. Standardmäßig werden Beobachtungen summiert, wenn sie gruppiert werden. Sie können über die Einstellungen für **Aggregation und Verteilung** auf der Registerkarte **Datenspezifikationen** eine andere Methode für die Gruppierung angeben, z. B. den Mittelwert der Beobachtungen.
- Bei einigen Zeitintervallen können über die zusätzlichen Einstellungen Unterbrechungen in den normalen gleichmäßigen Intervallen definiert werden. Wenn das Zeitintervall z. B. **Tage** ist, jedoch nur Werkzeuge berücksichtigt werden sollen, können Sie angeben, dass eine Woche fünf Tage hat und am Montag beginnt.

### Beobachtungen, die durch Perioden oder zyklische Perioden definiert werden

Beobachtungen können durch ein Feld oder mehrere Felder für ganze Zahlen definiert werden, die Perioden oder sich wiederholende Periodenzyklen bis zu einer beliebigen Anzahl Zyklusebenen darstellen. Mit dieser Struktur können Sie Zeitreihen von Beobachtungen beschreiben, die

keinem der Standardzeitintervalle entsprechen. So kann beispielsweise ein Geschäftsjahr mit nur zehn Monaten durch ein Zyklusfeld, das Jahre darstellt, und ein Periodenfeld, das Monate darstellt, beschrieben werden, wobei die Länge eines Zyklus den Wert 10 aufweist.

Felder, die zyklische Perioden angeben, definieren eine Hierarchie von Periodenebenen, von denen die niedrigste durch das Feld **Periode** definiert wird. Die nächsthöhere Ebene wird durch ein Zyklusfeld mit der Ebene 1 angegeben, gefolgt von einem Zyklusfeld mit der Ebene 2 usw. Feldwerte der einzelnen Ebenen mit Ausnahme der höchsten müssen bezogen auf die nächsthöhere Ebene periodisch sein. Werte für die höchste Ebene können nicht periodisch sein. Im Fall des zehnmonatigen Geschäftsjahrs z. B. sind Monate innerhalb von Jahren periodisch, Jahre dagegen sind nicht periodisch.

- Die Länge eines Zyklus auf einer bestimmten Ebene entspricht der Periodizität der nächstniedrigeren Ebene. Im Beispiel des Geschäftsjahrs gibt es nur eine einzige Zyklusebene und die Zykluslänge weist den Wert 10 auf, da die nächstniedrigere Ebene die Monate darstellt und das angegebene Geschäftsjahr zehn Monate umfasst.
- Geben Sie für alle Periodenfelder, die nicht mit 1 beginnen, den Startwert an. Diese Einstellung ist erforderlich, um fehlende Werte erkennen zu können. Wenn z. B. ein Periodenfeld mit 2 beginnt, obwohl der Startwert 1 angegeben wurde, geht die Prozedur davon aus, dass für die erste Periode in jedem Zyklus dieses Felds ein Wert fehlt.

### Beobachtungen, die durch die Datensatzreihenfolge definiert werden

Bei spaltenbasierten Daten können Sie angeben, dass die Beobachtungen durch die Datensatzreihenfolge definiert werden, sodass der erste Datensatz die erste Beobachtung darstellt, der zweite Datensatz die zweite Beobachtung usw. Daraufhin wird davon ausgegangen, dass die Datensätze Beobachtungen mit gleichmäßigen Zeitintervallen darstellen.

---

## Zeitintervall für die Analyse

Das für die Analyse verwendete Zeitintervall kann vom Zeitintervall der Beobachtungen abweichen. Wenn das Zeitintervall der Beobachtungen z. B. **Tage** ist, können Sie als Zeitintervall für die Analyse **Monate** auswählen. Die Daten werden dann vor der Erstellung des Modells von täglichen auf monatliche Daten aggregiert. Sie können auch festlegen, dass die Daten von einem längeren auf ein kürzeres Zeitintervall verteilt werden. Wenn z. B. die Beobachtungen vierteljährlich durchgeführt werden, können Sie die Daten von vierteljährlichen auf monatliche Daten verteilen.

Die Auswahlmöglichkeiten für das Zeitintervall, in dem die Analyse erfolgt, hängen davon ab, wie die Beobachtungen definiert werden, sowie vom Zeitintervall der Beobachtungen. Insbesondere, wenn die Beobachtungen durch zyklische Perioden definiert werden oder wenn für das aktive Dataset eine Datumspezifikation definiert ist, wird nur Aggregation unterstützt. In diesem Fall muss das Zeitintervall für die Analyse größer-gleich dem Zeitintervall für die Beobachtungen sein.

Das Zeitintervall für die Analyse wird über die Einstellungen für **Zeitintervall** auf der Registerkarte **Datenspezifikationen** angegeben. Die Methode, nach der die Daten aggregiert oder verteilt werden, wird über die Einstellungen für **Aggregation und Verteilung** auf der Registerkarte **Datenspezifikationen** angegeben.

---

## Aggregation und Verteilung

### Aggregationsfunktionen

Wenn das für die Analyse verwendete Zeitintervall länger ist als das Zeitintervall der Beobachtungen, werden die Eingabedaten aggregiert. Eine Aggregation erfolgt z. B. dann, wenn das Zeitintervall der Beobachtungen **Tage** und das Zeitintervall für die Analyse **Monate** ist. Folgende Aggregationsfunktionen sind verfügbar: "Mittelwert", "Summe", "Modalwert", "Min" oder "Max".

### Verteilungsfunktionen

Wenn das für die Analyse verwendete Zeitintervall kürzer ist als das Zeitintervall der Beobachtungen, werden die Eingabedaten verteilt. Eine Verteilung erfolgt z. B., wenn das Zeitintervall

der Beobachtungen **Quartale** und das Zeitintervall für die Analyse **Monate** ist. Folgende Verteilungsfunktionen sind verfügbar: "Mittelwert" oder "Summe".

### **Gruppierungsfunktionen**

Gruppierungen finden Anwendung, wenn Beobachtungen durch Datum/Zeit definiert werden und mehrere Beobachtungen in demselben Zeitintervall auftreten. Wenn das Zeitintervall der Beobachtungen z. B. **Monate** ist, werden mehrere Daten desselben Monats gruppiert und dem Monat ihres Auftretens zugeordnet. Folgende Gruppierungsfunktionen sind verfügbar: "Mittelwert", "Summe", "Modalwert", "Min" oder "Max". Eine Gruppierung erfolgt immer, wenn die Beobachtungen durch Datum/Zeit definiert werden und für das Zeitintervall der Beobachtungen **Unregelmäßig** angegeben ist.

**Anmerkung:** Obwohl die Gruppierung eine Form der Aggregation darstellt, wird sie vor der Behandlung fehlender Werte durchgeführt, während die formale Aggregation erst nach der Behandlung aller fehlenden Werte erfolgt. Wenn für das Zeitintervall der Beobachtungen **Unregelmäßig** angegeben ist, wird die Aggregation nur mit der Gruppierungsfunktion durchgeführt.

### **Tagübergreifende Beobachtungen auf vorherigen Tag aggregieren**

Gibt an, ob Beobachtungen, deren Uhrzeiten eine Tagesgrenze überschreiten, auf die Werte des Vortags aggregiert werden sollen. Wenn es z. B. um stündliche Beobachtungen bezogen auf einen Achtstundentag geht, der um 20.00 Uhr beginnt, wird mit dieser Einstellung festgelegt, ob Beobachtungen zwischen 00.00 Uhr und 04.00 Uhr in die aggregierten Ergebnisse für den vorherigen Tag eingeschlossen werden sollen. Diese Einstellung gilt nur, wenn das Zeitintervall der Beobachtungen **Stunden pro Tag**, **Minuten pro Tag** oder **Sekunden pro Tag** und das Zeitintervall für die Analyse **Tage** ist.

### **Benutzerdefinierte Einstellungen für angegebene Felder**

Sie können Aggregations-, Verteilungs- und Gruppierungsfunktionen für die einzelnen Felder angeben. Diese Einstellungen setzen die Standardeinstellungen für die Aggregations-, Verteilungs- und Gruppierungsfunktionen außer Kraft.

---

## **Fehlende Werte**

Fehlende Werte in den Eingabedaten werden durch imputierte Werte ersetzt. Folgende Ersetzungsmethoden sind verfügbar:

### **Lineare Interpolation**

Ersetzt fehlende Werte unter Anwendung einer linearen Interpolation. Für die Interpolation werden der letzte gültige Wert vor dem fehlenden Wert und der erste gültige Wert nach dem fehlenden Wert verwendet. Wenn die erste oder letzte Beobachtung in der Zeitreihe einen fehlenden Wert aufweist, werden die beiden nächstgelegenen nicht fehlenden Werte am Anfang oder am Ende der Zeitreihe verwendet.

### **Mittelwert der Zeitreihe**

Ersetzt fehlende Werte durch den Mittelwert der gesamten Zeitreihe.

### **Mittelwert benachbarter Punkte**

Ersetzt fehlende Werte durch den Mittelwert gültiger Umgebungswerte. Der Bereich benachbarter Punkte ist die Anzahl gültiger Werte über und unter dem fehlenden Wert, die für die Berechnung des Mittelwerts verwendet werden.

### **Median benachbarter Punkte**

Ersetzt fehlende Werte durch den Median gültiger Umgebungswerte. Der Bereich benachbarter Punkte ist die Anzahl gültiger Werte über und unter dem fehlenden Wert, die für die Berechnung des Medians verwendet werden.

### **Linearer Trend**

Diese Option verwendet alle nicht fehlenden Beobachtungen in der Zeitreihe, um ein einfaches lineares Regressionsmodell anzupassen, das anschließend zum Imputieren der fehlenden Werte verwendet wird.

Sonstige Einstellungen:

#### **Maximaler Prozentsatz fehlender Werte (%)**

Gibt den maximalen Prozentsatz fehlender Werte an, der für Zeitreihen zulässig ist. Zeitreihen, die mehr fehlende Werte aufweisen als das angegebene Maximum, werden von der Analyse ausgeschlossen.

#### **Benutzerdefiniert fehlende Werte**

Diese Option gibt an, ob benutzerdefiniert fehlende Werte wie gültige Daten behandelt und daher in die Zeitreihe eingeschlossen werden. Standardmäßig werden benutzerdefiniert fehlende Werte ausgeschlossen und wie systemdefiniert fehlende Werte behandelt, also anschließend imputiert.

---

## **Allgemeine Datenoptionen**

#### **Maximale Anzahl unterschiedlicher Werte pro Dimensionsfeld**

Diese Einstellung gilt für mehrdimensionale Daten und gibt die maximale Anzahl unterschiedlicher Werte an, die für ein Dimensionsfeld zulässig sind. Standardmäßig ist dieser Grenzwert auf 10.000 gesetzt, er kann jedoch auf eine beliebige Anzahl erhöht werden.

---

## **Allgemeine Erstellungsoptionen**

#### **Konfidenzintervallbreite (%)**

Diese Einstellung steuert die Konfidenzintervalle sowohl für Vorhersagen als auch für Modellparameter. Sie können einen beliebigen positiven Wert kleiner als 100 angeben. Standardmäßig wird ein Konfidenzintervall von 95 % verwendet.

#### **Maximale Anzahl an Eingaben für jedes Ziel**

Diese Einstellung gibt die maximale Anzahl zulässiger Eingaben je Ziel im Modell an. Sie können eine Ganzzahl im Bereich 1 - 20 angeben. Das Modell eines Ziels beinhaltet stets Lagwerte des Ziels selbst; wird dieser Wert auf 1 gesetzt, bedeutet dies, dass die einzig zulässige Eingabe das Ziel selbst ist.

#### **Modelltoleranz**

Diese Einstellung steuert den iterativen Prozess, der verwendet wird, um die optimalen Eingaben für die einzelnen Ziele zu bestimmen. Sie können einen beliebigen Wert größer als null angeben. Der Standardwert ist 0,001.

#### **Schwellenwert für Ausreißer (%)**

Eine Beobachtung wird als Ausreißer markiert, wenn die aufgrund des Modells berechnete Wahrscheinlichkeit, dass es sich hierbei um einen Ausreißer handelt, diesen Schwellenwert überschreitet. Sie können einen Wert im Bereich 50 bis 100 angeben.

#### **Die Anzahl der Lags für jede Eingabe**

Diese Einstellung gibt die Anzahl der Lagbegriffe für jede Eingabe in jedem Ziel des Modells an. Standardmäßig wird die Anzahl der Lagbegriffe automatisch durch das für die Analyse verwendete Zeitintervall festgelegt. Wenn das Zeitintervall z. B. **Monate** (mit einem Inkrement von einem Monat) ist, beträgt die Anzahl der Lags 12. Optional können Sie die Anzahl der Lags explizit angeben. Der angegebene Wert muss eine Ganzzahl im Bereich 1 - 20 sein.

---

## **Anzuzeigende Zeitreihe**

Die folgenden Optionen geben die Zeitreihen (Ziele oder Eingaben) an, für die eine Ausgabe angezeigt wird. Der Inhalt der Ausgabe für die angegebenen Zeitreihen wird über die Einstellungen für **Ausgabeoptionen** festgelegt.

#### **Mit am besten angepassten Modellen verbundene Ziele anzeigen**

Standardmäßig wird eine Ausgabe für die Ziele angezeigt, die den zehn laut R-Quadrat-Wert am besten angepassten Modellen zugeordnet sind. Sie können auch eine andere feste Anzahl von am

besten angepassten Modellen oder einen Prozentsatz für am besten angepasste Modelle angeben. Darüber hinaus können Sie folgende Maße für die Anpassungsgüte auswählen:

#### **R-Quadrat**

Ein Maß für die Anpassungsgüte eines linearen Modells. Wird auch als Bestimmtheitskoeffizient bezeichnet. Es gibt den Anteil der Variation der Zielvariablen an, der durch das Modell erklärt wird. Er liegt zwischen 0 und 1. Kleine Werte zeigen an, dass das Modell nicht gut zu den Daten passt.

#### **Wurzel der mittleren Fehlerquadratsumme (Prozent)**

Maß für die Differenz zwischen vom Modell vorhergesagten Werten und den beobachteten Werten der Zeitreihe. Es ist unabhängig von den verwendeten Maßeinheiten und kann daher verwendet werden, um Zeitreihen mit unterschiedlichen Einheiten zu vergleichen.

#### **Quadratwurzel des mittleren quadratischen Fehlers**

Die Quadratwurzel des mittleren quadratischen Fehlers. Ein Maß dafür, wie stark eine abhängige Zeitreihe von ihrem durch das Modell vorhergesagten Niveau abweicht, und zwar ausgedrückt in derselben Maßeinheit wie die abhängige Zeitreihe.

**BIC** Bayes-Informationskriterium. Ein Maß für die Auswahl und den Vergleich von Modellen, das auf  $-2$  Log-Likelihood beruht. Kleinere Werte stehen für bessere Modelle. BIC "penalisiert" überparametrisierte Modelle ebenfalls (beispielsweise Modelle mit einer großen Anzahl Eingaben). allerdings strikter als AIC.

**AIC** Akaike-Informationskriterium. Ein Maß für die Auswahl und den Vergleich von Modellen, das auf  $-2$  Log-Likelihood beruht. Kleinere Werte stehen für bessere Modelle. AIC "penalisiert" überparametrisierte Modelle (beispielsweise Modelle mit einer großen Anzahl Eingaben.)

### **Einzelne Zeitreihe angeben**

Sie können einzelne Zeitreihen angeben, für die eine Ausgabe erstellt werden soll.

- Bei spaltenbasierten Daten geben Sie die Felder mit den Zeitreihen an, die Sie benötigen. Die Reihenfolge der angegebenen Felder legt die Reihenfolge fest, in der diese in der Ausgabe angezeigt werden.
- Bei mehrdimensionalen Daten geben Sie eine bestimmte Zeitreihe an, indem Sie im Raster für das Metrikfeld, das die Zeitreihe enthält, eine Eingabe hinzufügen. Anschließend geben Sie die Werte der Dimensionsfelder an, die die Zeitreihe definieren.
  - Sie können die Werte für die einzelnen Dimensionsfelder direkt in das Raster eingeben oder sie aus der Liste der verfügbaren Dimensionswerte auswählen. Wenn Sie Werte aus der Liste der verfügbaren Dimensionswerte auswählen wollen, klicken Sie auf die Schaltfläche mit Auslassungspunkten in der Zelle für die gewünschte Dimension. Daraufhin wird das Unterdialogfeld **Dimensionswerte auswählen** geöffnet.
  - Sie können die Liste der Dimensionswerte durchsuchen, indem Sie im Unterdialogfeld **Dimensionswerte auswählen** auf das Fernglassymbol klicken und einen Suchbegriff angeben. Leerzeichen werden als Teil des Suchbegriffs behandelt. Sterne (\*) im Suchbegriff zeigen keine Platzhalterzeichen an.
  - Die Reihenfolge der Zeitreihen im Raster legt die Reihenfolge fest, in der diese in der Ausgabe angezeigt werden.

Sowohl für spaltenbasierte als auch für mehrdimensionale Daten ist die Ausgabe auf 30 Zeitreihen begrenzt. Dieser Höchstwert umfasst einzelne von Ihnen angegebene Zeitreihen (Eingaben oder Ziele) und Ziele, die mit am besten angepassten Modellen verbunden sind. Einzeln angegebene Zeitreihen haben Vorrang vor Zielen, die mit am besten angepassten Modellen verbunden sind.

---

## Ausgabeoptionen

Diese Optionen legen den Inhalt der Ausgabe fest. Die Optionen in der Gruppe **Ausgabe für Inhalte** generieren eine Ausgabe für die Ziele, die in den Einstellungen für **Anzuzeigende Zeitreihe** mit den am besten angepassten Modellen verbunden sind. Die Optionen in der Gruppe **Ausgabe für Inhalte** generieren eine Ausgabe für die einzelnen Zeitreihen, die in den Einstellungen für **Anzuzeigende Zeitreihe** angegeben sind.

### Gesamtmodellsystem

Zeigt eine grafische Darstellung der kausalen Beziehungen zwischen Zeitreihen im Modellsystem an. Tabellen sowohl für Statistiken zur Modellanpassungsgüte als auch und zu Ausreißern für die angezeigten Ziele sind in das Ausgabeelement eingeschlossen. Wenn diese Option in der Gruppe **Ausgabe für Zeitreihen** ausgewählt wird, wird ein separates Ausgabeelement für jede einzelne Zeitreihe generiert, die in den Einstellungen für **Anzuzeigende Zeitreihe** angegeben ist.

Kausale Beziehungen zwischen Zeitreihen haben ein zugeordnetes Signifikanzniveau, wobei ein kleineres Signifikanzniveau eine signifikantere Verbindung anzeigt. Sie können auswählen, dass Beziehungen mit einem Signifikanzniveau oberhalb eines bestimmten Wertes ausgeblendet werden.

### Statistik zur Anpassungsgüte des Modells und Ausreißer

Tabellen für Statistiken zur Modellanpassungsgüte und zu Ausreißern für die Zielzeitreihen, die zur Anzeige ausgewählt sind. Diese Tabellen enthalten dieselben Informationen wie die Tabellen zur Visualisierung des Gesamtmodellsystems. Sie unterstützen alle Standardfunktionen für das Drehen und Bearbeiten von Tabellen.

### Modelleffekte und Modellparameter

Tabellen zu Tests von Modelleffekten und zu Modellparametern für die Zielzeitreihen, die zur Anzeige ausgewählt sind. Tests zu Modelleffekten beinhalten die F-Statistik und zugehörige Signifikanzwerte für jede in das Modell eingeschlossene Eingabe.

### Wirkungsdiagramm

Zeigt eine grafische Darstellung der kausalen Beziehungen zwischen einer relevanten Zeitreihe und anderen Zeitreihen, auf die sie sich auswirkt oder die sich auf sie auswirken. Zeitreihen, die sich auf die relevante Zeitreihe auswirken, werden als *Ursachen* bezeichnet. Wenn Sie **Effekte** auswählen, wird ein für die Anzeige der Effekte initialisiertes Wirkungsdiagramm generiert. Wenn Sie **Ursachen** auswählen, wird ein für die Anzeige der Ursachen initialisiertes Wirkungsdiagramm generiert. Wenn Sie **Sowohl Ursachen als auch Effekte** auswählen, werden zwei separate Wirkungsdiagramme generiert, eins wird für die Ursachen und eines wird für die Effekte initialisiert. Im Ausgabeelement, das das Wirkungsdiagramm anzeigt, können Sie interaktiv zwischen Ursachen und Effekten wechseln.

Sie können die Anzahl der anzuzeigenden Ebenen von Ursachen oder Effekten angeben, wobei die erste Ebene nur die relevante Zeitreihe ist. Jede zusätzliche Ebene zeigt weitere indirekte Ursachen oder Effekte bezogen auf die relevante Zeitreihe an. So besteht z. B. die dritte Ebene in der Anzeige der Effekte aus den Zeitreihen, die Zeitreihen in der zweiten Ebene als direkte Eingabe enthalten. Auf Zeitreihen der dritten Ebene wirkt sich die relevante Zeitreihe dann indirekt aus, da die relevante Zeitreihe eine direkte Eingabe für die Zeitreihen in der zweiten Ebene ist.

### Zeitreihendiagramm

Diagramme beobachteter und vorhergesagter Werte für die Zielzeitreihen, die zur Anzeige ausgewählt sind. Wenn Vorhersagen angefordert werden, zeigt das Diagramm auch die vorhergesagten Werte und die Konfidenzintervalle für die Vorhersagen an.

### Residuumdiagramme

Diagramme der Modellresiduen für die Zielzeitreihen, die zur Anzeige ausgewählt sind.

### Top Eingaben

Diagramme der einzelnen angezeigten Ziele über einen bestimmten Zeitraum sowie die drei Top

Eingaben für das Ziel. Die Top Eingaben sind die Eingaben mit dem niedrigsten Signifikanzwert. Um unterschiedliche Maßstäbe für die Eingaben und Ziele veranschaulichen zu können, stellt die Y-Achse den Z-Score für die einzelnen Zeitreihen dar.

### **Vorhersagetabelle**

Tabellen mit vorhergesagten Werten und Konfidenzintervallen dieser Vorhersagen für die Zielzeitreihen, die zur Anzeige ausgewählt sind.

### **Ursachenanalyse für Ausreißer**

Legt fest, welche Zeitreihen die wahrscheinlichsten Ursachen für die einzelnen Ausreißer in einer relevanten Zeitreihe sind. Eine Ursachenanalyse für Ausreißer wird für jede Zielzeitreihe durchgeführt, die in der Liste der einzelnen Zeitreihen in den Einstellungen für **Anzuzeigende Zeitreihe** enthalten sind.

### **Ausgabe**

#### **Interaktive Ausreißertabelle und Diagramm**

Tabelle und Diagramm mit Ausreißern und Ursachen für diese Ausreißer für jede der relevanten Zeitreihen. Die Tabelle enthält eine Zeile je Ausreißer. Bei dem Diagramm handelt es sich um ein Wirkungsdiagramm. Bei Auswahl einer Tabellenzeile wird im Wirkungsdiagramm der Pfad von der relevanten Zeitreihe bis zu der Zeitreihe hervorgehoben, die die wahrscheinlichste Ursache für den zugehörigen Ausreißer ist.

#### **Pivot-Tabelle mit Ausreißern**

Tabelle mit Ausreißern und Ursachen für diese Ausreißer für jede der relevanten Zeitreihen. Diese Tabelle enthält dieselben Informationen wie die Tabelle in der interaktiven Anzeige. Sie unterstützt alle Standardfunktionen für das Drehen und Bearbeiten von Tabellen.

### **Kausale Ebenen**

Sie können die Anzahl der Ebenen angeben, die in die Suche nach den Ursachen eingeschlossen werden sollen. Das hier verwendete Konzept der Ebenen entspricht dem, das auch für Wirkungsdiagramme beschrieben wurde.

### **Modellübergreifende Anpassungsgüte des Modells**

Histogramm der Anpassungsgüte für alle Modelle und für ausgewählte Statistiken zur Anpassungsgüte. Folgende Statistiken zur Anpassungsgüte sind verfügbar:

#### **R-Quadrat**

Ein Maß für die Anpassungsgüte eines linearen Modells. Wird auch als Bestimmtheitskoeffizient bezeichnet. Es gibt den Anteil der Variation der Zielvariablen an, der durch das Modell erklärt wird. Er liegt zwischen 0 und 1. Kleine Werte zeigen an, dass das Modell nicht gut zu den Daten passt.

#### **Wurzel der mittleren Fehlerquadratsumme (Prozent)**

Maß für die Differenz zwischen vom Modell vorhergesagten Werten und den beobachteten Werten der Zeitreihe. Es ist unabhängig von den verwendeten Maßeinheiten und kann daher verwendet werden, um Zeitreihen mit unterschiedlichen Einheiten zu vergleichen.

#### **Quadratwurzel des mittleren quadratischen Fehlers**

Die Quadratwurzel des mittleren quadratischen Fehlers. Ein Maß dafür, wie stark eine abhängige Zeitreihe von ihrem durch das Modell vorhergesagten Niveau abweicht, und zwar ausgedrückt in derselben Maßeinheit wie die abhängige Zeitreihe.

**BIC** Bayes-Informationskriterium. Ein Maß für die Auswahl und den Vergleich von Modellen, das auf  $-2 \log$ -Likelihood beruht. Kleinere Werte stehen für bessere Modelle. BIC "penalisiert" überparametrisierte Modelle ebenfalls (beispielsweise Modelle mit einer großen Anzahl Eingaben). allerdings strikter als AIC.

**AIC** Akaike-Informationskriterium. Ein Maß für die Auswahl und den Vergleich von Model-

len, das auf  $-2$  Log-Likelihood beruht. Kleinere Werte stehen für bessere Modelle. AIC "penalisiert" überparametrisierte Modelle (beispielsweise Modelle mit einer großen Anzahl Eingaben.)

### Ausreißer über einen Zeitraum

Balkendiagramm zur Anzahl der Ausreißer in allen Zielen für jedes Zeitintervall in der Schätzperiode.

### Zeitreihentransformationen

Tabelle aller Transformationen, die auf die Zeitreihe im Modellsystem angewandt wurden. Zu den möglichen Transformationen gehören die Imputation fehlender Werte, die Aggregation und die Verteilung.

---

## Schätzperiode

Standardmäßig beginnt die Schätzperiode mit dem Zeitpunkt der frühesten Beobachtung und endet mit dem Zeitpunkt der spätesten Beobachtung bezogen auf alle Zeitreihen.

### Nach Start- und Endzeit

Sie können für die Schätzperiode entweder sowohl die Start- als auch die Endzeit angeben oder Sie können entweder die Start- oder die Endzeit angeben. Wenn Sie die Startzeit oder die Endzeit der Schätzperiode weglassen, wird der Standardwert verwendet.

- Wenn die Beobachtungen durch ein Datums-/Zeitfeld definiert werden, geben Sie Werte für den Anfang und für das Ende in demselben Format an, das für das Datums-/Zeitfeld verwendet wurde.
- Geben Sie für Beobachtungen, die durch zyklische Perioden definiert werden, einen Wert für jedes der Felder für die zyklische Periode an. Jedes Feld wird in einer separaten Spalte angezeigt.
- Gilt für das aktive Dataset eine Datumsspezifikation, müssen Sie einen Wert für jede Komponente (beispielsweise einen Monat) der Datumsspezifikation angeben. Jede Komponente wird in einer separaten Spalte angezeigt.
- Wenn die Beobachtungen nach Datensatzreihenfolge definiert werden, werden Anfang und Ende der Schätzperiode über die Zeilennummer (wie im Dateneditor angegeben) des betreffenden Falls definiert.

### Nach spätesten oder frühesten Zeitintervallen

Definiert die Schätzperiode mit einem optionalen Offset als eine bestimmte Anzahl Zeitintervalle, die mit dem frühesten Zeitintervall in den Daten beginnen oder mit dem spätesten Zeitintervall in den Daten enden. In diesem Kontext bezieht sich das Zeitintervall auf das Zeitintervall der Analyse. Beispiel: Die Beobachtungen erfolgen monatlich, doch das Zeitintervall für die Analyse ist **Quartale**. Wenn Sie **Aktuellste** und den Wert 24 für **Anzahl der Zeitintervalle** angeben, bedeutet dies die letzten 24 Quartale.

Optional können Sie eine bestimmte Anzahl Zeitintervalle ausschließen. Wenn Sie z. B. die aktuellsten 24 Zeitintervalle und für die auszuschließende Anzahl den Wert 1 angeben, bedeutet dies, dass die Schätzperiode aus den 24 Intervallen besteht, die dem letzten Intervall vorausgehen.

---

## Vorhersage

Die Option **Datensätze auf die Zukunft ausdehnen** legt die Anzahl der Zeitintervalle fest, für die über das Ende der Schätzperiode hinaus eine Vorhersage erstellt werden soll. In diesem Fall entspricht das Zeitintervall dem Zeitintervall für die Analyse, das auf der Registerkarte **Datenspezifikationen** angegeben ist. Bei Anforderung von Vorhersagen werden für alle Eingabezeitreihen, die nicht gleichzeitig auch Ziele sind, automatisch autoregressive Modelle erstellt. Diese Modelle werden dann verwendet, um für die Eingabezeitreihen in der Vorhersageperiode Werte zu generieren.

---

## Speichern

### Zieloptionen

Sie können sowohl Transformationen der Daten (beispielsweise Aggregation oder Imputation fehlender Werte) als auch neue Variablen (angegeben in den Einstellungen für **Ziele speichern**) in einer IBM SPSS Statistics-Datendatei oder in einem neuen Dataset in der aktuellen Sitzung speichern. Datums-/Zeitwerte in den gespeicherten Daten werden am Anfang jedes Zeitintervalls ausgerichtet, z. B. am Monatsersten, und stellen das Zeitintervall für die Analyse des Modellsystems dar. Sie können alle neuen Variablen nur dann im aktiven Dataset speichern, wenn die Beobachtungen durch eine Datumsspezifikation oder die Datensatzreihenfolge definiert werden und die Daten nicht aggregiert sind.

### Ziele speichern

Sie können Modellvorhersagen, Konfidenzintervalle und Residuen als neue Variablen speichern. Jedes zum Speichern angegebene Ziel generiert eine eigene Gruppe neuer Variablen und jede neue Variable enthält Werte für Schätz- und Vorhersageperioden. Für vorhergesagte Werte, Konfidenzintervalle und Restauschen können Sie den Stammmamen angeben, der als Präfix für die neuen Variablen verwendet werden soll. Der vollständige Variablenname ist die Verkettung des Stammmamens und des Namens des Felds, das die Zielzeitreihe enthält. Der Stammmame muss den Regeln für gültige Variablennamen entsprechen. Der Variablenname wird gegebenenfalls erweitert, um das Auftreten von Namenskonflikten zu vermeiden.

### Gibt Fälle mit Vorhersagen an

Erstellt eine Variable, die angibt, ob ein Datensatz Vorhersagedaten enthält. Sie können den Variablennamen angeben. Der Standardwert ist *ForecastIndicator*.

### Zu speichernde Ziele

Gibt an, ob für alle Zielzeitreihen im Modellsystem neue Variablen erstellt werden oder nur für die Zielzeitreihen, die in den Einstellungen für **Anzuzeigende Zeitreihe** angegeben sind.

### Modellsystem exportieren

Speichert das Modellsystem in einem komprimierten Dateiarchiv (ZIP-Datei). Die Modellsystemdatei kann von der Prozedur "Vorhersage temporaler kausaler Modelle" verwendet werden, um aktualisierte Vorhersagen abzurufen oder eine der verfügbaren Ausgaben zu generieren. Sie kann auch von der Prozedur "Temporale kausale Modellszenarios" verwendet werden, um eine Szenarioanalyse durchzuführen.

---

## Interaktive Ausgabe

Die Ausgabe aus temporaler kausaler Modellierung schließt eine Reihe interaktiver Ausgabeobjekte ein. Interaktive Funktionen sind durch Aktivieren (Doppelklicken) des Objekts im Ausgabeviewer verfügbar.

### Gesamtmodellsystem

Zeigt die kausalen Beziehungen zwischen Zeitreihen im Modellsystem an. Alle Linien, die ein bestimmtes Ziel mit seinen Eingaben verbinden, haben dieselbe Farbe. Die Stärke der Linie zeigt die Signifikanz der Kausalverbindung an, wobei dickere Linien eine signifikantere Verbindung darstellen. Eingaben, die nicht auch Ziele sind, werden durch ein schwarzes Quadrat gekennzeichnet.

- Sie können Beziehungen für beste Modelle, eine angegebene Zeitreihe, alle Zeitreihen oder Modelle ohne Eingaben anzeigen. Die besten Modelle sind die Modelle, die die Kriterien erfüllen, die für die am besten angepassten Modelle in den Einstellungen **Anzuzeigende Zeitreihe** festgelegt sind.
- Sie können Wirkungsdiagramme für mindestens eine Zeitreihe generieren, indem Sie die Zeitreihennamen im Diagramm auswählen, mit der rechten Maustaste klicken und dann **Wirkungsdiagramm erstellen** im Kontextmenü auswählen.

- Sie können auch kausale Beziehungen ausblenden, die ein Signifikanzniveau über einem festgelegten Wert haben. Kleinere Signifikanzniveaus weisen auf eine signifikantere kausale Beziehung hin.
- Sie können Beziehungen für eine bestimmte Zeitreihe anzeigen, indem Sie den Zeitreihennamen im Diagramm auswählen, mit der rechten Maustaste klicken und dann **Beziehungen für Zeitreihe hervorheben** im Kontextmenü auswählen.

### Wirkungsdiagramm

Zeigt eine grafische Darstellung der kausalen Beziehungen zwischen einer relevanten Zeitreihe und anderen Zeitreihen, auf die sie sich auswirkt oder die sich auf sie auswirken. Zeitreihen, die sich auf die relevante Zeitreihe auswirken, werden als *Ursachen* bezeichnet.

- Sie können die relevante Zeitreihe ändern, indem Sie den Namen der gewünschten Zeitreihe angeben. Durch Doppelklicken auf einen beliebigen Knoten im Wirkungsdiagramm ändern Sie die relevante Zeitreihe in die diesem Knoten zugeordnete Zeitreihe.
- Sie können in der Anzeige zwischen Ursachen und Wirkungen wechseln und die Anzahl der Ebenen von Ursachen oder Wirkungen, die angezeigt werden sollen, ändern.
- Durch Einmalklicken auf einen beliebigen Knoten wird ein detailliertes Sequenzdiagramm für die Zeitreihe geöffnet, die dem Knoten zugeordnet ist.

### Analyse der Ursachen für Ausreißer

Legt fest, welche Zeitreihen die wahrscheinlichsten Ursachen für die einzelnen Ausreißer in einer relevanten Zeitreihe sind.

- Sie können die Ursache für Ausreißer anzeigen, indem Sie die Zeile für den Ausreißer in der Tabelle **Ausreißer** auswählen. Sie können die Ursache auch anzeigen, indem Sie auf das Symbol für den Ausreißer im Sequenzdiagramm klicken.
- Durch Einmalklicken auf einen beliebigen Knoten wird ein detailliertes Sequenzdiagramm für die Zeitreihe geöffnet, die dem Knoten zugeordnet ist.

### Gesamtmodellqualität

Histogramm der Modellanpassungsgüte für alle Modelle für eine bestimmte Statistik zur Anpassungsgüte. Durch Klicken auf einen Balken im Balkendiagramm wird das Punktdiagramm gefiltert, sodass es nur die Modelle anzeigt, die mit dem ausgewählten Balken zusammenhängen. Sie können das Modell für eine bestimmte Zielzeitreihe im Punktdiagramm durch Angabe des Zeitreihennamens finden.

### Ausreißerverteilung

Balkendiagramm zur Anzahl der Ausreißer in allen Zielen für jedes Zeitintervall in der Schätzperiode. Durch Klicken auf einen Balken im Balkendiagramm wird das Punktdiagramm gefiltert, sodass es nur die zum ausgewählten Balken gehörigen Ausreißer anzeigt.

### Anmerkung:

- Wenn Sie ein Ausgabedokument speichern, das interaktive Ausgabedaten aus der temporalen kausalen Modellierung enthält, und Sie die interaktiven Funktionen beibehalten wollen, stellen Sie sicher, dass **Erforderliche Modellinformationen zusammen mit den Ausgabedokument speichern** im Dialog **Ausgabe speichern unter** ausgewählt ist.
- Einige interaktive Funktionen erfordern, dass das aktive Dataset aus den Daten besteht, aus denen das temporale kausale Modellsystem erstellt wurde.



---

## Kapitel 7. Anwenden temporaler kausaler Modelle

---

### Anwenden temporaler kausaler Modelle

Zwei Prozeduren sind verfügbar, um Modelle anzuwenden, die mit der Prozedur "Temporale kausale Modellierung" erstellt wurden. Beide Prozeduren erfordern die Modellsystemdatei, die als Teil der Prozedur "Temporale kausale Modellierung" gespeichert werden kann.

#### Vorhersage temporaler kausaler Modelle

Mit dieser Prozedur können Sie Vorhersagen für Zeitreihen abrufen, für die weitere aktuelle Daten verfügbar sind, ohne die Modelle erneut zu erstellen. Sie können auch die in der Prozedur "Temporale kausale Modellierung" verfügbaren Ausgaben generieren.

#### Temporale kausale Modellszenarios

Mit dieser Prozedur können Sie überprüfen, wie sich bestimmte Werte einer bestimmten Zeitreihe in einem Modellsystem auf die vorhergesagten Werte der Zeitreihe auswirken, die in kausaler Beziehung zu ihr stehen.

---

### Vorhersage temporaler kausaler Modelle

Die Prozedur "Vorhersage temporaler kausaler Modelle" lädt eine mit der Prozedur "Temporale kausale Modellierung" erstellte Modellsystemdatei und wendet die Modelle auf das aktive Dataset an. Mit dieser Prozedur können Sie Vorhersagen für Zeitreihen abrufen, für die weitere aktuelle Daten verfügbar sind, ohne die Modelle erneut zu erstellen. Sie können auch die in der Prozedur "Temporale kausale Modellierung" verfügbaren Ausgaben generieren.

#### Voraussetzungen

- Die Struktur der Daten im aktiven Dataset, ob spaltenbasiert oder mehrdimensional, muss dieselbe sein wie die, die bei der Modellsystemerstellung verwendet wurde. Bei mehrdimensionalen Daten müssen die Dimensionsfelder dieselben sein wie die, die bei der Erstellung des Modellsystems verwendet wurden. Darüber hinaus müssen die Dimensionswerte, die bei der Modellsystemerstellung verwendet wurden, im aktiven Dataset vorhanden sein.
- Modelle werden auf Felder im aktiven Dataset angewendet, die dieselben Namen wie die im Modellsystem angegebenen Felder aufweisen.
- Das Feld bzw. die Felder, durch die die Beobachtungen bei der Erstellung des Modellsystems definiert wurden, müssen im aktiven Dataset vorhanden sein. Es wird davon ausgegangen, dass das Zeitintervall zwischen den Beobachtungen dasselbe ist wie bei der Erstellung der Modelle. Wenn die Beobachtungen durch eine Datumsspezifikation definiert wurden, muss diese Datumsspezifikation auch im aktiven Dataset vorhanden sein. Datumsspezifikationen werden über das Dialogfeld **Datum definieren** oder den Befehl DATE erstellt.
- Das Zeitintervall für die Analyse und alle Einstellungen für Aggregation, Verteilung und fehlende Werte sind dieselben wie bei der Erstellung der Modelle.

### Verwenden von "Vorhersage temporaler kausaler Modelle"

Für diese Funktion ist die Option "Statistics Forecasting" erforderlich.

Wählen Sie in den Menüs Folgendes aus:

**Analysieren > Vorhersage > Temporale kausale Modelle anwenden...**

1. Geben Sie die Dateispezifikation für eine Modelldatei ein oder klicken Sie auf **Durchsuchen** und wählen Sie eine Modelldatei aus. Modellsystemdateien werden mit der Prozedur **Temporale kausale Modellierung** erstellt.

2. Klicken Sie auf die Option, um Modelle erneut zu schätzen, Vorhersagen zu erstellen und Ausgaben zu generieren.
3. Klicken Sie auf **Weiter**.
4. Geben Sie an, ob Sie vorhandene Modellparameter verwenden oder Modellparameter aus den Daten im aktiven Dataset erneut schätzen wollen.
5. Geben Sie an, wie weit die Vorhersage in die Zukunft reichen soll oder geben Sie an, dass keine Vorhersage erstellt werden soll.
6. Klicken Sie auf **Optionen**, um den Inhalt der Ausgabe festzulegen.
7. Klicken Sie auf **Speichern**, um Vorhersagen zu speichern, und exportieren Sie das aktualisierte Modellsystem in eine externe Datei, wenn Modellparameter erneut geschätzt werden.
8. Klicken Sie auf **Ausführen**, um die Prozedur auszuführen.

## Modellparameter und Vorhersagen

### Laden aus der Modelldatei

Vorhersagen werden ohne erneute Schätzung der Modellparameter durch Verwendung der Modellparameter aus der Modellsystemdatei und der Daten aus dem aktiven Dataset erzeugt. Maße für die Anpassungsgüte, die in der Ausgabe angezeigt und zum Auswählen der am besten angepassten Modelle verwendet werden, werden aus der Modellsystemdatei abgerufen. Diese Anpassungsmaße bilden dann die Daten ab, die bei der Entwicklung (oder letzten Aktualisierung) der einzelnen Modelle verwendet wurden. Diese Option eignet sich zur Erstellung von Vorhersagen und Ausgaben aus den Daten, die für die Erstellung des Modellsystems verwendet wurden.

### Erneut aus den Daten schätzen

Modellparameter werden mithilfe der Daten im aktiven Dataset erneut geschätzt. Die erneute Schätzung von Modellparametern wirkt sich nicht darauf aus, welche Eingaben in das Modell für die einzelnen Ziele eingeschlossen werden. Diese Option ist geeignet, wenn Sie über neue Daten verfügen, die über die ursprüngliche Schätzperiode hinausgehen, und Sie mit den aktualisierten Daten Vorhersagen oder andere Ausgaben generieren wollen.

### Alle Beobachtungen

Gibt an, dass die Schätzperiode mit dem Zeitpunkt der frühesten Beobachtung beginnt und mit dem Zeitpunkt der spätesten Beobachtung bezogen auf alle Zeitreihen endet.

### Nach Start- und Endzeit

Sie können für die Schätzperiode entweder sowohl die Start- als auch die Endzeit angeben oder Sie können entweder die Start- oder die Endzeit angeben. Wenn Sie die Startzeit oder die Endzeit der Schätzperiode weglassen, wird der Standardwert verwendet.

- Wenn die Beobachtungen durch ein Datums-/Zeitfeld definiert werden, geben Sie Werte für den Anfang und für das Ende in demselben Format an, das für das Datums-/Zeitfeld verwendet wurde.
- Geben Sie für Beobachtungen, die durch zyklische Perioden definiert werden, einen Wert für jedes der Felder für die zyklische Periode an. Jedes Feld wird in einer separaten Spalte angezeigt.
- Gilt für das aktive Dataset eine Datumsspezifikation, müssen Sie einen Wert für jede Komponente (beispielsweise einen Monat) der Datumsspezifikation angeben. Jede Komponente wird in einer separaten Spalte angezeigt.
- Wenn die Beobachtungen nach Datensatzreihenfolge definiert werden, werden Anfang und Ende der Schätzperiode über die Zeilennummer (wie im Dateneditor angegeben) des betreffenden Falls definiert.

### Nach spätesten oder frühesten Zeitintervallen

Definiert die Schätzperiode mit einem optionalen Offset als eine bestimmte Anzahl Zeitintervalle, die mit dem frühesten Zeitintervall in den Daten beginnen oder mit dem spätesten Zeitintervall in den Daten enden. In diesem Kontext bezieht sich das Zeitintervall auf

das Zeitintervall der Analyse. Beispiel: Die Beobachtungen erfolgen monatlich, doch das Zeitintervall für die Analyse ist **Quartale**. Wenn Sie **Aktuellste** und den Wert 24 für **Anzahl der Zeitintervalle** angeben, bedeutet dies die letzten 24 Quartale.

Optional können Sie eine bestimmte Anzahl Zeitintervalle ausschließen. Wenn Sie z. B. die aktuellsten 24 Zeitintervalle und für die auszuschließende Anzahl den Wert 1 angeben, bedeutet dies, dass die Schätzperiode aus den 24 Intervallen besteht, die dem letzten Intervall vorausgehen.

#### **Datensätze auf die Zukunft ausdehnen**

Legt die Anzahl der Zeitintervalle fest, für die über das Ende der Schätzperiode hinaus eine Vorhersage erstellt werden soll. In diesem Fall entspricht das Zeitintervall dem Zeitintervall für die Analyse. Bei Anforderung von Vorhersagen werden für alle Eingabezeitreihen, die nicht gleichzeitig auch Ziele sind, automatisch autoregressive Modelle erstellt. Diese Modelle werden dann zum Generieren von Werten für diese Eingabezeitreihen in der Vorhersageperiode verwendet, um die Vorhersagen für die Ziele der Eingaben abzurufen.

## **Allgemeine Optionen**

#### **Konfidenzintervallbreite (%)**

Diese Einstellung steuert die Konfidenzintervalle sowohl für Vorhersagen als auch für Modellparameter. Sie können einen beliebigen positiven Wert kleiner als 100 angeben. Standardmäßig wird ein Konfidenzintervall von 95 % verwendet.

#### **Schwellenwert für Ausreißer (%)**

Eine Beobachtung wird als Ausreißer markiert, wenn die aufgrund des Modells berechnete Wahrscheinlichkeit, dass es sich hierbei um einen Ausreißer handelt, diesen Schwellenwert überschreitet. Sie können einen Wert im Bereich 50 bis 100 angeben.

## **Anzuzeigende Zeitreihe**

Die folgenden Optionen geben die Zeitreihen (Ziele oder Eingaben) an, für die eine Ausgabe angezeigt wird. Der Inhalt der Ausgabe für die angegebenen Zeitreihen wird über die Einstellungen für **Ausgabeoptionen** festgelegt.

#### **Mit am besten angepassten Modellen verbundene Ziele anzeigen**

Standardmäßig wird eine Ausgabe für die Ziele angezeigt, die den zehn laut R-Quadrat-Wert am besten angepassten Modellen zugeordnet sind. Sie können auch eine andere feste Anzahl von am besten angepassten Modellen oder einen Prozentsatz für am besten angepasste Modelle angeben. Darüber hinaus können Sie folgende Maße für die Anpassungsgüte auswählen:

##### **R-Quadrat**

Ein Maß für die Anpassungsgüte eines linearen Modells. Wird auch als Bestimmtheitskoeffizient bezeichnet. Es gibt den Anteil der Variation der Zielvariablen an, der durch das Modell erklärt wird. Er liegt zwischen 0 und 1. Kleine Werte zeigen an, dass das Modell nicht gut zu den Daten passt.

##### **Wurzel der mittleren Fehlerquadratsumme (Prozent)**

Maß für die Differenz zwischen vom Modell vorhergesagten Werten und den beobachteten Werten der Zeitreihe. Es ist unabhängig von den verwendeten Maßeinheiten und kann daher verwendet werden, um Zeitreihen mit unterschiedlichen Einheiten zu vergleichen.

##### **Quadratwurzel des mittleren quadratischen Fehlers**

Die Quadratwurzel des mittleren quadratischen Fehlers. Ein Maß dafür, wie stark eine abhängige Zeitreihe von ihrem durch das Modell vorhergesagten Niveau abweicht, und zwar ausgedrückt in derselben Maßeinheit wie die abhängige Zeitreihe.

**BIC** Bayes-Informationskriterium. Ein Maß für die Auswahl und den Vergleich von Modellen,

das auf  $-2$  Log-Likelihood beruht. Kleinere Werte stehen für bessere Modelle. BIC "penalisiert" überparametrisierte Modelle ebenfalls (beispielsweise Modelle mit einer großen Anzahl Eingaben). allerdings strikter als AIC.

**AIC** Akaike-Informationskriterium. Ein Maß für die Auswahl und den Vergleich von Modellen, das auf  $-2$  Log-Likelihood beruht. Kleinere Werte stehen für bessere Modelle. AIC "penalisiert" überparametrisierte Modelle (beispielsweise Modelle mit einer großen Anzahl Eingaben.)

### Einzelne Zeitreihe angeben

Sie können einzelne Zeitreihen angeben, für die eine Ausgabe erstellt werden soll.

- Bei spaltenbasierten Daten geben Sie die Felder mit den Zeitreihen an, die Sie benötigen. Die Reihenfolge der angegebenen Felder legt die Reihenfolge fest, in der diese in der Ausgabe angezeigt werden.
- Bei mehrdimensionalen Daten geben Sie eine bestimmte Zeitreihe an, indem Sie im Raster für das Metrikfeld, das die Zeitreihe enthält, eine Eingabe hinzufügen. Anschließend geben Sie die Werte der Dimensionsfelder an, die die Zeitreihe definieren.
  - Sie können die Werte für die einzelnen Dimensionsfelder direkt in das Raster eingeben oder sie aus der Liste der verfügbaren Dimensionswerte auswählen. Wenn Sie Werte aus der Liste der verfügbaren Dimensionswerte auswählen wollen, klicken Sie auf die Schaltfläche mit Auslassungspunkten in der Zelle für die gewünschte Dimension. Daraufhin wird das Unterdialogfeld **Dimensionswerte auswählen** geöffnet.
  - Sie können die Liste der Dimensionswerte durchsuchen, indem Sie im Unterdialogfeld **Dimensionswerte auswählen** auf das Fernglassymbol klicken und einen Suchbegriff angeben. Leerzeichen werden als Teil des Suchbegriffs behandelt. Sterne (\*) im Suchbegriff zeigen keine Platzhalterzeichen an.
  - Die Reihenfolge der Zeitreihen im Raster legt die Reihenfolge fest, in der diese in der Ausgabe angezeigt werden.

Sowohl für spaltenbasierte als auch für mehrdimensionale Daten ist die Ausgabe auf 30 Zeitreihen begrenzt. Dieser Höchstwert umfasst einzelne von Ihnen angegebene Zeitreihen (Eingaben oder Ziele) und Ziele, die mit am besten angepassten Modellen verbunden sind. Einzeln angegebene Zeitreihen haben Vorrang vor Zielen, die mit am besten angepassten Modellen verbunden sind.

## Ausgabeoptionen

Diese Optionen legen den Inhalt der Ausgabe fest. Die Optionen in der Gruppe **Ausgabe für Inhalte** generieren eine Ausgabe für die Ziele, die in den Einstellungen für **Anzuzeigende Zeitreihe** mit den am besten angepassten Modellen verbunden sind. Die Optionen in der Gruppe **Ausgabe für Inhalte** generieren eine Ausgabe für die einzelnen Zeitreihen, die in den Einstellungen für **Anzuzeigende Zeitreihe** angegeben sind.

### Gesamtmodellsystem

Zeigt eine grafische Darstellung der kausalen Beziehungen zwischen Zeitreihen im Modellsystem an. Tabellen sowohl für Statistiken zur Modellanpassungsgüte als auch und zu Ausreißern für die angezeigten Ziele sind in das Ausgabeelement eingeschlossen. Wenn diese Option in der Gruppe **Ausgabe für Zeitreihen** ausgewählt wird, wird ein separates Ausgabeelement für jede einzelne Zeitreihe generiert, die in den Einstellungen für **Anzuzeigende Zeitreihe** angegeben ist.

Kausale Beziehungen zwischen Zeitreihen haben ein zugeordnetes Signifikanzniveau, wobei ein kleineres Signifikanzniveau eine signifikantere Verbindung anzeigt. Sie können auswählen, dass Beziehungen mit einem Signifikanzniveau oberhalb eines bestimmten Wertes ausgeblendet werden.

## Statistik zur Anpassungsgüte des Modells und Ausreißer

Tabellen für Statistiken zur Modellanpassungsgüte und zu Ausreißern für die Zielzeitreihen, die zur Anzeige ausgewählt sind. Diese Tabellen enthalten dieselben Informationen wie die Tabellen zur Visualisierung des Gesamtmodellsystems. Sie unterstützen alle Standardfunktionen für das Drehen und Bearbeiten von Tabellen.

## Modelleffekte und Modellparameter

Tabellen zu Tests von Modelleffekten und zu Modellparametern für die Zielzeitreihen, die zur Anzeige ausgewählt sind. Tests zu Modelleffekten beinhalten die F-Statistik und zugehörige Signifikanzwerte für jede in das Modell eingeschlossene Eingabe.

## Wirkungsdiagramm

Zeigt eine grafische Darstellung der kausalen Beziehungen zwischen einer relevanten Zeitreihe und anderen Zeitreihen, auf die sie sich auswirkt oder die sich auf sie auswirken. Zeitreihen, die sich auf die relevante Zeitreihe auswirken, werden als *Ursachen* bezeichnet. Wenn Sie **Effekte** auswählen, wird ein für die Anzeige der Effekte initialisiertes Wirkungsdiagramm generiert. Wenn Sie **Ursachen** auswählen, wird ein für die Anzeige der Ursachen initialisiertes Wirkungsdiagramm generiert. Wenn Sie **Sowohl Ursachen als auch Effekte** auswählen, werden zwei separate Wirkungsdiagramme generiert, eins wird für die Ursachen und eines wird für die Effekte initialisiert. Im Ausgabeelement, das das Wirkungsdiagramm anzeigt, können Sie interaktiv zwischen Ursachen und Effekten wechseln.

Sie können die Anzahl der anzuzeigenden Ebenen von Ursachen oder Effekten angeben, wobei die erste Ebene nur die relevante Zeitreihe ist. Jede zusätzliche Ebene zeigt weitere indirekte Ursachen oder Effekte bezogen auf die relevante Zeitreihe an. So besteht z. B. die dritte Ebene in der Anzeige der Effekte aus den Zeitreihen, die Zeitreihen in der zweiten Ebene als direkte Eingabe enthalten. Auf Zeitreihen der dritten Ebene wirkt sich die relevante Zeitreihe dann indirekt aus, da die relevante Zeitreihe eine direkte Eingabe für die Zeitreihen in der zweiten Ebene ist.

## Zeitreihendiagramm

Diagramme beobachteter und vorhergesagter Werte für die Zielzeitreihen, die zur Anzeige ausgewählt sind. Wenn Vorhersagen angefordert werden, zeigt das Diagramm auch die vorhergesagten Werte und die Konfidenzintervalle für die Vorhersagen an.

## Residuendiagramme

Diagramme der Modellresiduen für die Zielzeitreihen, die zur Anzeige ausgewählt sind.

## Top Eingaben

Diagramme der einzelnen angezeigten Ziele über einen bestimmten Zeitraum sowie die drei Top Eingaben für das Ziel. Die Top Eingaben sind die Eingaben mit dem niedrigsten Signifikanzwert. Um unterschiedliche Maßstäbe für die Eingaben und Ziele veranschaulichen zu können, stellt die Y-Achse den Z-Score für die einzelnen Zeitreihen dar.

## Vorhersagetabelle

Tabellen mit vorhergesagten Werten und Konfidenzintervallen dieser Vorhersagen für die Zielzeitreihen, die zur Anzeige ausgewählt sind.

## Ursachenanalyse für Ausreißer

Legt fest, welche Zeitreihen die wahrscheinlichsten Ursachen für die einzelnen Ausreißer in einer relevanten Zeitreihe sind. Eine Ursachenanalyse für Ausreißer wird für jede Zielzeitreihe durchgeführt, die in der Liste der einzelnen Zeitreihen in den Einstellungen für **Anzuzeigende Zeitreihe** enthalten sind.

## Ausgabe

### Interaktive Ausreißertabelle und Diagramm

Tabelle und Diagramm mit Ausreißern und Ursachen für diese Ausreißer für jede der relevanten Zeitreihen. Die Tabelle enthält eine Zeile je Ausreißer. Bei dem Diagramm handelt es sich um ein Wirkungsdiagramm. Bei Auswahl einer Tabel-

lenzeile wird im Wirkungsdiagramm der Pfad von der relevanten Zeitreihe bis zu der Zeitreihe hervorgehoben, die die wahrscheinlichste Ursache für den zugehörigen Ausreißer ist.

#### **Pivot-Tabelle mit Ausreißern**

Tabelle mit Ausreißern und Ursachen für diese Ausreißer für jede der relevanten Zeitreihen. Diese Tabelle enthält dieselben Informationen wie die Tabelle in der interaktiven Anzeige. Sie unterstützt alle Standardfunktionen für das Drehen und Bearbeiten von Tabellen.

#### **Kausale Ebenen**

Sie können die Anzahl der Ebenen angeben, die in die Suche nach den Ursachen eingeschlossen werden sollen. Das hier verwendete Konzept der Ebenen entspricht dem, das auch für Wirkungsdiagramme beschrieben wurde.

#### **Modellübergreifende Anpassungsgüte des Modells**

Histogramm der Anpassungsgüte für alle Modelle und für ausgewählte Statistiken zur Anpassungsgüte. Folgende Statistiken zur Anpassungsgüte sind verfügbar:

##### **R-Quadrat**

Ein Maß für die Anpassungsgüte eines linearen Modells. Wird auch als Bestimmtheitskoeffizient bezeichnet. Es gibt den Anteil der Variation der Zielvariablen an, der durch das Modell erklärt wird. Er liegt zwischen 0 und 1. Kleine Werte zeigen an, dass das Modell nicht gut zu den Daten passt.

##### **Wurzel der mittleren Fehlerquadratsumme (Prozent)**

Maß für die Differenz zwischen vom Modell vorhergesagten Werten und den beobachteten Werten der Zeitreihe. Es ist unabhängig von den verwendeten Maßeinheiten und kann daher verwendet werden, um Zeitreihen mit unterschiedlichen Einheiten zu vergleichen.

##### **Quadratwurzel des mittleren quadratischen Fehlers**

Die Quadratwurzel des mittleren quadratischen Fehlers. Ein Maß dafür, wie stark eine abhängige Zeitreihe von ihrem durch das Modell vorhergesagten Niveau abweicht, und zwar ausgedrückt in derselben Maßeinheit wie die abhängige Zeitreihe.

**BIC** Bayes-Informationskriterium. Ein Maß für die Auswahl und den Vergleich von Modellen, das auf  $-2 \log$ -Likelihood beruht. Kleinere Werte stehen für bessere Modelle. BIC "penalisiert" überparametrisierte Modelle ebenfalls (beispielsweise Modelle mit einer großen Anzahl Eingaben). allerdings strikter als AIC.

**AIC** Akaike-Informationskriterium. Ein Maß für die Auswahl und den Vergleich von Modellen, das auf  $-2 \log$ -Likelihood beruht. Kleinere Werte stehen für bessere Modelle. AIC "penalisiert" überparametrisierte Modelle (beispielsweise Modelle mit einer großen Anzahl Eingaben.)

#### **Ausreißer über einen Zeitraum**

Balkendiagramm zur Anzahl der Ausreißer in allen Zielen für jedes Zeitintervall in der Schätzperiode.

#### **Zeitreihentransformationen**

Tabelle aller Transformationen, die auf die Zeitreihe im Modellsystem angewandt wurden. Zu den möglichen Transformationen gehören die Imputation fehlender Werte, die Aggregation und die Verteilung.

## **Speichern**

### **Ziele speichern**

Sie können Modellvorhersagen, Konfidenzintervalle und Residuen als neue Variablen speichern. Jedes zum Speichern angegebene Ziel generiert eine eigene Gruppe neuer Variablen und jede neue Variable enthält Werte für Schätz- und Vorhersageperioden. Für vorhergesagte Werte, Konfi-

denzintervalle und Restauschen können Sie den Stammnamen angeben, der als Präfix für die neuen Variablen verwendet werden soll. Der vollständige Variablenname ist die Verkettung des Stammnamens und des Namens des Felds, das die Zielzeitreihe enthält. Der Stammname muss den Regeln für gültige Variablennamen entsprechen. Der Variablenname wird gegebenenfalls erweitert, um das Auftreten von Namenskonflikten zu vermeiden.

#### **Gibt Fälle mit Vorhersagen an**

Erstellt eine Variable, die angibt, ob ein Datensatz Vorhersagedaten enthält. Sie können den Variablennamen angeben. Der Standardwert ist *ForecastIndicator*.

#### **Zu speichernde Ziele**

Gibt an, ob für alle Zielzeitreihen im Modellsystem neue Variablen erstellt werden oder nur für die Zielzeitreihen, die in den Einstellungen für **Anzuzeigende Zeitreihe** angegeben sind.

#### **Zieloptionen**

Sie können sowohl Transformationen der Daten (beispielsweise Aggregation oder Imputation fehlender Werte) als auch neue Variablen (angegeben in den Einstellungen für **Ziele speichern**) in einer IBM SPSS Statistics-Datendatei oder in einem neuen Dataset in der aktuellen Sitzung speichern. Datums-/Zeitwerte in den gespeicherten Daten werden am Anfang jedes Zeitintervalls ausgerichtet, z. B. am Monatsersten, und stellen das Zeitintervall für die Analyse des Modellsystems dar. Sie können alle neuen Variablen nur dann im aktiven Dataset speichern, wenn die Beobachtungen durch eine Datumsspezifikation oder die Datensatzreihenfolge definiert werden und die Daten nicht aggregiert sind.

#### **Modellsystem exportieren**

Speichert das Modellsystem in einem komprimierten Dateiarhiv (ZIP-Datei). Die Modellsystemdatei kann von dieser Prozedur wiederverwendet werden. Sie kann auch von der Prozedur "Temporale kausale Modellszenarios" verwendet werden, um eine Szenarioanalyse durchzuführen. Diese Option ist nur verfügbar, wenn Modellparameter erneut geschätzt werden.

---

## **Temporale kausale Modellszenarios**

Die Prozedur "Temporale kausale Modellszenarios" führt mit den Daten aus dem aktiven Dataset benutzerdefinierte Szenarios für ein temporales kausales Modellsystem durch. Ein *Szenario* wird durch eine als *Stammzeitreihe* bezeichnete Zeitreihe und ein Set benutzerdefinierter Werte für diese Zeitreihe über einen bestimmten Zeitbereich definiert. Die angegebenen Werte werden dann verwendet, um Vorhersagen für die Zeitreihen zu generieren, auf die sich die Stammzeitreihe auswirkt. Die Prozedur erfordert eine Modellsystemdatei, die von der Prozedur "Temporale kausale Modellierung" erstellt wurde. Es wird vorausgesetzt, dass das aktive Dataset dieselben Daten enthält, die für die Erstellung der Modellsystemdatei verwendet wurden.

### **Beispiel**

Ein Entscheidungsträger in einem Unternehmen hat mit der Prozedur "Temporale kausale Modellierung" einen Schlüsselwert entdeckt, der sich auf eine Reihe von Leistungsindikatoren auswirkt. Da der Messwert steuerbar ist, will der Entscheidungsträger den Effekt mehrerer Sets von Werten für den Messwert innerhalb des nächsten Quartals untersuchen. Die Untersuchung kann ohne großen Aufwand durch Laden der Modellsystemdatei in die Prozedur "Temporale kausale Modellszenarios" und Angabe der Sets von Werten für den Schlüsselwert durchgeführt werden.

## **Ausführen von "Temporale kausale Modellszenarios"**

Für diese Funktion ist die Option "Statistics Forecasting" erforderlich.

Wählen Sie in den Menüs Folgendes aus:

**Analysieren > Vorhersage > Temporale kausale Modelle anwenden...**

1. Geben Sie die Dateispezifikation für eine Modelldatei ein oder klicken Sie auf **Durchsuchen** und wählen Sie eine Modelldatei aus. Modellsystemdateien werden mit der Prozedur **Temporale kausale Modellierung** erstellt.
2. Klicken Sie auf die Option, um Szenarios auszuführen.
3. Klicken Sie auf **Weiter**.
4. Klicken Sie auf der Registerkarte **Szenarios** im Dialogfeld **Temporale kausale Modellszenarios** auf **Szenarioperiode definieren** und geben Sie die Szenarioperiode an.
5. Klicken Sie für spaltenbasierte Daten auf **Szenario hinzufügen**, um die einzelnen Szenarios zu definieren. Klicken Sie für mehrdimensionale Daten auf **Szenario hinzufügen**, um die einzelnen Szenarios zu definieren, und klicken Sie auf **Szenariogruppe hinzufügen**, um die einzelnen Szenariogruppen zu definieren.
6. Klicken Sie auf **Optionen**, um den Inhalt der Ausgabe und Umfang der Zeitreihe anzugeben, auf die sich das Szenario auswirkt.
7. Klicken Sie auf **Ausführen**, um die Prozedur auszuführen.

## Definieren der Szenarioperiode

Die Szenarioperiode ist der Zeitraum, für den Sie die Werte angeben, die für die Ausführung Ihrer Szenarios verwendet werden. Sie kann vor oder nach dem Ende der Schätzperiode beginnen. Sie können optional angeben, dass über das Ende der Szenarioperiode hinaus Vorhersagen erstellt werden sollen. Standardmäßig werden Vorhersagen bis zum Ende der Szenarioperiode generiert. Alle Szenarios verwenden dieselbe Szenarioperiode und dieselben Spezifikationen dafür, bis wann Vorhersagen erstellt werden sollen.

**Anmerkung:** Vorhersagen beginnen in der ersten Zeitperiode nach dem Anfang der Szenarioperiode. Wenn die Szenarioperiode beispielsweise am 01.11.2014 beginnt und das Zeitintervall **Monate** ist, wird die erste Vorhersage für den 01.12.2014 erstellt.

### Nach Anfang, Ende und 'Vorhersagen bis' angeben

- Wenn die Beobachtungen durch ein Datums-/Zeitfeld definiert werden, geben Sie Werte für den Anfang, das Ende und für "Vorhersagen bis" in demselben Format an, das für das Datums-/Zeitfeld verwendet wird. Werte für Datums-/Zeitfelder werden am Anfang des zugeordneten Zeitintervalls ausgerichtet. Wenn das Zeitintervall für die Analyse z. B. **Monate** ist, wird der Wert 10/10/2014 in den Wert 10/01/2014 korrigiert, weil dies der Monatsbeginn ist.
- Geben Sie für Beobachtungen, die durch zyklische Perioden definiert werden, einen Wert für jedes der Felder für die zyklische Periode an. Jedes Feld wird in einer separaten Spalte angezeigt.
- Gilt für das aktive Dataset eine Datumsspezifikation, müssen Sie einen Wert für jede Komponente (beispielsweise einen Monat) der Datumsspezifikation angeben. Jede Komponente wird in einer separaten Spalte angezeigt.
- Wenn die Beobachtungen nach Datensatzreihenfolge definiert werden, werden die Werte für Anfang, Ende und "Vorhersagen bis" über die Zeilennummer (wie im Dateneditor angegeben) des betreffenden Falls definiert.

### Nach Zeitintervallen relativ zum Ende der Schätzperiode angeben

Definiert Anfang und Ende in Bezug auf die Anzahl der Zeitintervalle relativ zum Ende der Schätzperiode, wobei das Zeitintervall dem Zeitintervall der Analyse entspricht. Das Ende der Schätzperiode wird als Zeitintervall 0 definiert. Zeitintervalle vor dem Ende der Schätzperiode haben negative Werte und Intervalle nach dem Ende der Schätzperiode haben positive Werte. Sie können auch angeben, für wie viele Intervalle nach dem Ende der Szenarioperiode Vorhersagen erstellt werden sollen. Der Standardwert ist 0.

Beispiel: Das Zeitintervall für die Analyse ist **Monate** und Sie geben für das Startintervall den Wert 1, für das Endintervall den Wert 3 und für darüber hinaus reichende Vorhersagen den Wert 1 an. Die Szenarioperiode entspricht dann den drei Monaten nach dem Ende der Schätzperiode.

Es werden Vorhersagen für den zweiten und dritten Monat der Szenarioperiode und für einen weiteren Monat nach dem Ende der Szenarioperiode generiert.

## Hinzufügen von Szenarios und Szenariogruppen

Die Registerkarte **Szenarios** gibt die auszuführenden Szenarios an. Um Szenarios zu definieren, müssen Sie zunächst die Szenarioperiode definieren, indem Sie auf **Szenarioperiode definieren** klicken. Sie können Szenarios und Szenariogruppen (nur bei mehrdimensionalen Daten) erstellen, indem Sie auf die Schaltfläche **Szenario hinzufügen** bzw. **Szenariogruppe hinzufügen** klicken. Wenn Sie ein bestimmtes Szenario oder eine bestimmte Szenariogruppe im zugehörigen Raster auswählen, können Sie diese bearbeiten, eine Kopie davon erstellen oder sie löschen.

## Spaltenbasierte Daten

Die Spalte **Stammfeld** im Raster gibt das Zeitreihenfeld an, dessen Werte durch die Szenariowerte ersetzt werden. In der Spalte **Szenariowerte** werden die angegebenen Szenariowerte in der Reihenfolge vom frühesten bis zum aktuellsten angezeigt. Wenn die Szenariowerte durch einen Ausdruck definiert werden, zeigt die Spalte den Ausdruck an.

## Mehrdimensionale Daten

### Einzelne Szenarios

Jede Zeile im Raster **Einzelne Szenarios** gibt eine Zeitreihe an, deren Werte durch die angegebenen Szenariowerte definiert werden. Die Zeitreihe wird durch die Kombination aus dem in der Spalte **Stammmetrik** angegebenen Feld und den für die jeweiligen Dimensionsfelder angegebenen Werten definiert. Der Inhalt der Spalte **Szenariowerte** ist derselbe wie für spaltenbasierte Daten.

### Szenariogruppen

Eine *Szenariogruppe* definiert ein Set von Szenarios, die auf einem einzelnen Stammmetrikfeld basieren, und mehrere Sets von Dimensionswerten. Jedes Set von Dimensionswerten (ein Wert je Dimensionsfeld) definiert für das angegebene Metrikfeld eine Zeitreihe. Dann wird ein einzelnes Szenario für jede dieser Zeitreihen generiert, deren Werte anschließend durch die Szenariowerte ersetzt werden. Szenariowerte für eine Szenariogruppe werden durch einen Ausdruck angegeben, der anschließend auf jede Zeitreihe in der Gruppe angewendet wird.

In der Spalte **Anzahl Zeitreihen** wird die Anzahl der Sets von Dimensionswerten angezeigt, die einer Szenariogruppe zugeordnet sind. Der angezeigte Wert kann größer sein als die tatsächliche Anzahl an Zeitreihen, die der Szenariogruppe zugeordnet sind (eine Zeitreihe je Set). Diese Bedingung tritt auf, wenn einige der angegebenen Kombinationen von Dimensionswerten nicht den in der Stammmetrik für die Gruppe enthaltenen Zeitreihen entsprechen.

Ein Beispiel für eine Szenariogruppe wäre ein Metrikfeld *advertising* und zwei Dimensionsfelder *region* und *brand*. Sie können eine Szenariogruppe definieren, die auf *advertising* als Stammmetrik basiert und die alle Kombinationen von *region* und *brand* einschließt. Sie können dann  $advertising*1.2$  als Ausdruck angeben, um die Auswirkung einer Erhöhung von *advertising* um 20 % auf alle Zeitreihen zu prüfen, die dem Feld *advertising* zugeordnet sind. Bei vier Werten für *region* und zwei Werten für *brand* gäbe es acht solcher Zeitreihen und demnach acht durch die Gruppe definierten Szenarios.

## Szenariodefinition

Die Einstellungen zum Definieren eines Szenarios richten sich danach, ob es um spaltenbasierte oder mehrdimensionale Daten geht.

### Stammzeitreihen

Gibt die Stammzeitreihe für das Szenario an. Jedes Szenario basiert auf einer einzelnen Stammzeitreihe. Bei spaltenbasierten Daten wählen Sie das Feld aus, das die Stammzeitreihe definiert. Bei mehrdimensionalen Daten geben Sie die Stammzeitreihe an, indem Sie im Raster für das Met-

rikfeld, das die Zeitreihe enthält, eine Eingabe hinzufügen. Anschließend geben Sie die Werte der Dimensionsfelder an, die die Stammzeitreihe definieren. Für die Angabe der Dimensionswerte gilt Folgendes:

- Sie können die Werte für die einzelnen Dimensionsfelder direkt in das Raster eingeben oder sie aus der Liste der verfügbaren Dimensionswerte auswählen. Wenn Sie Werte aus der Liste der verfügbaren Dimensionswerte auswählen wollen, klicken Sie auf die Schaltfläche mit Auslassungspunkten in der Zelle für die gewünschte Dimension. Daraufhin wird das Unterdialogfeld **Dimensionswerte auswählen** geöffnet.
- Sie können die Liste der Dimensionswerte durchsuchen, indem Sie im Unterdialogfeld **Dimensionswerte auswählen** auf das Fernglassymbol klicken und einen Suchbegriff angeben. Leerzeichen werden als Teil des Suchbegriffs behandelt. Sterne (\*) im Suchbegriff zeigen keine Platzhalterzeichen an.

### **Betroffene Ziele angeben**

Verwenden Sie diese Option, wenn Sie wissen, auf welche Ziele sich die Stammzeitreihe auswirkt und Sie nur die Auswirkungen auf diese Ziele untersuchen wollen. Standardmäßig werden Ziele, auf die sich die Stammzeitreihe auswirkt, automatisch ermittelt. Sie können den Umfang der Zeitreihen, auf die sich das Szenario auswirkt, über Einstellungen auf der Registerkarte **Optionen** angeben.

Wählen Sie bei spaltenbasierten Daten die gewünschten Ziele aus. Bei mehrdimensionalen Daten geben Sie Zeitreihen an, indem Sie im Raster für das Zielmetrikfeld, das die Zeitreihe enthält, eine Eingabe hinzufügen. Standardmäßig werden alle in dem angegebenen Metrikfeld enthaltenen Zeitreihen eingeschlossen. Sie können das Set der eingeschlossenen Zeitreihen anpassen, indem Sie die eingeschlossenen Werte für mindestens eins der Dimensionsfelder anpassen. Wenn Sie die eingeschlossenen Dimensionswerte anpassen wollen, klicken Sie auf die Schaltfläche mit Auslassungspunkten für die betreffende Dimension. Daraufhin wird das Dialogfeld **Dimensionswerte auswählen** geöffnet.

Die Spalte **Anzahl Zeitreihen** (für mehrdimensionale Daten) zeigt die Anzahl der Sets von Dimensionswerten an, die aktuell für die zugehörige Zielmetrik angegeben sind. Der angezeigte Wert kann größer als die tatsächliche Anzahl an betroffenen Zielzeitreihen sein (eine Zeitreihe pro Set). Diese Bedingung tritt auf, wenn einige der angegebenen Kombinationen von Dimensionswerten nicht den in der zugeordneten Zielmetrik enthaltenen Zeitreihen entsprechen.

### **Szenario-ID**

Jedes Szenario muss über eine eindeutige ID verfügen. Die ID wird in der Ausgabe angezeigt, die dem Szenario zugeordnet ist. Außer der Eindeutigkeit gibt es keine Einschränkungen für den Wert dieser ID.

### **Szenariowerte für Stammzeitreihen angeben**

Mit dieser Option können Sie explizite Werte für die Stammzeitreihe in der Szenarioperiode angeben. Sie müssen für jedes Zeitintervall, das im Raster angezeigt ist, einen numerischen Wert angeben. Sie können die Werte der Stammzeitreihe (tatsächliche oder vorhergesagte) für die einzelnen Intervalle in der Szenarioperiode abrufen, indem Sie auf **Lesen, Vorhersage** oder **Lesen/Vorhersage** klicken.

### **Ausdruck für Szenariowerte für Stammzeitreihen angeben**

Sie können für die Berechnung der Werte der Stammzeitreihe in der Szenarioperiode einen Ausdruck definieren. Diesen Ausdruck können Sie direkt eingeben oder auf die Schaltfläche für den Taschenrechner klicken und den Ausdruck über den Expression Builder für Szenariowerte erstellen.

- Der Ausdruck kann ein beliebiges Ziel oder eine beliebige Eingabe aus dem Modellsystem enthalten.
- Wenn die Szenarioperiode über die vorhandenen Daten hinausgeht, wird der Ausdruck auf vorhergesagte Werte der Felder im Ausdruck angewendet.

- Bei mehrdimensionalen Daten gibt jedes Feld im Ausdruck eine Zeitreihe an, die durch das für die Stammmetrik angegebene Feld und die Dimensionswerte definiert wird. Diese Zeitreihen werden zum Auswerten des Ausdrucks verwendet.

Beispiel: Das Stammfeld ist *advertising* und der Ausdruck ist *advertising\*1.2*. Die im Szenario verwendeten Werte für *advertising* stellen eine 20-prozentige Steigerung gegenüber den vorhandenen Werten dar.

**Anmerkung:** Sie können Szenarios erstellen, indem Sie auf der Registerkarte **Szenarios** auf die Option **Szenario hinzufügen** klicken.

**Expression Builder für Szenariowerte:** Verwenden Sie den Expression Builder für Szenariowerte, um einen Ausdruck zu erstellen, der zur Berechnung von Szenariowerten für ein einzelnes Szenario oder eine Szenariogruppe verwendet wird. Um einen Ausdruck zu erstellen, müssen Sie Komponenten aus der Zwischenablage in das Feld **Ausdruck** einfügen oder direkt in dieses Feld eingeben.

- Der Ausdruck kann ein beliebiges Ziel oder eine beliebige Eingabe aus dem Modellsystem enthalten.
- Sie können Funktionen einfügen, indem Sie eine Gruppe aus der Liste **Funktionsgruppe** auswählen und in der Liste **Funktionen** auf die Funktion doppelklicken (oder die Funktion auswählen und auf den Pfeil neben der Liste "Funktionsgruppe" klicken). Geben Sie alle durch Fragezeichen gekennzeichneten Parameter ein. Die Funktionsgruppe mit der Beschriftung **Alle** bietet eine Auflistung aller verfügbaren Funktionen. Eine kurze Beschreibung der aktuell ausgewählten Funktion wird in einem speziellen Bereich des Dialogfelds angezeigt.
- Zeichenfolgekonstanten müssen in Anführungszeichen eingeschlossen werden.
- Wenn die Werte Dezimalstellen enthalten, muss ein Punkt (.) als Dezimaltrennzeichen verwendet werden.

Klicken Sie für den Zugriff auf Expression Builder für Szenariowerte auf die Schaltfläche für den Taschenrechner im Dialogfeld **Szenariodefinition** oder **Szenariogruppendifinition**.

**Auswählen von Dimensionswerten:** Bei mehrdimensionalen Daten können Sie die Dimensionswerte anpassen, die die Ziele definieren, auf die sich ein Szenario oder eine Szenariogruppe auswirkt. Sie können auch die Dimensionswerte anpassen, die das Set von Stammzeitreihen für eine Szenariogruppe definieren.

#### **Alle Werte**

Gibt an, dass alle Werte des aktuellen Dimensionsfelds eingeschlossen sind. Dies ist die Standardoption.

#### **Werte auswählen**

Mit dieser Option können Sie das Set von Werten für das aktuelle Dimensionsfeld angeben. Sie können das Set mit auswählbaren Werten filtern. Werte, die mit der Filterbedingung übereinstimmen, sind auf der Registerkarte **Übereinstimmung** aufgeführt, während Werte, die mit dieser Bedingung nicht übereinstimmen, auf der Registerkarte **Keine Übereinstimmung** der Liste **Nicht ausgewählte Werte** angezeigt werden. Auf der Registerkarte **Alle** werden alle nicht ausgewählten Werte unabhängig von Filterbedingungen aufgelistet.

- Sie können Sterne (\*) verwenden, um beim Setzen von Filtern Platzhalterzeichen anzugeben.
- Wenn Sie den aktuellen Filter löschen wollen, geben Sie im Dialogfeld **Angezeigte Werte filtern** für den Suchbegriff einen leeren Wert an.

So passen Sie Dimensionswerte für betroffene Ziele an:

1. Wählen Sie im Dialogfeld **Szenariodefinition** oder **Szenariogruppendifinition** die Zielmetrik aus, für die Sie die Dimensionswerte anpassen wollen.
2. Klicken Sie auf die Schaltfläche mit Auslassungspunkten in der Spalte für die Dimension, die Sie anpassen wollen.

So passen Sie Dimensionswerte für die Stammzeitreihen einer Szenariogruppe an:

1. Klicken Sie im Dialogfeld **Szenariogruppendifinition** auf die Schaltfläche mit Auslassungspunkten (im Raster der Stammzeitreihe) für die Dimension, die Sie anpassen wollen.

## Szenariogruppendifinition

### Stammzeitreihen

Gibt das Set von Stammzeitreihen für die Szenariogruppe an. Für jede Zeitreihe im Set wird ein einzelnes Szenario generiert. Sie geben die Stammzeitreihe an, indem Sie im Raster für das Metrikfeld, das die gewünschte Zeitreihe enthält, eine Eingabe hinzufügen. Anschließend geben Sie die Werte der Dimensionsfelder an, die das Set definieren. Standardmäßig werden alle in dem angegebenen Stammmetrikfeld enthaltenen Zeitreihen eingeschlossen. Sie können das Set der eingeschlossenen Zeitreihen anpassen, indem Sie die eingeschlossenen Werte für mindestens eins der Dimensionsfelder anpassen. Wenn Sie die eingeschlossenen Dimensionswerte anpassen wollen, klicken Sie auf die Schaltfläche mit Auslassungspunkten für eine Dimension. Daraufhin wird das Dialogfeld **Dimensionswerte auswählen** geöffnet.

Die Spalte **Anzahl Zeitreihen** zeigt die Anzahl der Sets von Dimensionswerten an, die aktuell für die zugehörige Stammmetrik eingeschlossen sind. Der angezeigte Wert kann größer sein als die tatsächliche Anzahl an Stammzeitreihen für die Szenariogruppe (eine Zeitreihe je Set). Diese Bedingung tritt auf, wenn einige der angegebenen Kombinationen von Dimensionswerten nicht den in der Stammmetrik enthaltenen Zeitreihen entsprechen.

### Betroffene Ziel-Zeitreihe angeben

Verwenden Sie diese Option, wenn Sie wissen, auf welche Ziele sich das Set von Stammzeitreihen auswirkt und Sie nur die Auswirkungen auf diese Ziele untersuchen wollen. Standardmäßig werden Ziele, auf die sich die einzelnen Stammzeitreihen auswirken, automatisch ermittelt. Sie können den Umfang der Zeitreihen, auf die sich die einzelnen Szenarios auswirken, über Einstellungen auf der Registerkarte **Optionen** angeben.

Sie geben Zielzeitreihen an, indem Sie im Raster für das Metrikfeld, das die Zeitreihe enthält, eine Eingabe hinzufügen. Standardmäßig werden alle in dem angegebenen Metrikfeld enthaltenen Zeitreihen eingeschlossen. Sie können das Set der eingeschlossenen Zeitreihen anpassen, indem Sie die eingeschlossenen Werte für mindestens eins der Dimensionsfelder anpassen. Wenn Sie die eingeschlossenen Dimensionswerte anpassen wollen, klicken Sie auf die Schaltfläche mit Auslassungspunkten für die betreffende Dimension. Daraufhin wird das Dialogfeld **Dimensionswerte auswählen** geöffnet.

Die Spalte **Anzahl Zeitreihen** zeigt die Anzahl der Sets von Dimensionswerten an, die aktuell für die zugehörige Zielmetrik angegeben sind. Der angezeigte Wert kann größer als die tatsächliche Anzahl an betroffenen Zielzeitreihen sein (eine Zeitreihe pro Set). Diese Bedingung tritt auf, wenn einige der angegebenen Kombinationen von Dimensionswerten nicht den in der zugeordneten Zielmetrik enthaltenen Zeitreihen entsprechen.

### Szenario-ID-Präfix

Jede Szenariogruppe muss über ein eindeutiges Präfix verfügen. Das Präfix wird zum Erstellen einer ID verwendet, die in der den einzelnen Szenarios in der Szenariogruppe zugeordneten Ausgabe angezeigt wird. Die ID für ein einzelnes Szenario besteht aus dem Präfix, gefolgt von einem Unterstrich und dem Wert des jeweiligen Dimensionsfelds, das die Stammzeitreihe angibt. Die Dimensionswerte sind durch Unterstriche voneinander getrennt. Außer der Eindeutigkeit gibt es keine Einschränkungen für den Wert des Präfixes.

### Ausdruck für Szenariowerte für Stammzeitreihe

Szenariowerte für eine Szenariogruppe werden durch einen Ausdruck angegeben, der anschließend verwendet wird, um die Werte für die einzelnen Stammzeitreihen in der Gruppe zu berechnen. Sie können einen Ausdruck direkt eingeben oder auf die Schaltfläche für den Taschenrechner klicken und den Ausdruck über den Expression Builder für Szenariowerte erstellen.

- Der Ausdruck kann ein beliebiges Ziel oder eine beliebige Eingabe aus dem Modellsystem enthalten.

- Wenn die Szenarioperiode über die vorhandenen Daten hinausgeht, wird der Ausdruck auf vorhergesagte Werte der Felder im Ausdruck angewendet.
- Für die einzelnen Stammzeitreihen in der Gruppe geben die Felder im Ausdruck Zeitreihen an, die durch diese Felder definiert werden, und die Dimensionswerte, die die Stammzeitreihen definieren. Diese Zeitreihen werden zum Auswerten des Ausdrucks verwendet. Wenn z. B. eine Stammzeitreihe durch `region='north'` und `brand='X'` definiert wird, werden die im Ausdruck verwendeten Zeitreihen durch dieselben Dimensionswerte definiert.

Beispiel: Das Stammmetrikfeld ist *advertising* und es gibt die beiden Dimensionsfelder *region* und *brand*. Die Szenariogruppe schließt alle Kombinationen von Dimensionsfeldwerten ein. Sie können dann `advertising*1.2` als Ausdruck angeben, um die Auswirkung einer Erhöhung von *advertising* um 20 % auf alle Zeitreihen zu prüfen, die dem Feld *advertising* zugeordnet sind.

**Anmerkung:** Szenariogruppen werden nur auf mehrdimensionale Daten angewendet. Sie können sie erstellen, indem Sie auf der Registerkarte **Szenarios** auf die Option **Szenario hinzufügen** klicken.

## Optionen

### Maximale Ebene für betroffene Ziele

Gibt die maximale Anzahl der Ebenen für betroffene Ziele an. Bis zu fünf aufeinanderfolgende Ebenen schließen jeweils Ziele ein, auf die sich die Stammzeitreihen eher indirekt auswirken. Die erste Ebene schließt Ziele ein, die die Stammzeitreihe als direkte Eingabe aufweisen. Ziele der zweiten Ebene weisen Ziele der ersten Ebene als direkte Eingaben auf usw. Wenn der Wert für diese Einstellung erhöht wird, steigt auch die Komplexität der Berechnung, was die Leistung beeinträchtigen kann.

### Maximale Anzahl automatisch erkannter Ziele

Gibt die maximale Anzahl der betroffenen Ziele an, die automatisch für jede einzelne Stammzeitreihe berücksichtigt werden. Wenn der Wert für diese Einstellung erhöht wird, steigt auch die Komplexität der Berechnung, was die Leistung beeinträchtigen kann.

### Wirkungsdiagramm

Zeigt eine grafische Darstellung der kausalen Beziehungen zwischen der Stammzeitreihe der einzelnen Szenarios und den Zeitreihen, auf die diese sich auswirkt. Tabellen sowohl zu den Szenariowerten als auch zu den für die betroffenen Ziele vorhergesagten Werten sind in das Ausgabeelement eingeschlossen. Die Grafik enthält Diagramme der vorhergesagten Werte für die betroffenen Ziele. Durch Einmalklicken auf einen beliebigen Knoten im Wirkungsdiagramm wird ein detailliertes Ablaufdiagramm für die Zeitreihe geöffnet, die dem Knoten zugeordnet ist. Für jedes Szenario wird ein separates Wirkungsdiagramm generiert.

### Zeitreihendiagramme

Generiert Zeitreihendiagramme mit den vorhergesagten Werten für jedes betroffene Ziel in jedem der Szenarios.

### Vorhersage- und Szenariotabellen

Tabellen mit vorhergesagten Werten und Szenariowerten für jedes der Szenarios. Diese Tabellen enthalten dieselben Informationen wie die Tabellen im Wirkungsdiagramm. Sie unterstützen alle Standardfunktionen für das Drehen und Bearbeiten von Tabellen.

### Konfidenzintervalle in Diagramme und Tabellen einschließen

Gibt an, ob Konfidenzintervalle für Szenariovorhersagen sowohl in die Diagramm- als auch in die Tabellenausgabe eingeschlossen werden sollen.

### Konfidenzintervallbreite (%)

Diese Einstellung steuert die Konfidenzintervalle für die Szenariovorhersagen. Sie können einen beliebigen positiven Wert kleiner als 100 angeben. Standardmäßig wird ein Konfidenzintervall von 95 % verwendet.



---

## Kapitel 8. Maße für die Anpassungsgüte

In diesem Abschnitt finden Sie Definitionen der Maße für die Anpassungsgüte, die bei der Zeitreihenmodellierung verwendet werden.

- *R-Quadrat für stationären Teil*. Ein Maß, das den stationären Teil des Modells mit einem einfachen Mittelwertmodell vergleicht. Dieses Maß ist dem gewöhnlichen R-Quadrat vorzuziehen, wenn ein Trend oder ein saisonales Muster vorliegt. R-Quadrat für den stationären Teil kann auch negativ sein, es nimmt Werte zwischen minus unendlich und 1 an. Negative Werte bedeuten, dass das betrachtete Modell schlechter ist als das Basismodell. Positive Werte bedeuten, dass das betrachtete Modell besser ist als das Basismodell.
- *R-Quadrat*. Eine Schätzung für den Anteil der Gesamtvariation der Zeitreihe, der durch das Modell erklärt wird. Dieses Maß ist sehr nützlich, wenn die Zeitreihe stationär ist. R-Quadrat kann auch negativ sein, es nimmt Werte zwischen minus unendlich und 1 an. Negative Werte bedeuten, dass das betrachtete Modell schlechter ist als das Basismodell. Positive Werte bedeuten, dass das betrachtete Modell besser ist als das Basismodell.
- *RMSE*. Steht für Root Mean Square Error. Die Quadratwurzel des mittleren quadratischen Fehlers. Ein Maß dafür, wie stark eine abhängige Zeitreihe von ihrem durch das Modell vorhergesagten Niveau abweicht, und zwar ausgedrückt in derselben Maßeinheit wie die abhängige Zeitreihe.
- *MAPE*. Mittlerer absoluter Fehler in Prozent. Ein Maß dafür, wie stark eine abhängige Zeitreihe von ihrem durch das Modell vorhergesagten Niveau abweicht. Es ist unabhängig von den verwendeten Maßeinheiten und kann daher verwendet werden, um Zeitreihen mit unterschiedlichen Einheiten zu vergleichen.
- *MAE*. Mittlerer absoluter Fehler. Er misst, wie stark die Zeitreihe von ihrem durch das Modell vorhergesagten Niveau abweicht. MAE wird in derselben Maßeinheit angegeben wie die ursprüngliche Zeitreihe.
- *MaxAPE*. Maximaler absoluter Fehler in Prozent (Maximum Absolute Percentage Error, also maximaler Betrag des relativen Fehlers). Dies ist der größte vorhergesagte Fehler, ausgedrückt in Prozent. Dieses Maß hilft dabei, sich ein Worst-Case-Szenario für die Vorhersagen vorzustellen.
- *MaxAE*. Maximaler absoluter Fehler (Maximum Absolute Error, also maximaler Betrag des Fehlers). Dies ist der größte vorhergesagte Fehler, ausgedrückt in derselben Maßeinheit wie die abhängige Zeitreihe. Genau wie MaxAPE hilft er dabei, sich ein Worst-Case-Szenario für die Vorhersagen vorzustellen. Der maximale absolute Fehler und der maximale absolute Fehler in Prozent können an verschiedenen Punkten in der Zeitreihe auftreten, beispielsweise wenn der absolute Fehler für einen großen Zeitreihenwert geringfügig größer ist als der absolute Fehler für einen kleinen Zeitreihenwert. In diesem Fall tritt der maximale absolute Fehler beim größeren Zeitreihenwert und der maximale absolute Fehler in Prozent beim kleineren Zeitreihenwert auf.
- *Normalisiertes BIC*. Normalisiertes Bayes-Informationskriterium (BIC). Ein allgemeines Maß der insgesamt erreichten Güte der Anpassung, das auch die Komplexität des Modells zu berücksichtigen versucht. Es ist ein Score, der auf dem mittleren quadratischen Fehler beruht und eine Penalisierung für die Anzahl der Modellparameter und die Länge der Zeitreihe enthält. Die Penalisierung neutralisiert die Überlegenheit von Modellen mit einer größeren Anzahl von Parametern und macht die Statistik damit gut vergleichbar für verschiedene Modelle derselben Zeitreihe.



---

## Kapitel 9. Ausreißertypen

In diesem Abschnitt finden Sie Definitionen der Ausreißertypen, die bei der Zeitreihenmodellierung verwendet werden.

- *Additiv*. Ein Ausreißer, der eine einzelne Beobachtung beeinflusst. So könnte z. B. ein Fehler bei der Datencodierung als additiver Ausreißer identifiziert werden.
- *Niveauverschiebung*. Ein Ausreißer, der beginnend an einer bestimmten Stelle der Zeitreihe alle Beobachtungen um eine Konstante verschiebt. Eine Niveauverschiebung könnte sich aus einem Strategiewechsel ergeben.
- *Innovativ*. Ein Ausreißer, der als Zuwachs zum Rauschen an einer bestimmten Stelle der Zeitreihe wirkt. Bei stationären Zeitreihen beeinflusst ein innovativer Ausreißer mehrere Beobachtungen. Bei nicht stationären Zeitreihen kann er alle Beobachtungen ab einer bestimmten Stelle der Zeitreihe beeinflussen.
- *Transient*. Ein Ausreißer, dessen Auswirkung exponentiell gegen null strebt.
- *Saisonal additiv*. Ein Ausreißer, der eine bestimmte Beobachtung sowie alle folgenden Beobachtungen beeinflusst, die von dieser mindestens eine saisonale Periode entfernt sind. Alle diese Beobachtungen werden in gleicher Weise beeinflusst. Ein saisonal additiver Ausreißer könnte auftreten, wenn ab einem bestimmten Jahr die Umsätze jeden Januar höher sind.
- *Lokaler Trend*. Ein Ausreißer, der an einer bestimmten Stelle der Zeitreihe einen lokalen Trend auslöst.
- *Additiver Bereich*. Eine Gruppe von zwei oder mehr aufeinanderfolgenden additiven Ausreißern. Wenn dieser Typ von Ausreißern gewählt wird, werden neben Gruppen von Ausreißern auch einzelne additive Ausreißer entdeckt.



---

## Kapitel 10. Übersicht über ACF-/PACF-Diagramme

Die hier gezeigten Diagramme gehören zu reinen oder theoretischen ARIMA-Prozessen. Hier einige allgemeine Richtlinien zur Identifizierung des Prozesses:

- Nicht stationäre Zeitreihen weisen eine ACF auf, die über mindestens ein halbes Dutzend Lags signifikant bleibt, anstatt sich schnell dem Wert 0 anzunähern. Sie müssen eine solche Zeitreihe vor dem Identifizieren des Prozesses differenzieren, bis sie stationär ist.
- Autoregressive Prozesse weisen eine exponentiell abnehmende ACF und Spikes im ersten Lag (oder den ersten Lags) der PACF auf. Die Anzahl der Spikes gibt die Ordnung der Autoregression an.
- Prozesse mit gleitendem Durchschnitt Verfahren weisen Spikes im ersten Lag (oder den ersten Lags) der ACF und eine exponentiell abnehmende PACF auf. Die Anzahl der Spikes gibt die Ordnung des gleitenden Durchschnitts an.
- Gemischte (ARMA-)Prozesse weisen normalerweise eine exponentielle Abnahme sowohl bei der ACF als auch bei der PACF auf.

In der Identifizierungsphase brauchen Sie sich nicht um das Vorzeichen der ACF bzw. PACF oder um die Geschwindigkeit zu kümmern, mit der eine exponentiell abnehmende ACF bzw. PACF sich dem Wert 0 annähert. Diese hängt vom Vorzeichen und dem tatsächlichen Wert der AR- und MA-Koeffizienten ab. In einigen Fällen wechselt eine exponentiell abnehmende ACF zwischen positiven und negativen Werten.

*ACF- und PACF-Diagramme aus echten Daten sind nie so sauber wie die hier gezeigten Diagramme. Sie müssen lernen, aus dem jeweiligen Diagramm das Wesentliche herauszusuchen. Überprüfen Sie stets die ACF und PACF der Residuen, für den Fall, dass Ihre Identifizierung falsch ist. Beachten Sie:*

- Saisonale Prozesse zeigen diese Muster in den saisonalen Lags (den Mehrfachen der Saison-Periode).
- Sie dürfen nicht signifikante Werte als 0 behandeln, d. h., Sie können Werte ignorieren, die im Diagramm innerhalb der Konfidenzintervalle liegen. Sie müssen sie jedoch nicht ignorieren, insbesondere, wenn Sie das Muster der statistisch signifikanten Werte fortsetzen.
- Einzelne Autokorrelationen sind schon durch Zufall statistisch signifikant. Sie können eine statistisch signifikante Autokorrelation ignorieren, wenn sie isoliert vorkommt, insbesondere in hohen Lags, und wenn sie nicht in einem saisonalen Lag auftritt.

Detailliertere Erläuterungen zu ACF- und PACF-Diagrammen finden Sie in jedem beliebigen Text zur ARIMA-Analyse.

Tabelle 3. ARIMA(0,0,1),  $q>0$

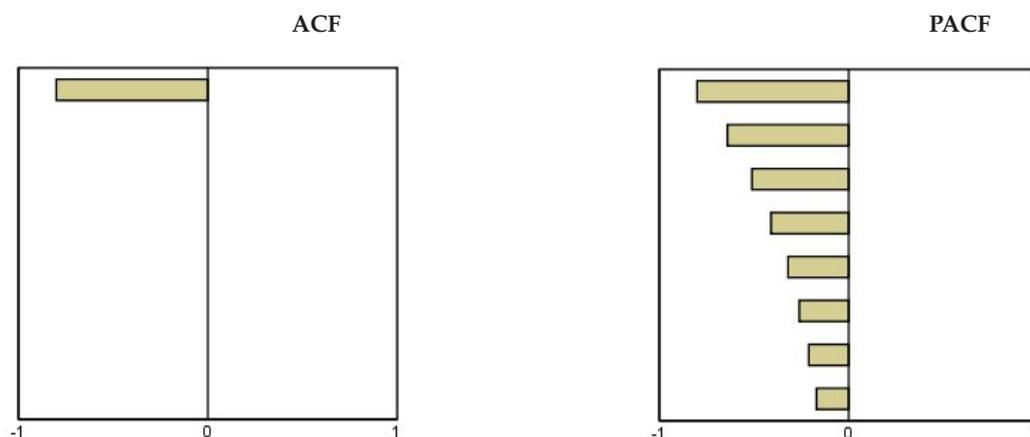
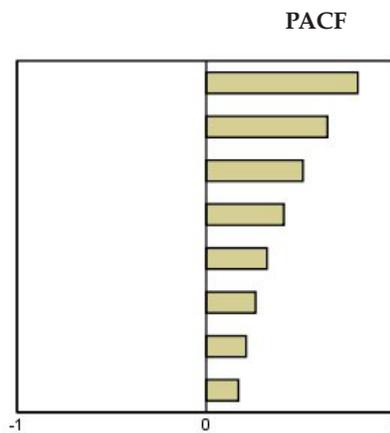
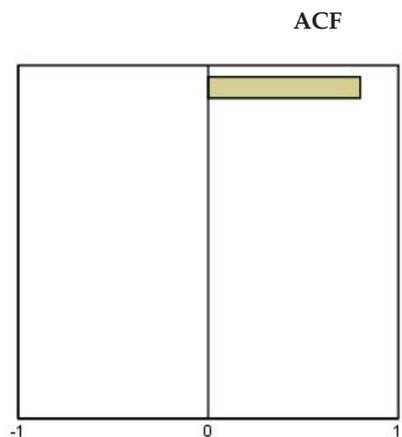


Tabelle 4. ARIMA(0,0,1),  $q < 0$



ARIMA(0,0,2),  $\rho_1 > 0$

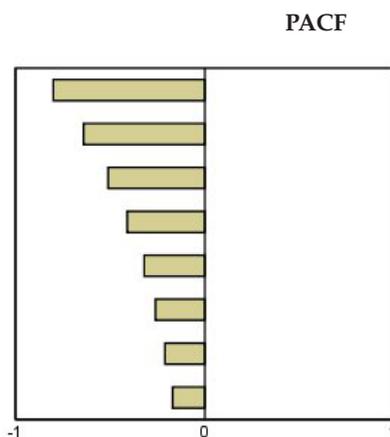
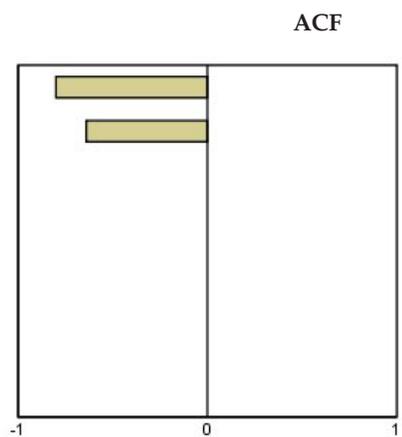


Tabelle 5. ARIMA(1,0,0),  $f > 0$

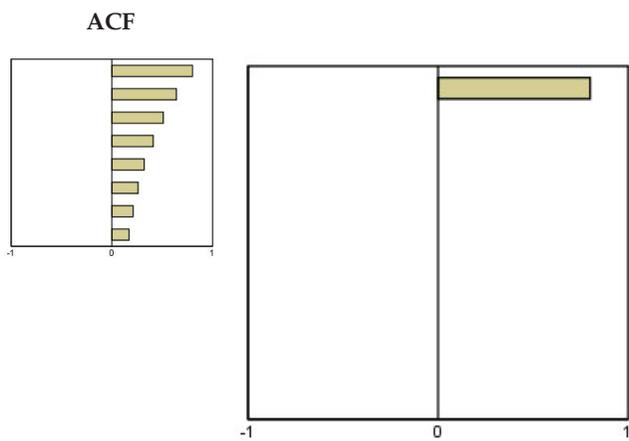
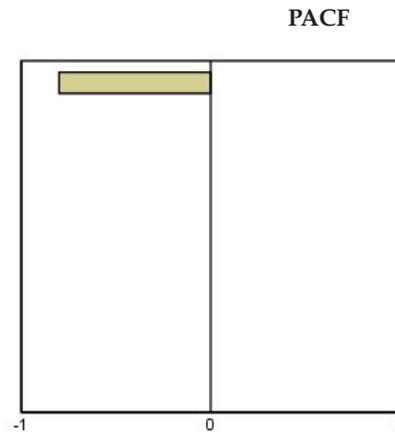
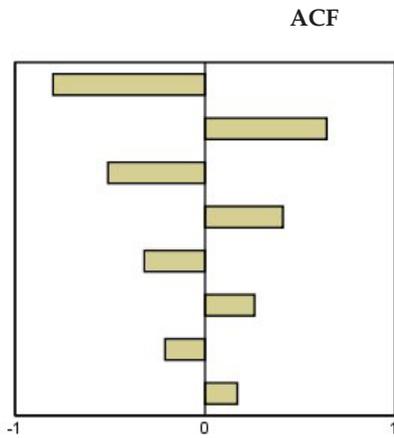
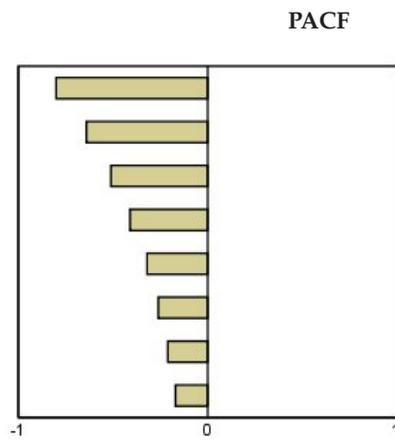
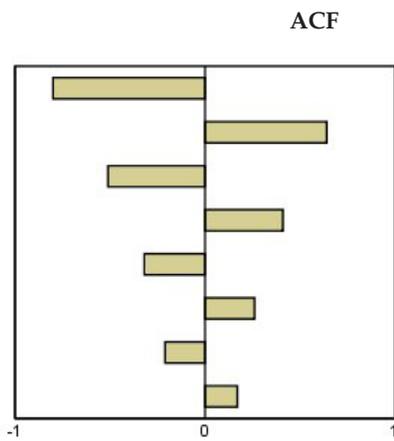


Tabelle 6. ARIMA(1,0,0),  $f < 0$



ARIMA(1,0,1),  $<0, >0$



ARIMA(2,0,0),  $1_2 > 0$

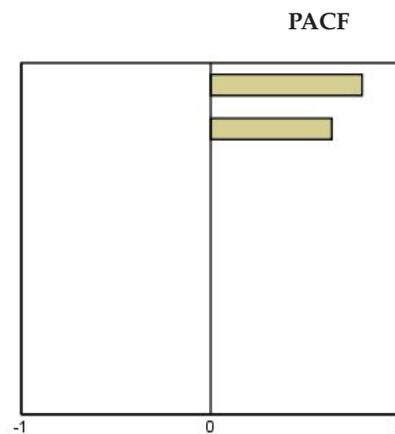
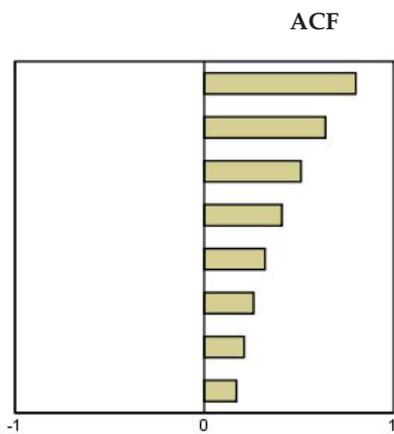
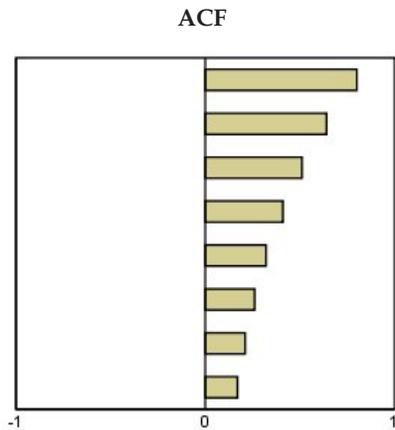


Tabelle 7. ARIMA(0,1,0) (integrierte Zeitreihe)



---

## Bemerkungen

Die vorliegenden Informationen wurden für Produkte und Services entwickelt, die auf dem deutschen Markt angeboten werden.

Möglicherweise bietet IBM die in dieser Dokumentation beschriebenen Produkte, Services oder Funktionen in anderen Ländern nicht an. Informationen über die gegenwärtig im jeweiligen Land verfügbaren Produkte und Services sind beim zuständigen IBM Ansprechpartner erhältlich. Hinweise auf IBM Lizenzprogramme oder andere IBM Produkte bedeuten nicht, dass nur Programme, Produkte oder Services von IBM verwendet werden können. Anstelle der IBM Produkte, Programme oder Services können auch andere, ihnen äquivalente Produkte, Programme oder Services verwendet werden, solange diese keine gewerblichen oder anderen Schutzrechte von IBM verletzen. Die Verantwortung für den Betrieb von Produkten, Programmen und Services anderer Anbieter liegt beim Kunden.

Für in diesem Handbuch beschriebene Erzeugnisse und Verfahren kann es IBM Patente oder Patentanmeldungen geben. Mit der Auslieferung dieses Handbuchs ist keine Lizenzierung dieser Patente verbunden. Lizenzanforderungen sind schriftlich an folgende Adresse zu richten (Anfragen an diese Adresse müssen auf Englisch formuliert werden):

IBM Director of Licensing  
IBM Europe, Middle East & Africa  
Tour Descartes  
2, avenue Gambetta  
92066 Paris La Defense  
France

Verweise in diesen Informationen auf Websites anderer Anbieter werden lediglich als Service für den Kunden bereitgestellt und stellen keinerlei Billigung des Inhalts dieser Websites dar. Das über diese Websites verfügbare Material ist nicht Bestandteil des Materials für dieses IBM Produkt. Die Verwendung dieser Websites geschieht auf eigene Verantwortung.

Werden an IBM Informationen eingesandt, können diese beliebig verwendet werden, ohne dass eine Verpflichtung gegenüber dem Einsender entsteht.

Lizenznehmer des Programms, die Informationen zu diesem Produkt wünschen mit der Zielsetzung: (i) den Austausch von Informationen zwischen unabhängig voneinander erstellten Programmen und anderen Programmen (einschließlich des vorliegenden Programms) sowie (ii) die gemeinsame Nutzung der ausgetauschten Informationen zu ermöglichen, wenden sich an folgende Adresse:

IBM Software Group  
ATTN: Licensing  
200 W. Madison St.  
Chicago, IL; 60606  
USA

Die Bereitstellung dieser Informationen kann unter Umständen von bestimmten Bedingungen - in einigen Fällen auch von der Zahlung einer Gebühr - abhängig sein.

Die Lieferung des in diesem Dokument beschriebenen Lizenzprogramms sowie des zugehörigen Lizenzmaterials erfolgt auf der Basis der IBM Rahmenvereinbarung bzw. der Allgemeinen Geschäftsbedingungen von IBM, der IBM Internationalen Nutzungsbedingungen für Programmpakete oder einer äquivalenten Vereinbarung.

Alle in diesem Dokument enthaltenen Leistungsdaten stammen aus einer kontrollierten Umgebung. Die Ergebnisse, die in anderen Betriebsumgebungen erzielt werden, können daher erheblich von den hier erzielten Ergebnissen abweichen. Einige Daten stammen möglicherweise von Systemen, deren Entwicklung noch nicht abgeschlossen ist. Eine Gewährleistung, dass diese Daten auch in allgemein verfügbaren Systemen erzielt werden, kann nicht gegeben werden. Darüber hinaus wurden einige Daten unter Umständen durch Extrapolation berechnet. Die tatsächlichen Ergebnisse können davon abweichen. Benutzer dieses Dokuments sollten die entsprechenden Daten in ihrer spezifischen Umgebung prüfen.

Alle Informationen zu Produkten anderer Anbieter stammen von den Anbietern der aufgeführten Produkte, deren veröffentlichten Ankündigungen oder anderen allgemein verfügbaren Quellen. IBM hat diese Produkte nicht getestet und kann daher keine Aussagen zu Leistung, Kompatibilität oder anderen Merkmalen machen. Fragen zu den Leistungsmerkmalen von Produkten anderer Anbieter sind an den jeweiligen Anbieter zu richten.

Aussagen über Pläne und Absichten von IBM unterliegen Änderungen oder können zurückgenommen werden und repräsentieren nur die Ziele von IBM.

Diese Veröffentlichung enthält Beispiele für Daten und Berichte des alltäglichen Geschäftsablaufs. Sie sollen nur die Funktionen des Lizenzprogramms illustrieren und können Namen von Personen, Firmen, Marken oder Produkten enthalten. Alle diese Namen sind frei erfunden; Ähnlichkeiten mit tatsächlichen Namen und Adressen sind rein zufällig.

#### COPYRIGHTLIZENZ:

Diese Veröffentlichung enthält Beispielanwendungsprogramme, die in Quellsprache geschrieben sind und Programmier Techniken in verschiedenen Betriebsumgebungen veranschaulichen. Sie dürfen diese Beispielprogramme kostenlos kopieren, ändern und verteilen, wenn dies zu dem Zweck geschieht, Anwendungsprogramme zu entwickeln, zu verwenden, zu vermarkten oder zu verteilen, die mit der Anwendungsprogrammierschnittstelle für die Betriebsumgebung konform sind, für die diese Beispielprogramme geschrieben werden. Diese Beispiele wurden nicht unter allen denkbaren Bedingungen getestet. Daher kann IBM die Zuverlässigkeit, Wartungsfreundlichkeit oder Funktion dieser Programme weder zusagen noch gewährleisten. Die Beispielprogramme werden ohne Wartung (auf "as-is"-Basis) und ohne jegliche Gewährleistung zur Verfügung gestellt. IBM übernimmt keine Haftung für Schäden, die durch die Verwendung der Beispielprogramme entstehen.

Kopien oder Teile der Beispielprogramme bzw. daraus abgeleiteter Code müssen folgenden Copyrightvermerk beinhalten:

© (Name Ihrer Firma) (Jahr). Teile des vorliegenden Codes wurden aus Beispielprogrammen der IBM Corporation abgeleitet.

© Copyright IBM Corp. \_Jahr/Jahre angeben\_. Alle Rechte vorbehalten.

---

## Marken

IBM, das IBM Logo und [ibm.com](http://ibm.com) sind Marken oder eingetragene Marken der IBM Corp in den USA und/oder anderen Ländern. Weitere Produkt- und Servicennamen können Marken von IBM oder anderen Unternehmen sein. Eine aktuelle Liste der IBM Marken finden Sie auf der Webseite „Copyright and trademark information“ unter [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

Adobe, das Adobe-Logo, PostScript und das PostScript-Logo sind Marken oder eingetragene Marken der Adobe Systems Incorporated in den USA und/oder anderen Ländern.

Intel, das Intel-Logo, Intel Inside, das Intel Inside-Logo, Intel Centrino, das Intel Centrino-Logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium und Pentium sind Marken oder eingetragene Marken der Intel Corporation oder der Tochtergesellschaften des Unternehmens in den USA und anderen Ländern.

Linux ist eine eingetragene Marke von Linus Torvalds in den USA, anderen Ländern oder beidem.

Microsoft, Windows, Windows NT und das Windows-Logo sind Marken der Microsoft Corporation in den USA und/oder anderen Ländern.

UNIX ist eine eingetragene Marke von The Open Group in den USA und anderen Ländern.

Java und alle auf Java basierenden Marken und Logos sind Marken oder eingetragene Marken der Oracle Corporation und/oder ihrer verbundenen Unternehmen.



---

# Index

## A

- ACF
  - Diagramme für reine ARIMA-Prozesse 63
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13
- Additiver Ausreißer 61
- in Zeitreihenmodellierung 8, 11
- Anpassungsgüte
  - Definitionen 59
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13
- Anpassungswerte
  - in "Zeitreihenmodell zuweisen" 20
  - in Zeitreihenmodellierung 13
- ARIMA-Modelle 5
- Ausreißer 11
- Transferfunktionen 10
- Ausreißer
  - ARIMA-Modelle 11
  - Definitionen 61
  - Expert Modeler 8
- Ausreißer mit additivem Bereich 61
- in Zeitreihenmodellierung 8, 11
- Ausreißer mit lokalem Trend 61
- in Zeitreihenmodellierung 8, 11
- Autokorrelationsfunktion
  - Diagramme für reine ARIMA-Prozesse 63
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13

## B

- Box-Ljung-Statistik
  - in "Zeitreihenmodell zuweisen" 19
  - in Zeitreihenmodellierung 12

## E

- Einfaches Modell mit exponentiellem Glätten 8
- Einfaches saisonales Modell mit exponentiellem Glätten 8
- Ereignisse 7
- in Zeitreihenmodellierung 7
- Expert Modeler 5
- Ausreißer 8
- Modellraum einschränken 7

## F

- Fehlende Werte
  - in "Zeitreihenmodell zuweisen" 22
  - in Zeitreihenmodellierung 15

## G

- Gedämpftes Modell mit exponentiellem Glätten 8

## H

- Harmonische Analyse 27
- Historische Daten
  - in "Zeitreihenmodell zuweisen" 20
  - in Zeitreihenmodellierung 13
- Historische Periode 2
- Holdout-Fälle 2

## I

- Im Niveau verschobener Ausreißer 61
- in Zeitreihenmodellierung 8, 11
- Innovativer Ausreißer 61
- in Zeitreihenmodellierung 8, 11

## K

- Konfidenzintervalle
  - in "Zeitreihenmodell zuweisen" 20, 22
  - in Zeitreihenmodellierung 13, 15

## L

- Log-Transformation
  - in Zeitreihenmodellierung 8, 9, 10

## M

- MAE 59
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13
- MAPE 59
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13
- MaxAE 59
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13
- MaxAPE 59
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13
- Maximaler absoluter Fehler 59
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13
- Maximaler absoluter Fehler in Prozent 59
  - in "Zeitreihenmodell zuweisen" 19, 20

- Maximaler absoluter Fehler in Prozent (*Forts.*)
  - in Zeitreihenmodellierung 12, 13
- Mittlerer absoluter Fehler 59
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13
- Mittlerer absoluter Fehler in Prozent 59
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13
- Modell mit exponentiellem Glätten nach Brown 8
- Modell mit exponentiellem Glätten nach Holt 8
- Modell mit exponentiellem Glätten nach Winters
  - additiv 8
  - multiplikativ 8
- Modelle
  - ARIMA (X11 ARIMA) 5
  - Expert Modeler 5
  - exponentielles Glätten 5, 8
- Modelle mit exponentiellem Glätten 5, 8
- Modellnamen
  - in Zeitreihenmodellierung 15
- Modellparameter
  - in "Zeitreihenmodell zuweisen" 19
  - in Zeitreihenmodellierung 12
- Modellparameter erneut schätzen
  - in "Zeitreihenmodell zuweisen" 17

## N

- Normalisiertes BIC (bayessches Informationskriterium) 59
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13

## P

- PACF
  - Diagramme für reine ARIMA-Prozesse 63
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13
- Partielle Autokorrelationsfunktion
  - Diagramme für reine ARIMA-Prozesse 63
  - in "Zeitreihenmodell zuweisen" 19, 20
  - in Zeitreihenmodellierung 12, 13
- Periodizität
  - in Zeitreihenmodellierung 7, 8, 9, 10

## Q

Quadratwurzeltransformation  
in Zeitreihenmodellierung 8, 9, 10

## R

R2 59  
in "Zeitreihenmodell zuweisen" 19, 20  
in Zeitreihenmodellierung 12, 13  
Residuen  
in "Zeitreihenmodell zuweisen" 19, 20  
in Zeitreihenmodellierung 12, 13  
RMSE 59  
in "Zeitreihenmodell zuweisen" 19, 20  
in Zeitreihenmodellierung 12, 13

## S

Saisonal additiver Ausreißer 61  
in Zeitreihenmodellierung 8, 11  
Saisonale Zerlegung 25, 26  
Annahmen 25  
Berechnen der gleitenden Durchschnitts-  
schnitts 25  
Erstellen von Variablen 26  
Modelle 25  
Speichern von neuen Variablen 26  
Schätzperiode 2  
in "Zeitreihenmodell zuweisen" 17  
in Zeitreihenmodellierung 5  
Speichern  
erneut geschätzte Modelle als  
XML 21  
Modellspezifikationen als XML 14  
Modellvorhersagen 14, 21  
neue Variablennamen 14, 21  
Spektraldiagramme 27, 29  
Annahmen 27  
bivariate Spektralanalyse 27  
Spektralfenster 27  
Zentriertransformation 27  
Stationäres R-Quadrat 59  
in "Zeitreihenmodell zuweisen" 19, 20  
in Zeitreihenmodellierung 12, 13

## T

Temporale kausale Modelle 31, 32, 34,  
35, 36, 37, 39, 41, 42  
Temporale kausale Modellszenarios 51,  
52, 53, 55, 56, 57  
Transferfunktionen 10  
Nennerterme 10  
Ordnung der Differenzen 10  
saisonale Ordnungen 10  
Verzögerung 10  
Zählerterme 10  
Transformation mit natürlichem Logarith-  
mus  
in Zeitreihenmodellierung 8, 9, 10  
Transienter Ausreißer 61

Transienter Ausreißer (*Forts.*)  
in Zeitreihenmodellierung 8, 11

## V

Validierungsperiode 2  
Variablennamen  
in "Zeitreihenmodell zuweisen" 21  
in Zeitreihenmodellierung 14  
Vorhersage temporaler kausaler Model-  
le 45, 46, 47, 48, 50  
Vorhersagen  
in "Zeitreihenmodell zuweisen" 19,  
20  
in Zeitreihenmodellierung 12, 13  
Vorhersageperiode  
in "Zeitreihenmodell zuweisen" 17  
in Zeitreihenmodellierung 5, 15

## W

Wurzel der mittleren Fehlerquadratsum-  
me 59  
in "Zeitreihenmodell zuweisen" 19,  
20  
in Zeitreihenmodellierung 12, 13

## X

XML  
erneut geschätzte Modelle als XML  
speichern 21  
Zeitreihenmodelle als XML spei-  
chern 14

## Z

Zeitreihenanalyse  
Temporale kausale Modelle 31  
Zeitreihenmodell zuweisen  
Box-Ljung-Statistik 19  
Modellparameter 19  
partielle Residuen-Autokorrelations-  
funktion 19  
Residuen-Autokorrelationsfunktio-  
nen 19  
Statistik für alle Modelle 19  
Statistik für Anpassungsgüte 19  
Vorhersagen 19  
Zeitreihenmodellierung 5  
am besten/schlechtesten angepasste  
Modelle 14  
Anpassungswerte 13  
ARIMA (X11 ARIMA) 5, 9  
Ausreißer 8, 11  
Box-Ljung-Statistik 12  
Ereignisse 7  
Expert Modeler 5  
exponentielles Glätten 5, 8  
fehlende Werte 15  
Konfidenzintervalle 13, 15  
Modellnamen 15  
Modellparameter 12  
Modellspezifikationen als XML spei-  
chern 14

Zeitreihenmodellierung (*Forts.*)  
neue Variablennamen 14  
partielle Residuen-Autokorrelations-  
funktion 12, 13  
Periodizität 7, 8, 9, 10  
Residuen-Autokorrelationsfunktio-  
nen 12, 13  
Schätzperiode 5  
Statistik für alle Modelle 12, 13  
Statistik für Anpassungsgüte 12, 13  
Transferfunktionen 10  
Vorhersagen 12, 13  
Vorhersagen speichern 14  
Vorhersageperiode 5, 15  
Zeitreihentransformation 8, 9, 10  
Zuweisen von Zeitreihenmodellen 17  
am besten/schlechtesten angepasste  
Modelle 21  
Anpassungswerte 20  
erneut geschätzte Modelle als XML  
speichern 21  
fehlende Werte 22  
Konfidenzintervalle 20, 22  
Modellparameter erneut schätzen 17  
neue Variablennamen 21  
partielle Residuen-Autokorrelations-  
funktion 20  
Residuen-Autokorrelationsfunktio-  
nen 20  
Schätzperiode 17  
Statistik für alle Modelle 20  
Statistik für Anpassungsgüte 20  
Vorhersagen 20  
Vorhersagen speichern 21  
Vorhersageperiode 17



