

Seasonal ARIMA: forecasting a series with a **yearly rhythm**

Monthly sales, call volumes, energy use: series with a trend and a pattern that repeats every year. SARIMA models both, forecasts with a band, simulates futures for a Monte Carlo model, and can be tested against the months it did not see. Five XL Edge functions cover the whole workflow, each a live formula.

Functions `fx.RiskSARIMARankλ` · `fx.RiskSARIMAFitλ` · `fx.RiskSARIMAForecastλ` · `fx.RiskSARIMAλ` · `fx.RiskSARIMABacktestλ`

In XL Edge XL Edge tab > Productivity Tools > Stat Tools > Insert Time Series > Seasonal ARIMA (SARIMA) section

Example *XL Edge Example - SARIMA.xlsx* (same folder), data from *SARIMA - Example Data.xlsx*. Every number in this guide is read from it.

Method Box-Jenkins SARIMA up to (2,1,2)(1,1,1); conditional least squares; ψ -weight forecast bands.

1 The idea, and the airline model

2 Using the ribbon

3 The data

4 Shortlist orders: Rank

5 The Model block: Fit

6 The forecast table

7 Simulated futures and the fan chart

8 Would it have worked? Backtest

9 Practical advice

10 Argument reference

1 The idea, and the airline model

A SARIMA model is written $(p, d, q)(P, D, Q)_s$. The first bracket describes month-to-month behavior, the second year-to-year behavior, and s is the season length (12 for monthly data):

Letter	Name	In plain words
d, D	Differencing	d = 1 models the change from last month; D = 1 the change from the same month last year. Together they remove the trend and the season.
p, P	Autoregressive terms	This month's change depends on recent changes (p) or on the change a year ago (P).
q, Q	Moving-average terms	This month's change partly corrects recent surprises (q) or last year's surprise in the same month (Q).

The default, and the best place to start, is the **airline model (0,1,1)(0,1,1)₁₂**, named after Box and Jenkins' classic airline-passenger example. It differences both ways and keeps one moving-average term of each kind: two coefficients in all. In effect it forecasts next January as this January plus the recent year-over-year change, smoothed so that one odd month does not throw it off. It fits a remarkable range of business series. The *Correlogram* guide shows how this example's correlogram points straight at it.

2 Using the ribbon

XL Edge tab > Productivity Tools > Stat Tools (arrow) > Insert Time Series > Seasonal ARIMA (SARIMA)		
Menu item	XL Edge asks for	Result
Rank SARIMA orders	The history (one column, oldest first); d and D (0 or 1, default 1); season length; Log.	13 × 3
Fit SARIMA	The history; the order as six cells or {0,1,1,0,1,1} (none = airline); season length (none = 12); Log (TRUE to model the logarithm).	19 × 2 Model block
SARIMA forecast table	The history; the Model block; periods ahead; the band's percentiles (0.1 and 0.9); labels for the future periods.	(Periods + 1) × 4
SARIMA simulated paths	The history; the Model block; periods ahead; optional Student t degrees of freedom; then trials (MC_Trials), VarID and seed as for any simulator.	Trials × Periods
SARIMA backtest	The history; how many months to hold out; the order, season and Log to test; the band's percentiles.	(Holdout + 6) × 6

For a fan chart, select the simulated paths and use *Insert Monte Carlo Chart > Fan Chart (Time Series)*. Every result is one live formula: add a month to the history and everything refits.

3 The data

108 months of sales, January 2017 to December 2025 (sheet *Data*, column C, copied from *SARIMA - Example Data.xlsx*). Sales peak each summer and dip each winter. The level rose until 2020 (an average month of 152,448) and has eased every year since (2025: 126,915). The last value, for December 2025, is 109,280; December 2024 was 113,511. The history is passed as a single column, oldest first: `Data!C5:C112`.

4 Shortlist orders: `fx.RiskSARIMARankλ`

```
history; d, D, season and Log omitted = 1, 1, 12, FALSE
=fx.RiskSARIMARankλ(Data!C5:C112)
```

Rank fits twelve common orders and sorts them by **AIC** (Akaike's information criterion: fit quality, penalized for each extra coefficient; lower is better). All twelve are scored on the same months, so their AICs compare fairly.

Sheet 1 Rank, first and eighth of twelve

Place	Order	AIC	StDev
1	(1,1,1)(1,1,1)12	1,429.4	6,509
8	(0,1,1)(0,1,1)12 (airline)	1,446.9	8,277

The leader, (1,1,1)(1,1,1)12, uses all four kinds of term. The airline model comes 8th. On AIC alone you would pick the leader; section 8 shows why you should not, yet. The same ranking on the logarithm of sales (`Log` = TRUE, right of the first table) has the same leader, (1,1,1)(1,1,1)12.

Turning the leader into an Order. Rank shows orders as text. To fit or backtest the leader, the example converts the text to six numbers with a formula, so it follows the ranking if the data change:

```
"(1,1,1)(1,1,1)12" becomes 1, 1, 1, 1, 1, 1 in six cells (sheet 1 Rank, B21)
=--TEXTSPLIT(TEXTBEFORE(MID(SUBSTITUTE(A7, "(" , ","), 2, 99), ","), ",")
```

5 The Model block: `fx.RiskSARIMAFitλ`

```
=fx.RiskSARIMAFitλ(Data!C5:C112)          airline model
=fx.RiskSARIMAFitλ(Data!C5:C112, , , TRUE)  airline model on LN(sales)
=fx.RiskSARIMAFitλ(Data!C5:C112, '1 Rank'!B21#) Rank's leader
```

The result is a 19 × 2 block that the forecast, the paths and the residual tools all read: the order (p, d, q, P, D, Q, Season, Log), the Constant, the six possible coefficients (AR1, AR2, MA1, MA2, SAR1, SMA1; zero when not in the model), StDev, AIC and Observations.

Sheet 2 Fit

Model	AR1	MA1	SAR1	SMA1	StDev
Airline	-	-0.642	-	-0.315	8,277
Airline, Log	-	-0.580	-	-0.293	0.0503
Rank's leader	0.320	-0.817	0.080	-0.807	6,509

Reading it. StDev is the typical size of a one-month forecast error: about 8,277 for the airline model, some 6% of an average month (141,062). On logs, StDev is a proportion: 0.0503 means about 5% a month. The leader's coefficients carry a warning: its AR1 (0.320) and MA1 (-0.817) nearly cancel each other, as do its SAR1 and SMA1. Terms that offset each other add flexibility the data cannot really support, which is overfitting.

DON'T COMPARE FIT'S AIC ACROSS ORDERS

Fit reports AIC on its own residuals: 1,718.0 for the airline model, 1,446.0 for the leader. A model with autoregressive terms loses its first months (13 here) and is scored on fewer residuals, so its AIC is lower for a reason that has nothing to do with fit. Compare orders only with Rank, which scores every candidate on the same window.

WHY OTHER SOFTWARE GIVES DIFFERENT COEFFICIENTS

Fit uses conditional least squares, treating the errors before the first month as zero. Packages such as statsmodels use exact maximum likelihood, so on a short series their coefficients and StDev differ somewhat. The forecasts usually agree closely; judge models by their backtests (section 8), not by one coefficient.

6 The forecast table: `fx.RiskSARIMAForecast`

```
sheet 3 Forecast: C4 holds 24 month labels; the table reads the airline Model block
C4: =TEXT(EOMONTH(INDEX(Data!A5:A112, ROWS(Data!A5:A112)), SEQUENCE(1, 24)), "mmm-yy")
A7: =fx.RiskSARIMAForecast(Data!C5:C112, '2 Fit'!A6#, 24, , , TRANSPOSE(C4#))
```

Sheet 3 Forecast (24 rows; four shown)

Month	Forecast	P10	P90
Jan-26	105,946	95,339	116,553
Jun-26	134,577	120,994	148,160
Dec-26	103,937	87,479	120,396
Dec-27	98,929	69,892	127,966

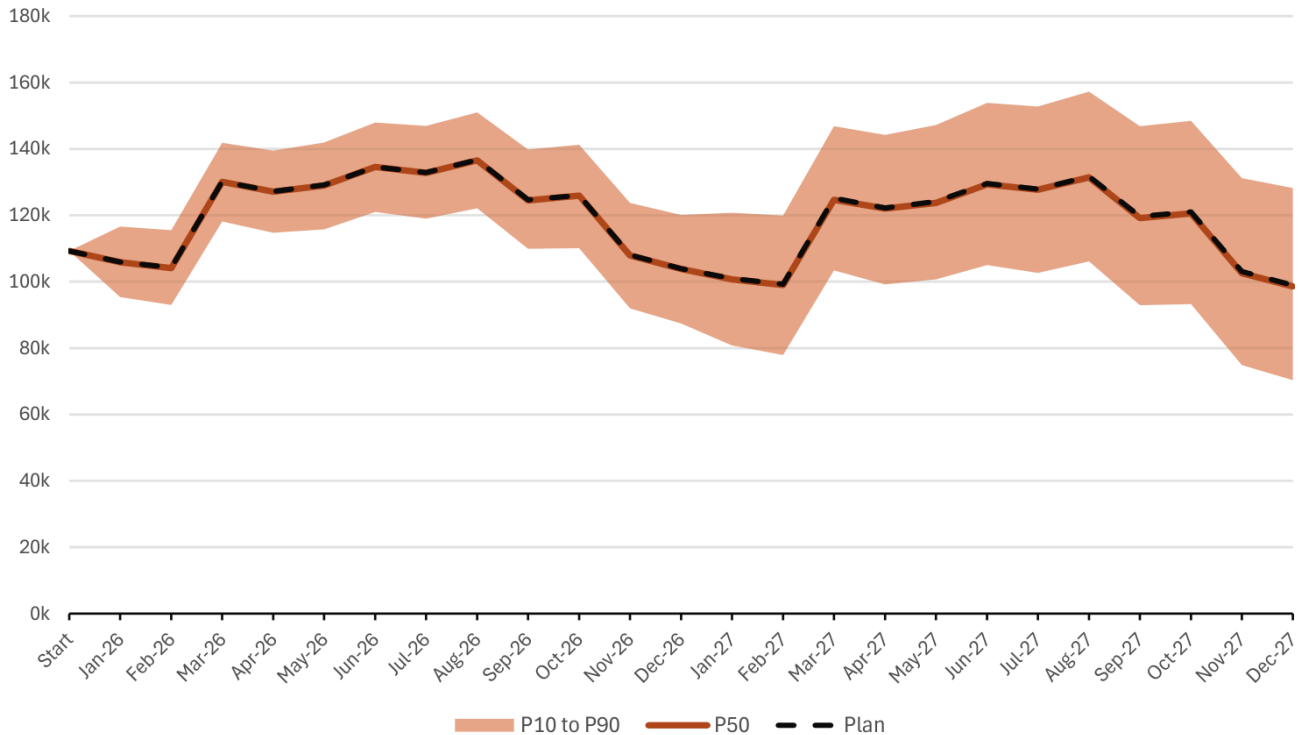
Reading it. Each row gives the forecast and a P10 to P90 band: an 80% range, so one month in five should fall outside it. The band widens with the horizon ($\pm 10,607$ in the first month, $\pm 29,037$ by the 24th), because each month's uncertainty adds to the last. December 2026 is forecast at 103,937, below December 2025's 109,280, which was itself below 2024's 113,511: the model carries the latest year-over-year decline forward. That is what differencing does. Whether the decline continues is a business question, and the forecast is only as good as that assumption.

7 Simulated futures and the fan chart: `fx.RiskSARIMAX`

For a Monte Carlo model you need whole futures, not just a band: 10,000 possible paths of the next 24 months, each internally consistent (a weak spring tends to stay weak into summer). `fx.RiskSARIMAX` draws them from the same Model block, one row per trial:

```
=fx.RiskSARIMAX(Data!C5:C112, '2 Fit'!A6#, 24, , 10000)          sheet Paths, A3
=fx.RiskChartFan(Paths!A3#, , , "Monthly sales", INDEX(Data!C5:C112, ROWS(Data!C5:C112)),
                  DROP(CHOOSECOLS(A7#, 2), 1), TRANSPOSE(C4#))    start, plan =
the forecast, labels
```

Monthly sales



XL Edge's Fan Chart (sheet *3 Forecast*): the P10 to P90 band of the 10,000 paths, their median (P50), and the forecast table's values as the dashed Plan line. The fan opens from December 2025's actual sales.

The simulated band matches the table: in Jan-26 the paths' P50 is 105,850 against the forecast's 105,946, and their P10 to P90 is 95,399 to 116,598 against the table's 95,339 to 116,553 (simulation noise accounts for the difference). Feed the paths into the rest of a model, such as $\text{revenue} = \text{sales} \times \text{price}$, and every result inherits a realistic, seasonally shaped uncertainty.

8 Would it have worked? `fx.RiskSARIMABacktestA`

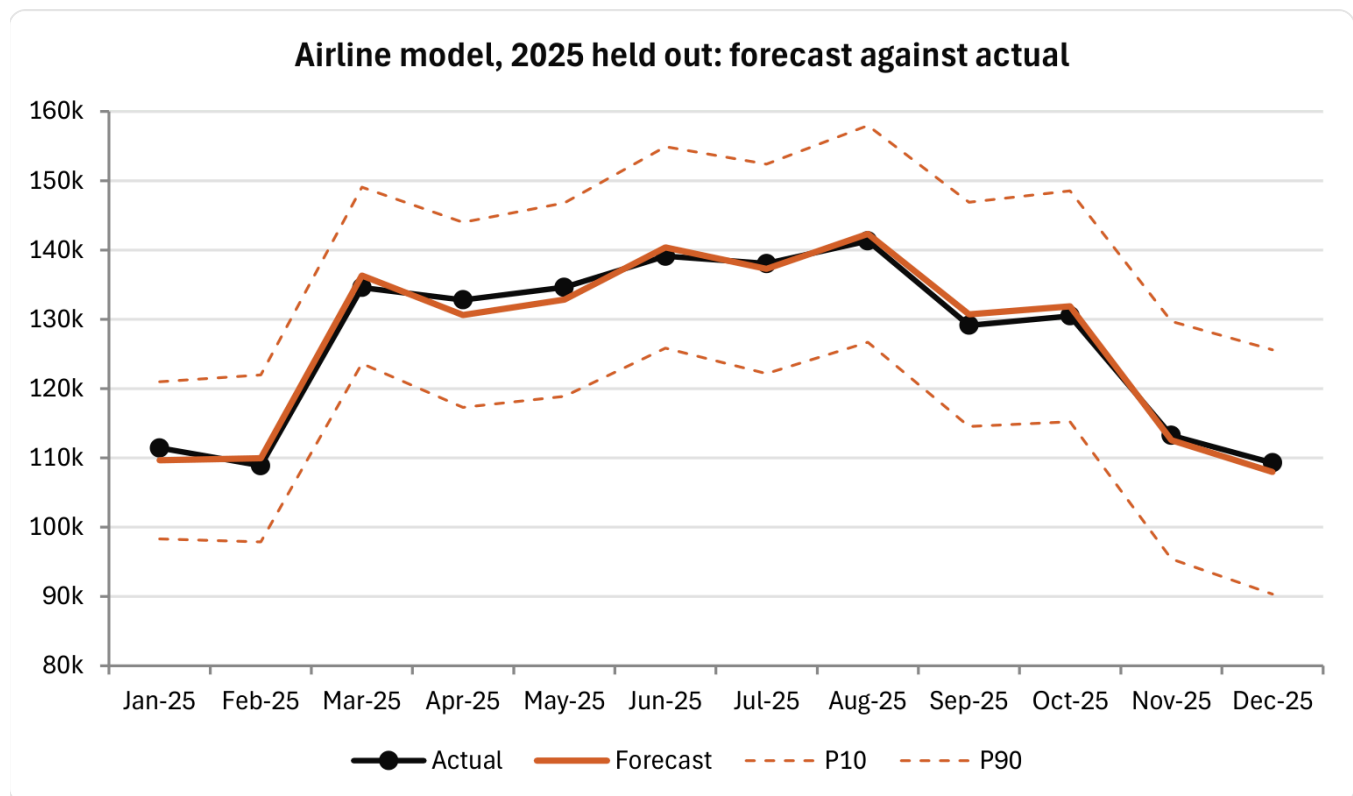
A backtest hides the latest months, refits on the rest, forecasts the hidden months and scores the forecast against what actually happened. It is the honest test of a forecasting model, because the model never saw the answers.

```
=fx.RiskSARIMABacktestA(Data!C5:C112, 12)
=fx.RiskSARIMABacktestA(Data!C5:C112, 12, '1 Rank'!B21#)
=fx.RiskSARIMABacktestA(Data!C5:C112, 12, , , TRUE)
=fx.RiskSARIMABacktestA(Data!C5:C112, 12, {0,1,2,0,1,1})
term
```

```
airline, 2025 held out
Rank's leader
airline on logs
airline plus a second MA
```

Sheet 4 *Backtest*. MAPE: mean absolute percentage error. Coverage: share of actual months inside the P10 to P90 band.

Candidate	MAPE, 12 months	MAPE, 24 months	RMSE, 24 months	Coverage, 12 / 24
Airline (0,1,1)(0,1,1)12	1.09%	2.50%	3,653	100% / 100%
Rank's leader (1,1,1)(1,1,1)12	2.10%	3.41%	5,235	100% / 100%
Airline, Log = TRUE	0.91%	2.51%	3,735	100% / 100%
(0,1,2)(0,1,1)12	1.06%	2.16%	3,066	100% / 100%



The airline model's 2025 backtest (sheet 4 *Backtest*, an ordinary Excel chart of the block): forecast made from data to December 2024, against what happened.

Reading it.

- **AIC's favorite forecasts worst.** Rank's leader misses by 2.10% against the airline model's 1.09% over 12 months, nearly double, and by 3.41% against 2.50% over 24. Its extra, offsetting terms fitted the history's noise, not its pattern. That is the overfitting the Fit section warned about, and the reason to backtest before choosing.
- **The airline model is hard to beat.** Under 1.1% error a year ahead is very good for monthly sales.
- **Small refinements help a little.** Logs improve the 12-month error to 0.91%; the second MA term, which the airline residuals' correlogram hinted at, gives 1.06% over 12 months and 2.16% over 24. The differences are modest, and a different holdout could reorder them. Prefer the simplest model unless the gain holds across holdouts.
- **The bands are generous.** Every candidate covered all the held-out months, where a P10 to P90 band should miss about one month in five. Conditional least squares tends to overstate StDev on short

series, so read the band as a cautious range, not an exact 80%.

9 Practical advice

- **Start with the airline model**, then let Rank suggest a shortlist and the backtest choose. Never choose on AIC alone.
- **Use Log = TRUE** when the seasonal swings grow with the level of the series (the summer peak gets bigger as sales grow). Forecasts come back in sales units; the forecast becomes a median and the band skews upward.
- **Have enough history**: at least four or five full years for monthly data. Here there are nine.
- **One-off events** (a pandemic year, a lost customer) distort the coefficients. Consider adjusting them in a copy of the history and comparing.
- **Check the residuals** with the correlogram (ModelDF = the number of fitted coefficients); see the *Correlogram* guide.
- **Compare with Holt-Winters**. XL Edge's Holt-Winters functions use the same Model-block, forecast and backtest layouts, so the two methods sit side by side.
- **Refit as data arrive**. Because every block is a live formula, extending the history range refits the model, the forecast and the paths.

10 Argument reference

fx.RiskSARIMARankλ(History, [Diff], [SeasonalDiff], [Season], [Log])

Argument	Default	Meaning
History	required	The series, oldest first, one column (all > 0 when Log is TRUE).
Diff, SeasonalDiff	1, 1	d and D, each 0 or 1.
Season	12	Season length.
Log	FALSE	TRUE models LN(History).

Output, 13 × 3: Order (text), AIC and StDev for 12 common orders, lowest AIC first, scored on a common window.

fx.RiskSARIMAFitλ(History, [Order], [Season], [Log])

Argument	Default	Meaning
Order	{0,1,1,0,1,1}	p, d, q, P, D, Q as six cells or an array: p, q ≤ 2; P, Q ≤ 1; d, D ≤ 1.

Output, 19 × 2: Parameter / Estimate; p, d, q, P, D, Q, Season, Log, Constant, AR1, AR2, MA1, MA2, SAR1, SMA1, StDev, AIC, Observations. Constant is the mean of the differenced series when d + D = 0, else 0. Estimation: a coordinate grid, then 20 Levenberg-Marquardt steps.

fx.RiskSARIMAForecastλ(History, Model, Horizon, [Lower], [Upper], [Labels])

Argument	Default	Meaning
Model	required	A Model block from Fit (or typed in its layout).
Horizon	required	Periods ahead.
Lower, Upper	0.10, 0.90	Percentiles of the band.
Labels	1, 2, 3 ...	One label per future period (a column).

Output, (Horizon + 1) × 4: Period, Forecast, Lower, Upper. Band: forecast ± z × StDev × √(Σψ²).

fx.RiskSARIMAX(History, Model, Periods, [DegFree], [Trials], [VarID], [Seed])

Argument	Default	Meaning
DegFree	normal	Student t shocks with this many degrees of freedom (> 2), for fatter tails.
Trials, VarID, Seed	10000, 1, 0	As for every XL Edge simulator; Trials may be a Trials × Periods block of uniforms.

Output: Trials × Periods, one row per trial; percentiles match the forecast table's band.

```
fx.RiskSARIMABacktestλ(History, Holdout, [Order], [Season], [Log], [Lower], [Upper])
```

Argument	Default	Meaning
Holdout	required	How many of the latest periods to hide, e.g. 12.
Output, (Holdout + 6) × 6: MAPE, RMSE, MAE, Bias (forecast – actual) and Coverage; then Period (position in History), Actual, Forecast, Error, Lower, Upper.		

References: Box, Jenkins and Reinsel, *Time Series Analysis: Forecasting and Control* (the airline model, conditional least squares); Akaike (1974), *IEEE Transactions on Automatic Control* 19(6). Example data: *SARIMA - Example Data.xlsx*. XL Edge and the fx.Risk... functions: © 2026 W Phillips, edgewisedata.com, MIT License.