Time Series Demo (Forecast in ARIMA)

Accuracy versus Fit

- *Information criteria* are used to compare how well different models perform on the data that is used to build the model (that is, the fit sample).
  They are penalized versions of fit measures.
  - For example: AIC, SBC
- *Accuracy statistics* are used to evaluate forecasting error, either in the *fit sample* or in data not used to build the model (that is, the *holdout sample*).
  - For example: MAPE, MAE, RMSE
- In general, an accuracy statistic on out-of-sample data (such as a holdout sample) provides an unbiased estimate of implementation accuracy, that is, the accuracy likely to be experienced when the forecast model is deployed.
  - Assessing a predictive model using accuracy statistics calculated for a *holdout sample* is called *honest assessment*.

Accuracy versus Fit

	Quarter	t	
Ultimate Goal: Forecast the next four quarters.	4Q2018	Y_{t+4}	
	3Q2018	Y_{t+3}	
	2Q2018	Y_{t+2}	
	1Q2018	Y_{t+1}	
Accuracy statistics, such as MAPE, MAE, RMSE, and so on, evaluate forecasting error (or lack thereof) on data that is not used to build the model.	4Q2017	Y_t	Holdout Sample
	3Q2017	Y_{t-1}	
	2Q2017	Y_{t-2}	
	1Q2017	Y_{t-3}	
Information criteria evaluate penalized model fit on data that is used to build the model.	4Q2016	Y_{t-4}	Fit Sample
	…	…	

Forecast in ARIMA

```
/* Identify, Estimate and then Forecast AR(1) Model */
Ods graphics on/imagemap=on;
Title 'Forecasting holdout sample on weekly solar power data';
proc ARIMA data=COURSE.Solarpv plots(only)= forecast(forecast forecastonly);
identify var=kW_Gen nlags=12;
estimate p=1 method= ML;
forecast lead=6 back=6 id=EDT out=work.AR1;
run;
ods graphics off;
```

4