

شرح منوی
Apply Time Series Models
در نرم افزار آماری
SPSS

تهیه:

نیما سلطانی محمدی
کارشناس آمار

Apply Time Series Models

به کار بردن مدل های سری زمانی

رویه Apply Time Series Models مدل های سری زمانی موجود را از یک فایل خارجی بارگذاری کرده و آن ها را بر داده های فعال بکار می برد. شما می توانید این رویه را برای بدست آوردن پیش بینی هایی برای سری هایی برای داده های جدید و یا اصلاح شده موجود استفاده کنید بدون آنکه مدل های خود را دوباره سازی کنید. مدل های تولید شده از رویه Time Series Modeler استفاده می کنند.

مثال: شما یک مدیر اموال هستید با یک خرده فروش بزرگ و پاسخگو برای همه ۵۰۰۰ کالا. شما از Expert Modeler برای ساخت مدل هایی که فروش هر کالا در سه ماه آینده را پیش بینی می کنند استفاده کرده اید. داده های انبار شما هر ماه با داده های فروش واقعی که مایلید برای پیش بینی های به روز شده ماهانه استفاده کنید تغییر می کند. رویه Apply Time Series Models به شما اجازه می دهد از مدل های اصلی استفاده کنید و به سادگی پارامترهای مدل را برای محاسبه داده های جدید برآورد مجدد کنید.

آماره ها: مقیاس های نیکویی برازش: مربع R مانا، مربع R، خطای مربع میانگین ریشه (RMSE)، خطای مطلق میانگین (MAE)، خطای درصد مطلق میانگین (MAPE)، ماکسیم خطای مطلق (MaxAE)، ماکسیم خطای درصد مطلق (MaxAPE)، معیار اطلاعات بیزی نرمال شده (BIC)، باقیمانده ها: تابع خود همبستگی، تابع خود همبستگی جزئی، Ljung-Box Q

نمودارها: نمودارهای خلاصه در تمام مدل ها: هیستوگرام مربع R مانا، مربع R، خطای مربع میانگین ریشه (RMSE)، خطای مطلق میانگین (MAE)، خطای درصد مطلق میانگین (MAPE)، ماکسیم خطای مطلق (MaxAE)، ماکسیم خطای درصد مطلق (MaxAPE)، معیار اطلاعات بیزی نرمال شده (BIC)، نمودارهای جعبه ای باقیمانده خودهمبستگی و خودهمبستگی جزئی، نتایج برای مدل های منحصر بفرد: مقادیر پیش بینی، مقادیر برازش، مقادیر مشاهده شده، حدود اطمینان بالا و پایین، خودهمبستگی و خودهمبستگی جزئی باقیمانده ها.

ملاحظات داده ها در Apply Time Series Models

داده ها: متغیرهایی (وابسته و مستقل) که در مدل به کار می روند باید عددی باشند.

مفروضات: متغیرهایی که در دیتاست فعال بکار می رود با متغیرهای مشخص شده در مدل هم نام هستند. همه متغیرها به عنوان سری های زمانی مورد استفاده هستند، یعنی هر حالت یک نقطه زمانی را نمایش می دهد با حالت های متوالی جدا از هم با یک بازه زمانی ثابت.

پیش بینی: برای ساخت پیش بینی ها از مدل هایی با متغیرهای مستقل (پیشگو) استفاده می شود که دیتاست فعال باید شامل مقادیری از این متغیرها برای تمام حالات در دوره پیش بینی باشد. اگر پارامترهای مدل برآورد مجدد شده باشند، پس متغیرهای مستقل نباید شامل مقادیر گم شده در دوره برآورد باشند.

تعریف زمان

رویه Apply Time Series Models نیازمند تناوب می باشد، اگر هر کدام از دیتاست های فعال با تناوب مدل ها هماهنگ شود بکار برده می شود. اگر شما بطور ساده از همان دیتاست برای پیش بینی استفاده می کنید که برای ساخت مدل استفاده شده است، پس این وضعیت قابل قبول است. اگر هیچ تناوبی برای دیتاست فعال موجود نباشد شما فرصت خواهید داشت که از جعبه محاوره ای Define Dates برای ساخت یک دوره تناوب استفاده کنید. اگر چه، اگر یک مدل بدون نیاز به دوره تناوب ساخته شده باشد پس دیتاست فعال باید بدون آن باشد.

برای اجرای مدل ها

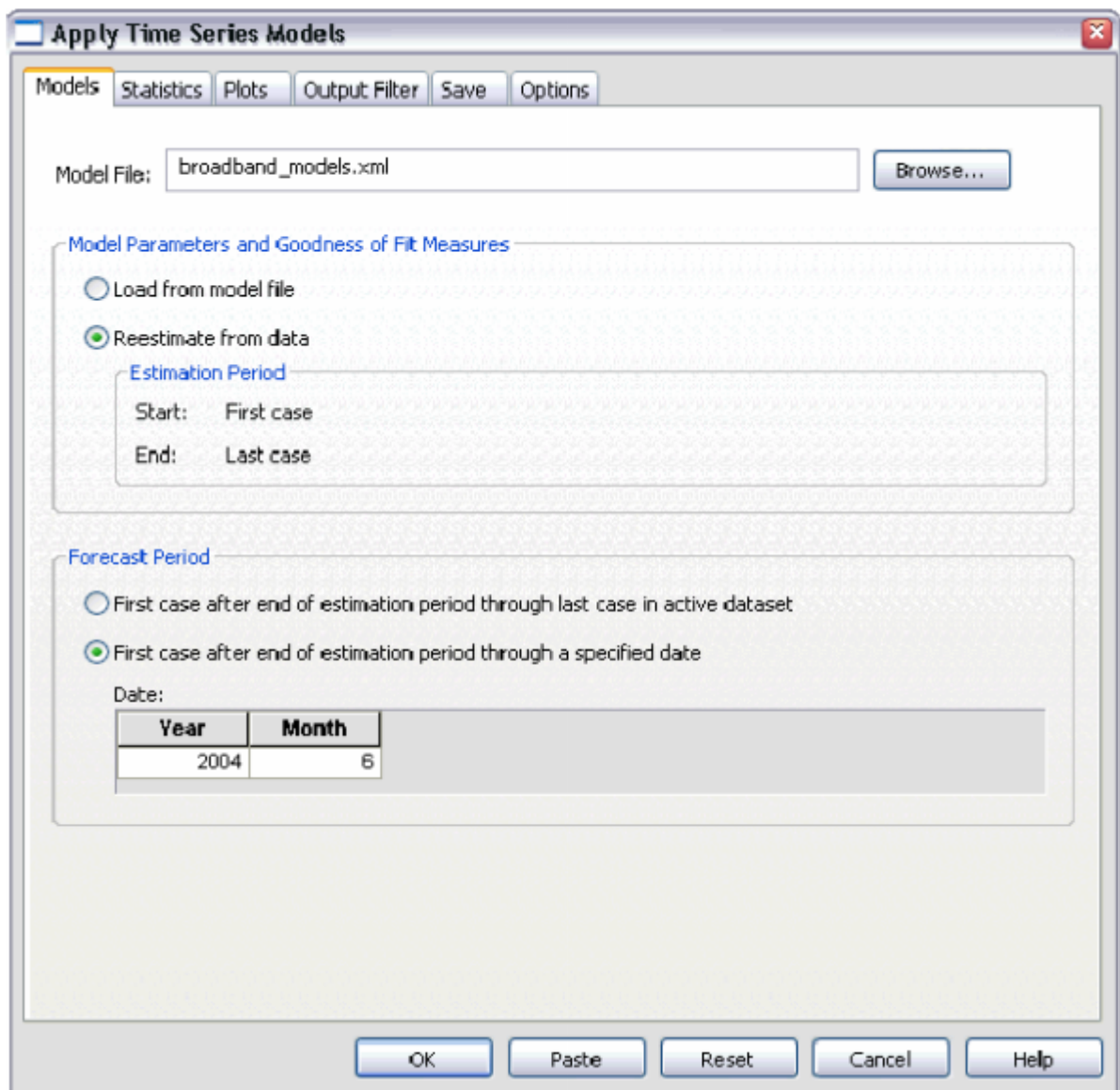
از منو انتخاب کنید:

Analyze

Forecasting

Apply Models . . .

Models، سربرگ Apply Time Series Models



مشخصات فایل مورد نظر را وارد کنید و یا با کلیک بر روی Browse یک فایل مدل انتخاب کنید. (فایل های مدل به وسیله رویه Time Series Modeler ساخته شده اند.)

بطور اختیاری می توانید:

- با استفاده از داده های موجود در دیتاست فعال پارامترهای مدل را برآورد مجدد نمایید.
- پیش بینی ها برای برآورد مجدد پارامترها ساخته شده اند.
- ذخیره پیشگویی ها، بازه های اطمینان و مانده های نوفه (noise residual)
- ذخیره مدل های برآورد مجدد شده در قالب XLM

پارامترهای مدل و اندازه های نیکویی برازش

بارگذاری از فایل مدل: پیش بینی ها با استفاده از پارامترهای موجود در فایل مدل و بدون برآورد مجدد آن پارامترها ساخته شده اند. اندازه های نیکویی برازش در خروجی نشان داده شده اند و برای فیلتر کردن مدل ها (بهترین یا ضعیف ترین برازش) استفاده شده اند که از فایل مدل گرفته شده و داده های استفاده شده را وقتی هر مدل توسعه داده شده یا بروز شده منعکس می کند. با این گزینه پیش بینی ها در داده های تاریخی _چه متغیرهای وابسته و یا مستقل_ در دیتاست فعال به حساب نمی آیند. بعلاوه پیش بینی ها در مقادیر سری های وابسته در دوره پیش بینی به حساب نمی آیند اما در مقادیر متغیرهای مستقل در دوره پیش بینی به حساب می آیند. اگر شما مقادیر بیشتری از سری های وابسته دارید و می خواهیم آنها را در پیش بینی وارد کنید، نیاز به برآورد مجدد، تعدیل دوره برآورد برای وارد کردن آنها دارید.

برآورد مجدد از داده ها: پارامترهای مدل با استفاده از داده های موجود در دیتاست فعال برآورد مجدد می شوند. برآورد مجدد پارامترهای مدل هیچ اثری بر ساختار مدل ندارد. بطور مثال، یک مدل $ARIMA(1,0,1)$ باقی می ماند اما پارامترهای اتورگرسیو و میانگین متحرک برآورد مجدد خواهند شد. برآورد مجدد در کشف مقادیر دورافتاده جدید نتیجه نمی دهد. مقادیر دورافتاده همیشه از فایل مدل گرفته می شوند.

• **دوره برآورد.** دوره برآورد مجموعه مواردی که برای برآورد مجدد پارامترهای مدل استفاده می شوند را تعریف می کند. بطور پیش فرض، دوره برآورد تمام حالت های موجود در دیتاست فعال را شامل می شود. برای اجرای دوره برآورد، گزینه **Based on Time** یا **Case range** را در جعبه محاوره ای **Select Cases** انتخاب کنید. بسته به داده های موجود، دوره برآورد استفاده شده توسط رویه ممکن است با مدل تغییر کند پس متفاوت از مقدار نمایش داده شده است. برای یک مدل داده شده، دوره برآورد صحیح پس از حذف مقادیر گم شده پیوسته، از متغیر وابسته مدل می باشد که در ابتدا یا انتهای دوره برآورد مشخص شده است.

دوره پیش بینی

دوره پیش بینی برای هر مدل همیشه با اولین مورد بعد از پایان دوره برآورد شروع می شود و تا آخرین مورد در دیتاست فعال یا زمان مشخص شده توسط کاربر پیش می رود. اگر پارامترها برآورد مجدد نشده باشند (بطور پیش فرض این چنین است)، پس دوره برآورد برای هر مدل، مجموعه مواردی است که وقتی مدل توسعه یافته بود استفاده شد.

First case after end of estimation period through last case in active dataset. این

گزینه را زمانی انتخاب کنید که پایان دوره برآورد پیش از آخرین مورد در دیتاست فعال است، و می خواهید تا آخرین مورد پیش بینی کنید.

First case after end of estimation period through a specified date این گزینه را برای مشخص کردن پایان دوره پیش بینی به طور صریح انتخاب کنید. مقادیر را برای همه خانه ها در Date grid وارد نمایید.

اگر هیچ زمانی برای دیتاست فعال مشخص نشود، Date grid ستون تکی Observation را نمایش می دهد. برای مشخص کردن پایان دوره پیش بینی، شماره ردیف مقدار مربوطه را وارد کنید.

ستون Cycle در Date grid مربوط به ارزش متغیر CYCLE در دیتاست فعال می باشد.

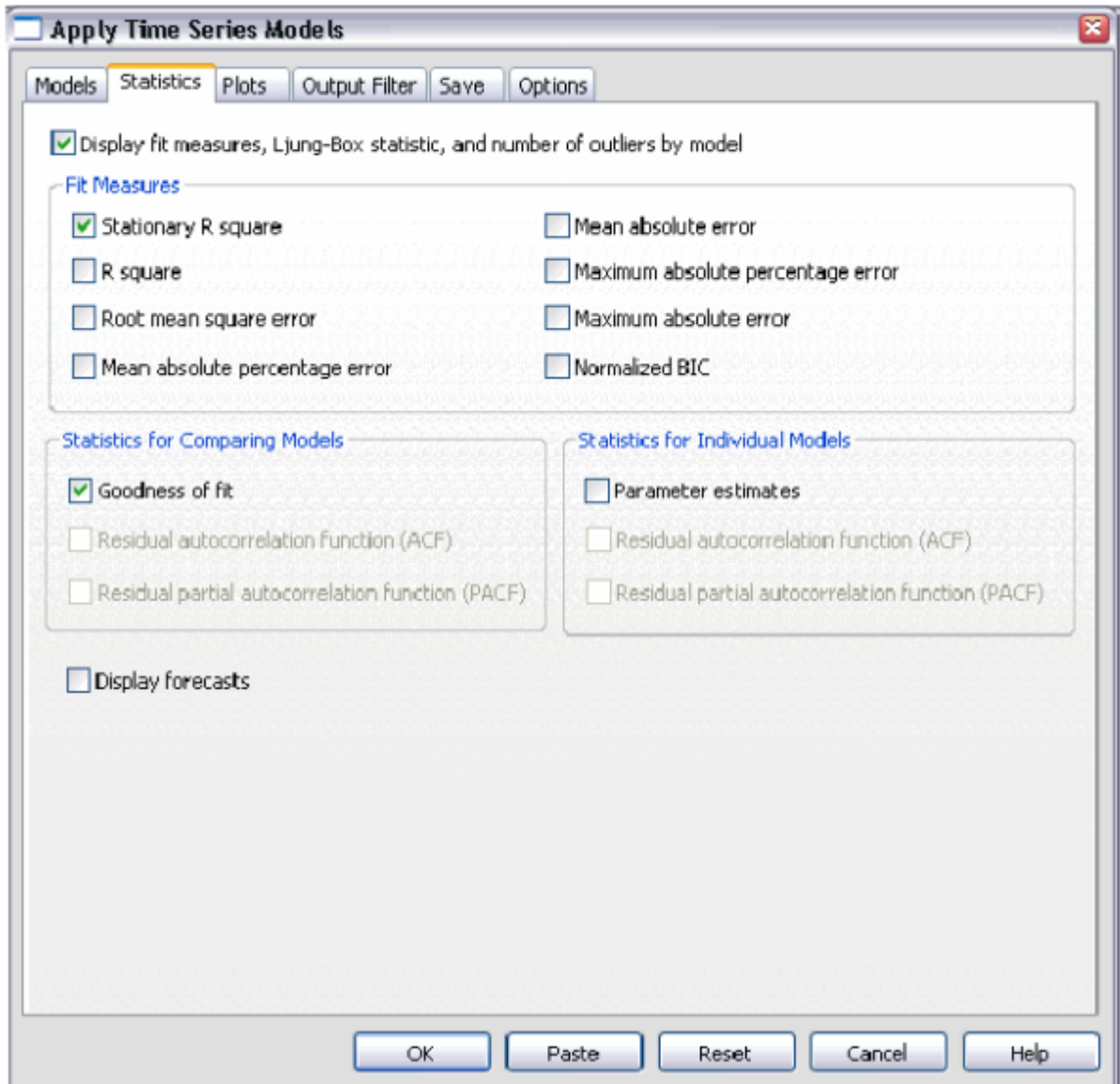
خروجی

خروجی حاضر شامل نتایجی برای مدل های تکی به خوبی تمام مدل ها می باشد. نتایج برای مدل های تکی می تواند محدود به مجموعه مدل های با بهترین یا بدترین برازش مبتنی بر معیارهای مشخص شده توسط کاربر باشد.

آماره ها و جداول پیش بینی

تصویر ۲

Statistics، سربرگ Apply Time Series Models



سربرگ Statistics گزینه هایی برای نمایش جداول و آماره های برازش مدل، آماره های مدل، توابع اتوکواریانس و پیش بینی ها ارائه می دهد. مقادیر نمایش داده شده از اندازه های برازش، مقادیر Ljung-Box و پارامترهای مدل آنهایی هستند که از فایل مدل گرفته شده و بر داده ها اثر می گذارد زمانی که هر مدل توسعه یافته است مگر آنکه پارامترهای مدل برآورد مجدد شده باشند. مقادیر دورافتاده همیشه از فایل مدل گرفته می شوند.

Display fit measures, Ljung-Box statistics, and number of outliers by model

این گزینه را برای نمایش جداول شامل برازش، مقدار Ljung-Box و تعداد مقادیر دورافتاده در هر مدل انتخاب کنید.

Fit measures شما می توانید یک یا چند مورد از مقادیر برازش را برای قرار دادن در جداول از لیست زیر انتخاب کنید:

- ⌚ Stationary R-square
- ⌚ R-square
- ⌚ Root mean square error
- ⌚ Mean absolute percentage error
- ⌚ Mean absolute error
- ⌚ Maximum absolute percentage error
- ⌚ Maximum absolute error
- ⌚ Normalized BIC

Statistics for comparing model این گروه از گزینه ها نمایش جداول شامل آماره ها در میان مدل ها را کنترل می کند. هر گزینه یک جدول مجزا تولید می کند. شما می توانید یک یا چند گزینه را انتخاب کنید.

Goodness of fit جدول خلاصه آماره ها و درصدها برای مربع R مانا، مربع R ، خطای مربع میانگین ریشه (RMSE) ، خطای مطلق میانگین (MAE)، خطای درصد مطلق میانگین (MAPE)، ماکسیمم خطای مطلق (MaxAE)، ماکسیمم خطای درصد مطلق (MaxAPE)، معیار اطلاعات بیزی نرمال شده (BIC).

Residual autocorrelation function (ACF) جدول خلاصه آماره ها و درصدها برای خودهمبستگی مانده ها در میان تمام مدل های برآورد شده. این جدول تنها وقتی پارامترهای مدل برآورد مجدد شده باشند در دسترس است.

Residual partial autocorrelation function (PACF) جدول خلاصه آماره ها و درصدها برای خودهمبستگی جزئی مانده ها در میان تمام مدل های برآورد شده. این جدول تنها وقتی پارامترهای مدل برآورد مجدد شده باشند در دسترس است.

Statistics for individual models این گروه از گزینه ها نمایش جداول شامل جزئیات اطلاعات برای هر مدل را کنترل می کنند. هر گزینه یک جدول مجزا تولید می کند. شما می توانید یک یا چند مورد از گزینه های زیر را انتخاب نمایید:

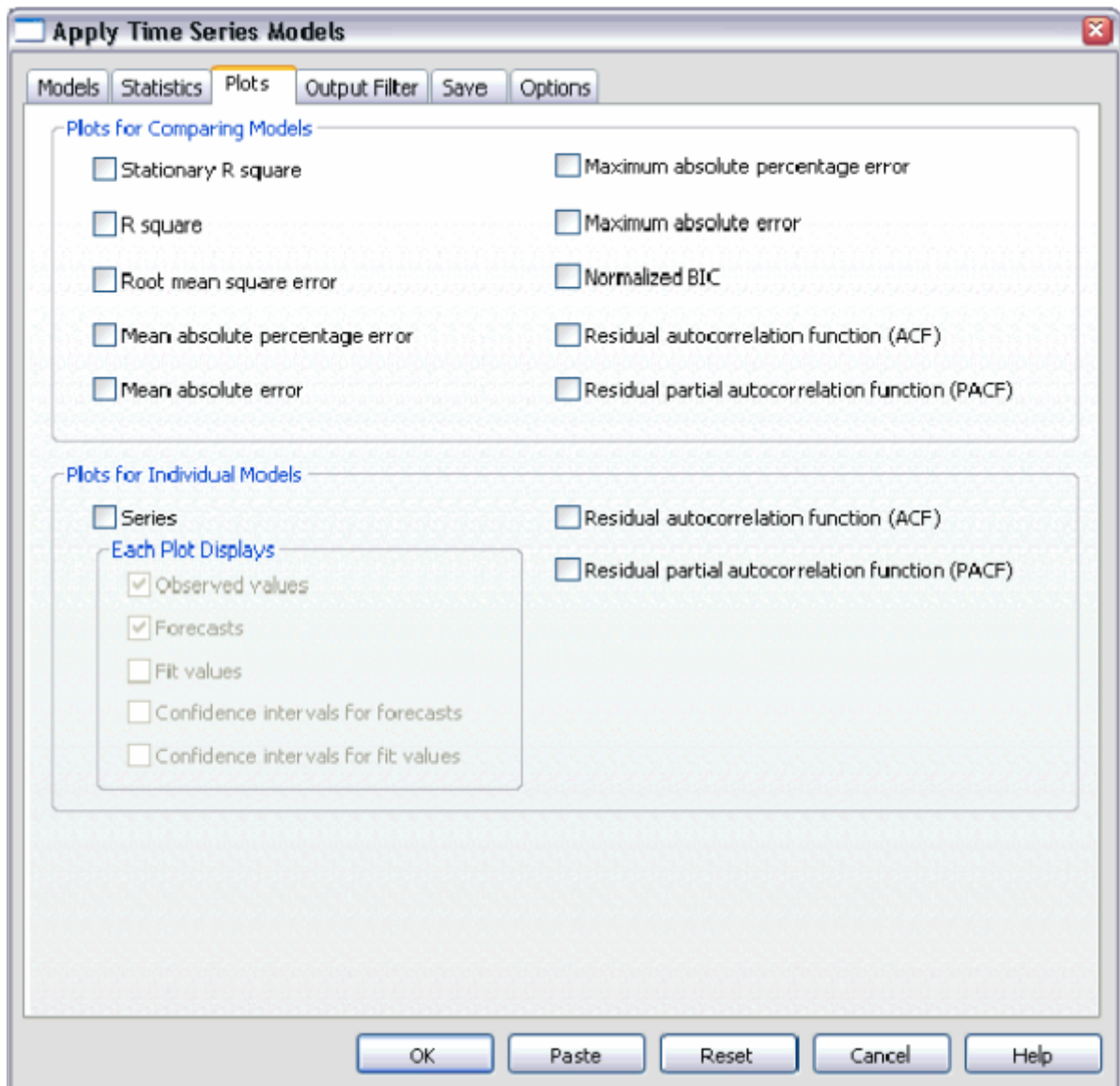
- **Parameter estimates.** یک جدول از پارامترهای برآورد شده را برای هر مدل نشان می دهد. جداول مجزا برای مدل های هموارسازی نمایی و ARIMA نمایش داده می شوند. اگر مقدار دورافتاده ای وجود داشته باشد، پارامترهای برآورد شده برای آنها نیز در جدول جداگانه ای نمایش داده می شود.

- Residual autocorrelation function (ACF). یک جدول از خودهمبستگی مانده ها با تأخیر برای هر مدل برآورد شده نمایش می دهد. جدول شامل بازه های اطمینان برای خودهمبستگی می باشد. این جدول تنها زمانی که پارامترهای مدا برآورد مجدد شده باشند در دسترس است.
- Residual partial autocorrelation function (PACF). یک جدول از خودهمبستگی جزئی مانده ها با تأخیر برای هر مدل برآورد شده نمایش می دهد. جدول شامل بازه های اطمینان برای خودهمبستگی جزئی می باشد. این جدول تنها زمانی که پارامترهای مدا برآورد مجدد شده باشند در دسترس است.
- Display forecasts. یک جدول از پیش بینی های مدل و بازه های اطمینان برای هر مدل نمایش می دهد.

نمودارها

تصویر ۳

Plots، سربرگ Apply Time Series Models



سربرگ Plots گزینه هایی برای نمایش نمودارهایی از آماره های برازش مدل، توابع خودهمبستگی و مقادیر سری ها (شامل پیش بینی ها) را ارائه می دهد.

Plots for comparing Models

این گروه از گزینه ها، نمایش نمودارهایی شامل آماره ها در میان تمام مدل ها را کنترل می کند. مقادیر نمایش داده شده همان هایی هستند که از فایل مدل بوده و بر داده های مورد استفاده تأثیر می گذارد وقتی که هر مدل توسعه داده شده است مگر آنکه پارامترهای مدل برآورد مجدد شده باشند. بعلاوه نمودارهای خودهمبستگی تنها زمانی در

دسترس هستند که پارامترهای مدل برآورد مجدد شده باشند. هر گزینه یک نمودار مجزا را تولید می کند. شما می توانید یک یا چند مورد از گزینه های زیر را انتخاب کنید:

- ⌚ Stationary R -square
- ⌚ R -square
- ⌚ Root mean square error
- ⌚ Mean absolute percentage error
- ⌚ Mean absolute error
- ⌚ Maximum absolute percentage error
- ⌚ Maximum absolute error
- ⌚ Normalized BIC
- ⌚ Residual autocorrelation function (ACF)
- ⌚ Residual partial autocorrelation function (PACF)

Plots for Individual Models

Series. این گزینه را برای بدست آوردن نمودارهایی از مقادیر پیش بینی شده برای هر مدل انتخاب کنید. مقادیر مشاهده شده، مقادیر برازش ، بازه های اطمینان برای مقایسه برازش و خودهمبستگی ها تنها اگر پارامترهای مدل برآورد مجدد شده باشند در دسترس می باشند. شما می توانید یک یا چند مورد از موارد زیر برای وجود در نمودار انتخاب کنید:

- Observed Values. مقادیر مشاهده سری های وابسته
- Forecasts. مقادیر پیش بینی شده مدل برای دوره پیش بینی
- Fit Values. مقادیر پیش بینی شده مدل برای دوره برآورد
- Confidence intervals for forecasts. بازه های اطمینان برای دوره پیش بینی
- Confidence intervals for fit values. بازه های اطمینان برای دوره برآورد

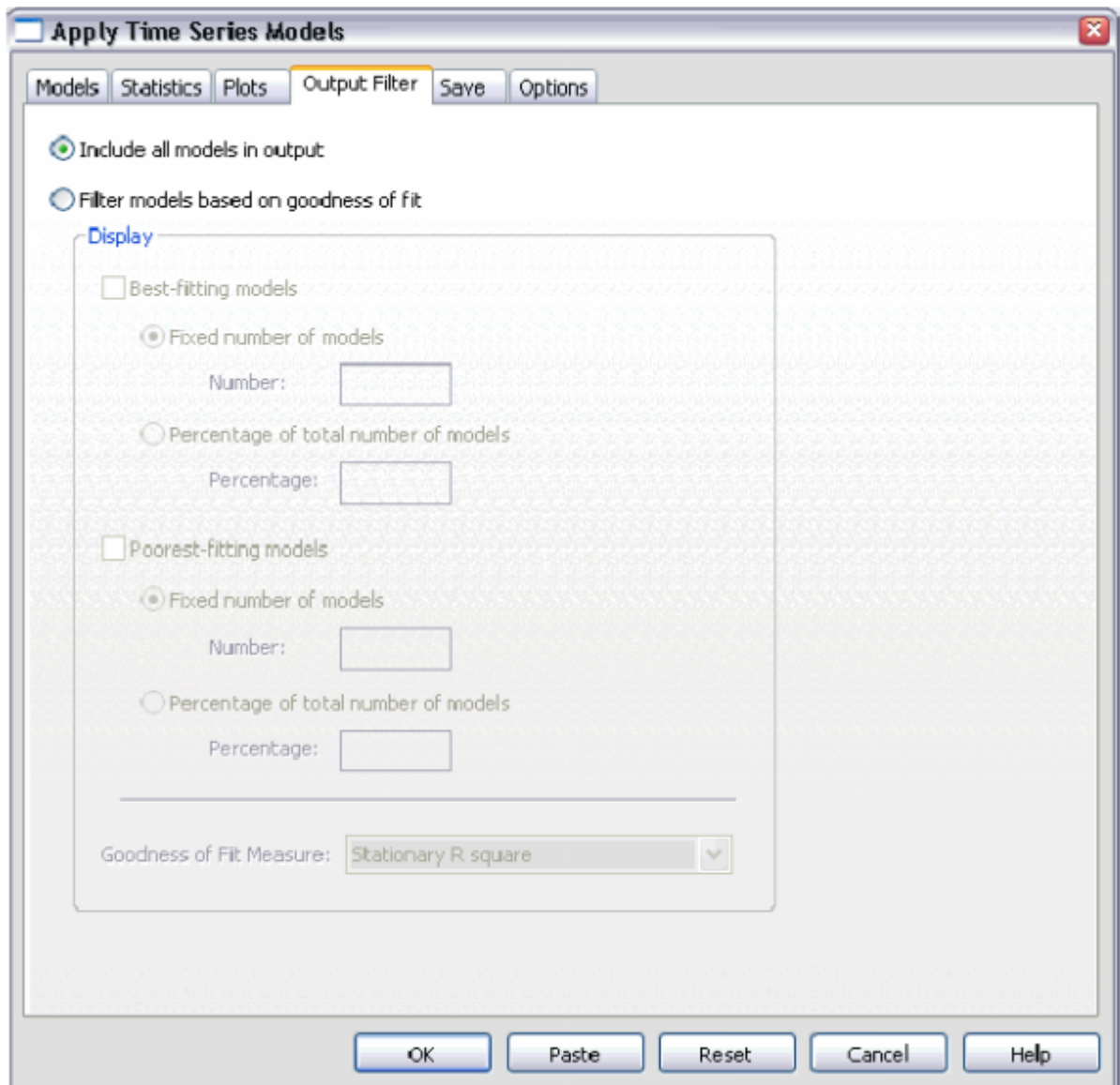
Residual autocorrelation function (ACF). یک نمودار از خودهمبستگی های مانده ها برای هر مدل برآورد شده نمایش می دهد.

Residual partial autocorrelation function (PACF). یک نمودار از خودهمبستگی های جزئی مانده ها برای هر مدل برآورد شده نمایش می دهد.

محدود کردن خروجی به مدل‌ها با بهترین یا ضعیف‌ترین برازش

تصویر ۴

Output Filter، سربرگ Apply Time Series Models



سربرگ Output Filter گزینه‌هایی برای محدود کردن خروجی جدولی و نموداری بر یک زیرمجموعه از مدل‌ها ارائه می‌دهد. شما می‌توانید برای محدود کردن خروجی به بهترین یا بدترین برازش با توجه به ضوابط برازش که خود مهیا می‌کنید، انتخاب کنید. بطور پیش فرض تمام مدل‌ها در خروجی موجود است. به جز زمانی که پارامترهای مدل برآورد مجدد شده‌اند ارزش مقادیر برازش مورد استفاده برای محدود کردن مدل‌ها همان‌هایی هستند که در فایل مدل بوده و بر داده اثر می‌گذارد.

Best-fitting models. این گزینه را برای نمایش مدل هایی با بهترین برازش در خروجی، انتخاب کنید. یک اندازه نیکویی برازش انتخاب و تعداد مدل ها را مشخص کنید. انتخاب این گزینه مانعی برای انتخاب همزمان مدل ها با کمترین برازش ایجاد نمی کند. در این صورت خروجی شامل مدل های با کمترین برازش به خوبی مدل های با بهترین برازش خواهد بود.

- Fixed number of models نتایج نمایش داده شده برای n مدل با بهترین برازش را مشخص می کند. اگر عدد از جمع تمام مدل ها تجاوز کند، تمام مدل ها نمایش داده می شوند.
- Percentage of total number of models نتایج نمایش داده شده برای مدل های با مقادیر نیکویی برازش با n درصد بالایی در میان تمام مدل ها را نمایش می دهد.

Poorest-fitting models. این گزینه را برای نمایش مدل هایی با ضعیفترین برازش در خروجی، انتخاب کنید. یک اندازه نیکویی برازش انتخاب و تعداد مدل ها را مشخص کنید. انتخاب این گزینه مانعی برای انتخاب همزمان مدل ها با بهترین برازش ایجاد نمی کند. در این صورت خروجی شامل مدل های با بهترین برازش به خوبی مدل های با ضعیف ترین برازش خواهد بود.

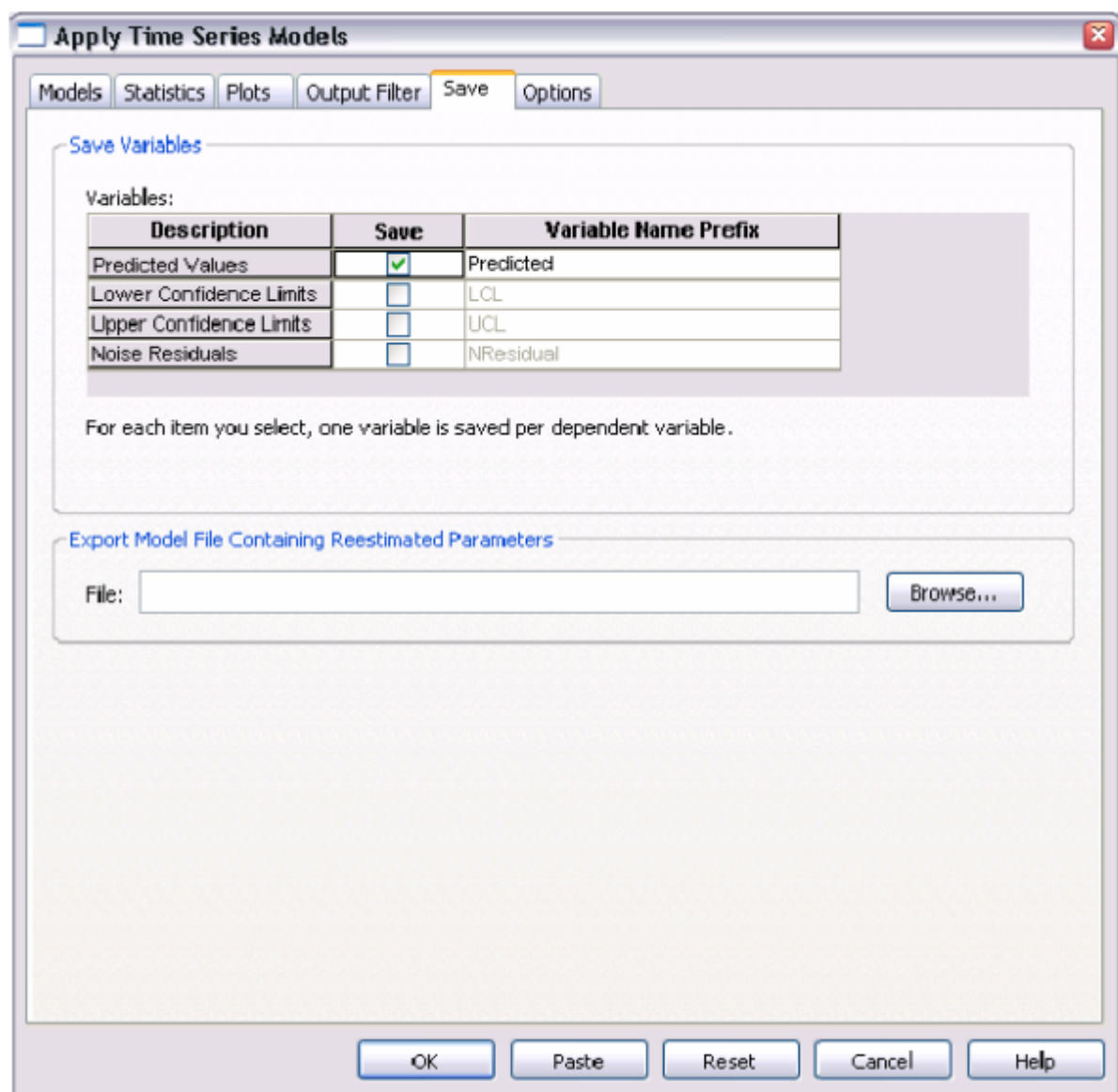
- Fixed number of models نتایج نمایش داده شده برای n مدل با ضعیف ترین برازش را مشخص می کند. اگر عدد از جمع تمام مدل ها تجاوز کند، تمام مدل ها نمایش داده می شوند.
- Percentage of total number of models نتایج نمایش داده شده برای مدل های با مقادیر نیکویی برازش با n درصد پائینی در میان تمام مدل ها را نمایش می دهد.

Goodness of Fit Measure. مقدار نیکویی برازش را برای فیلتر کردن مدل ها انتخاب کنید. پیش فرض مربع R مانا می باشد.

ذخیره پیشگویی های مدل و مشخصات مدل

تصویر ۵

Save برگ ، Apply Time Series Models



سربگ Save به شما اجازه می دهد پیشگویی های مدل را به عنوان متغیرهای جدید در دیتاست فعال ذخیره کرده و مشخصات مدل را در یک فایل خروجی با قالب XLM ذخیره نمایید.

Save Variables . شما می توانید پیش بینی ها، بازه های اطمینان و مانده ها را به عنوان متغیرهای جدید در دیتاست فعال ذخیره کنید. هر مدل به دسته جدید متغیرهای خود ارتقا می یابد. موارد جدید اگر دوره پیش بینی در طول سری متغیر وابسته با مدل همبسته باشد، اضافه می شوند. مقادیر پیش بینی شده و محدوده های اطمینان فقط برای

دوره پیش بینی ساخته شده اند مگر آنکه پارامترها برآورد مجدد شده باشند. متغیرهای جدید را با انتخاب جعبه انتخاب Save برای هر کدام، ذخیره کنید. به طور پیش فرض هیچ متغیر جدیدی ذخیره نشده است.

- Predicted Values. مقادیر پیش بینی شده مدل
- Lower Confidence Limits. حدود پایین اطمینان برای مقادیر پیش بینی شده
- Upper Confidence Limits. حدود بالای اطمینان برای مقادیر پیش بینی شده
- Noise Residuals. مانده های مدل. وقتی تبدیلات متغیر وابسته اجرا شده (بطور مثال لگاریتم طبیعی) این ها مانده هایی برای مدل تبدیل شده هستند. این انتخاب تنها زمانی در دسترس است که پارامترهای مدل برآورد مجدد شده باشند.
- Variable Name Prefix. پیشوندها را برای استفاده در نام های متغیرهای جدید مشخص کرده و یا پیشوندهای پیش فرض را در نظر بگیرید. نام های متغیرها شامل پیشوند، نام متغیر وابسته و مشخص کننده مدل است. نام متغیر تمديد می شود اگر لازم باشد از تضاد نام گذاری متغیر دوری شود. پیشوند باید با قواعد نام های مجاز متغیرها مطابقت داشته باشد.

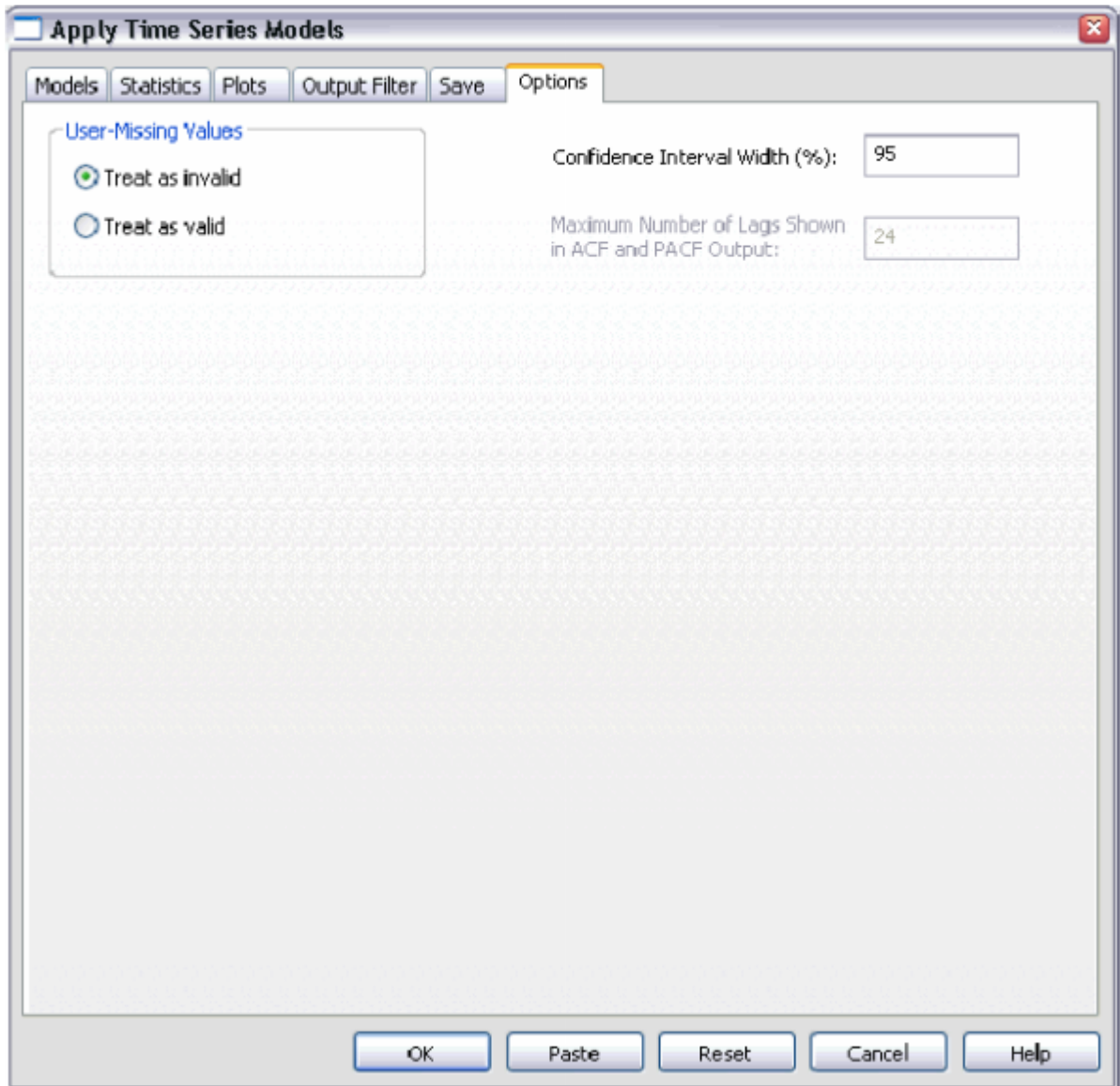
Export Model File Containing Reestimated Parameters. مشخصه های مدل شامل

پارامترهای برآورد شده و آماره های برازش به یک فایل مشخص با قالب XLM ارسال می شود. این گزینه تنها اگر پارامترهای مدل برآورد مجدد شده باشند در دسترس است.

گزینه ها

تصویر ۶

Options، Apply Time Series Models، سربرگ



سربرگ Options به شما اجازه می دهد نحوه بررسی مقادیر گمشده را کنترل کرده، اندازه بازه اطمینان را مشخص کنید، و تعداد تأخیرها برای خودهمبستگی ها را مشخص کنید.

User-missing Values. این گزینه بررسی مقادیر گمشده کاربر را کنترل می کند.

- Treat as invalid. مقادیر گمشده کاربر مانند مقادیر گمشده سیستم تلقی می شود.
- Treat as valid. مقادیر گمشده کاربر بعنوان داده معتبر تلقی می شود.

Missing Value Policy. قوانین زیر برای رفتار مقادیر گم شده به کار می روند. (شامل مقادیر گم شده سیستم و مقادیر گم شده کاربر که غیر معتبر تلقی می شود):

- مواردی که با مقادیر گم شده از یک متغیر وابسته که در دوره برآورد اتفاق می افتند در مدل گنجانده شده اند. بررسی خاصی از مقادیر گم شده بستگی به شیوه برآورد دارد.
- برای مدل های ARIMA، یک اخطار صادر می شود اگر یک پیشگو هر مقدار گم شده ای در دوره برآورد داشته باشد. هر مدل شامل پیشگو، برآورد مجدد شده نیست.
- اگر هر متغیر مستقل در دوره پیش بینی مقادیر گم شده داشته باشد، روند یک خطا می دهد و پیش بینی ها تا حد ممکن دور می شوند.

Confidence Interval width (%). بازه های اطمینان برای پیشگویی های مدل و خودهمبستگی های مانده ها محاسبه شده اند. می توانید هر مقدار مثبت کمتر از ۱۰۰ را مشخص کنید. بطور پیش فرض یک بازه اطمینان ۹۵٪ استفاده شده است.

Maximum Number of Lags Shown in ACF and PACF output. می توانید حداکثر مقدار تأخیرهای مشاهده شده در جدول ها و نمودار خودهمبستگی ها و خودهمبستگی جزئی را مشخص کنید. این گزینه تنها زمانی در دسترس است که پارامترهای مدل برآورد مجدد شده باشند.