KOF

Uses of Short Term-indicators for Forecasting

International Workshop Beijing

Klaus Abberger

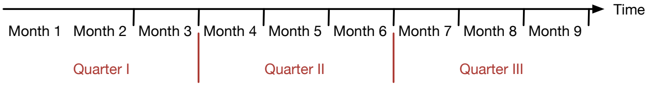
Swiss Economic Institute (KOF) of the ETH Zurich

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

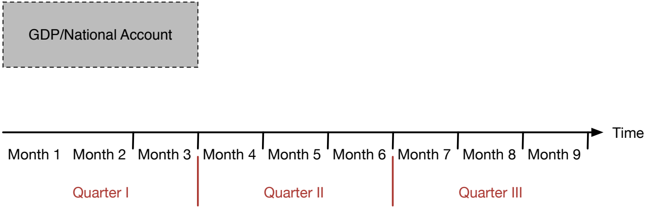
Outline

Outline

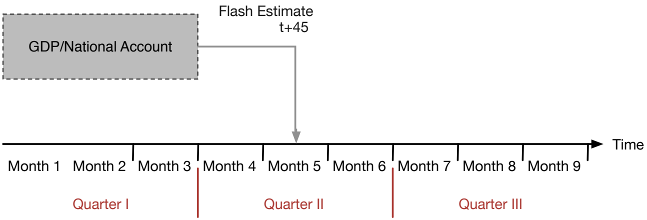
????casting


Time
Month 1
Month 2
Month 3
Month 4
Month 5
Month 6
Month 7
Month 8
Month 9
Quarter I
Quarter II
Quarter III

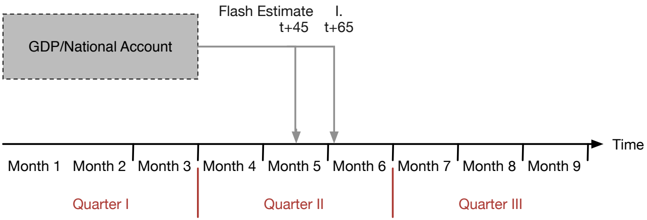
????casting


GDP/National Account
Time
Month 1
Month 2
Month 3
Month 4
Month 5
Month 6
Month 7
Month 8
Month 9
Quarter I
Quarter II
Quarter III

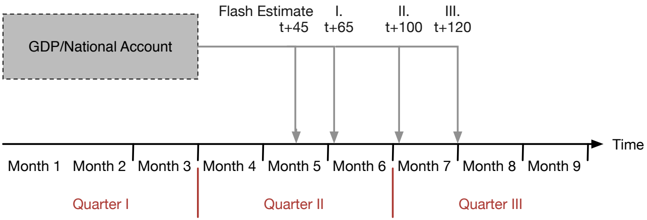
????casting


GDP/National Account
Flash Estimate
t+45
Time
Month 1
Month 2
Month 3
Month 4
Month 5
Month 6
Month 7
Month 8
Month 9
Quarter I
Quarter II
Quarter III

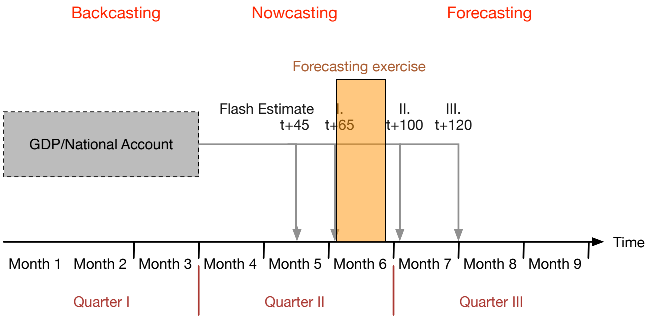
????casting


GDP/National Account
Flash Estimate
t+45
I.
t+65
Time
Month 1
Month 2
Month 3
Month 4
Month 5
Month 6
Month 7
Month 8
Month 9
Quarter I
Quarter II
Quarter III

????casting


GDP/National Account
Flash Estimate
t+45
I.
t+65
II.
t+100
III.
t+120
Time
Month 1
Month 2
Month 3
Month 4
Month 5
Month 6
Month 7
Month 8
Month 9
Quarter I
Quarter II
Quarter III

????casting


Backcasting
Nowcasting
Forecasting
Forecasting exercise
GDP/National Account
Flash Estimate
t+45
I.
t+65
II.
t+100
III.
t+120
Time
Month 1
Month 2
Month 3
Month 4
Month 5
Month 6
Month 7
Month 8
Month 9
Quarter I
Quarter II
Quarter III

Composite Indicators

- Leading indicators
- Coincident indicators (but publication lead)

Composite Indicators: Uses

- Nowcasting
- Short term forecasting (e.g. next quarter)
- (Backcasting, revision prediction)

GDP Vintages

BIP	2012Q2	2012Q3	2012Q4	2013Q1	2013Q2	2013Q3	2013Q4	2014Q1	2014Q2	2014Q3	2014Q4
2010Q1	0.99	0.99	0.99	0.99	1.07	1.06	1.04	1.04	0.91	0.92	0.92
2010Q2	0.90	0.90	0.89	0.88	0.83	0.83	0.82	0.82	0.89	0.90	0.90
2010Q3	0.70	0.70	0.69	0.69	0.63	0.62	0.59	0.60	0.36	0.34	0.34
2010Q4	0.97	0.96	1.00	1.01	0.94	0.96	1.02	1.02	0.86	0.87	0.86
2011Q1	0.25	0.29	0.28	0.26	0.28	0.28	0.27	0.26	0.38	0.38	0.39
2011Q2	0.50	0.50	0.50	0.52	0.41	0.40	0.38	0.38	0.56	0.57	0.57
2011Q3	-0.20	-0.25	-0.25	-0.24	-0.18	-0.19	-0.22	-0.21	-0.11	-0.14	-0.14
2011Q4	0.37	0.34	0.31	0.30	0.24	0.25	0.32	0.32	0.39	0.41	0.40
2012Q1	0.53	0.47	0.48	0.40	0.47	0.46	0.43	0.42	0.10	0.10	0.10
2012Q2	-0.06	-0.12	-0.11	-0.06	-0.04	-0.06	-0.08	-0.08	0.27	0.28	0.28
2012Q3	.	0.57	0.57	0.61	0.68	0.74	0.71	0.72	0.66	0.62	0.62
2012Q4	.	.	0.24	0.26	0.30	0.28	0.39	0.39	0.38	0.41	0.40
2013Q1	.	.	.	0.56	0.58	0.55	0.60	0.60	0.15	0.16	0.16
2013Q2	.	.	.	.	0.52	0.55	0.57	0.54	1.04	1.05	1.05
2013Q3	.	.	.	.	.	0.52	0.51	0.54	0.37	0.34	0.34
2013Q4	.	.	.	.	.	.	0.16	0.18	0.50	0.52	0.52
2014Q1	.	.	.	.	.	.	.	0.46	0.44	0.45	0.46
2014Q2	.	.	.	.	.	.	.	.	0.19	0.29	0.29
2014Q3	.	.	.	.	.	.	.	.	.	0.63	0.66
2014Q4	.	.	.	.	.	.	.	.	.	.	0.60

Target of forecasts

- Which frequency (quarterly growth, annual growth, quarterly growth as an instrument for annual growth where quarterly figures are only partly published)
- Which vintage/publication (first publication, last publication)

GDP Nowcasting at KOF

Target	Nowcast	Actual (SECO)[a]	Target	Nowcast	Actual (SECO)
2013Q1	0.55 [-0.18, 1.28]	**0.56**	2014Q1	0.69 [-0.03, 1.41]	**0.46**
Released on	01.03.2013	30.05.2013	Released on	04.03.2014	28.05.2014
2013Q2	0.50 [-0.21, 1.21]	**0.52**	2014Q2	0.56 [-0.15, 1.27]	**-0.04 (0.19)**[b]
Released on	03.06.2013	03.09.2013	Released on	29.05.2014	02(30).09.2014
2013Q3	0.44 [-0.27, 1.15]	**0.52**	2014Q3	0.39 [-0.33, 1.11]	**0.63**
Released on	03.09.2013	28.11.2013	Released on	02.09.2014	03.12.2014
2013Q4	0.55 [-0.18, 1.28]	**0.16**	2014Q4	0.28 [-0.41, 0.97]	**0.60**
Released on	28.11.2013	27.02.2014	Released on	03.12.2014	03.03.2015

Forecatsing system of KOF

- Macroeconomic model consisting of Error Correction Models and equations to ensure consistency. This model is used for quarterly forecasts of various variables up to two years. Consistency of the model forecasts is crucial.
- Nowcasting and short term forecasting with indicator model. These forecasts are used to adjust the model forecasts in the short term.

Composite Indicators

Some specific topics in quantification of composite indicators

- Target/reference series?
  - Which series (e.g. GDP)
  - Which transformation (e.g. annual growth, quarter-on-quarter growth)
  - Which release (first release, release after some revisions)
- Statistical methods (e.g. time series regressions)
- Frequency of indicators and target the same or different?

Outline

Autoregressive Distributed Lag Model

With Y_t the target variable at time t and IN_t an (composite) indicator at time t one can try to estimate the model

$$Y_t = \alpha + \sum_{i=1}^{m} \beta_i Y_{t-i} + \sum_{j=0}^{n} \gamma_j IN_{t-j} + \epsilon_t,$$

with error term ϵ.
This approach can be used to test whether the coefficients γ are significant different from zero. In addition this model can be used to calculate indicator based forecasts.

Error Correction Model

With Y_t the target variable at time t and X_t an independent variable. An Error Correction Model is

$$\Delta_1 Y_t = \alpha + \zeta Z_{t-1} + \sum_{i=1}^{m} \beta_i \Delta_1 Y_{t-i} + \sum_{j=0}^{n} \gamma_j \Delta_1 X_{t-j} + \epsilon_t,$$

with error term ϵ and

$$Z_t = Y_t - bX_t.$$

$\Delta_1 Y$ will respond negatively to Z since Z is the deviation from the long run equilibrium state ($Z = 0$). ζ gives the speed of adjustment to the equilibrium.

Error Correction Model

- Estimates short term and long term effects
- Applications to stationary and nonstationary data
- Estimation with OLS

Error Correction Model

- Motivating ECM with cointegrated data
  - Handles nonstationary cointegrated variables
  - Two step estimation procedure
  - Models with only differenced variables may ignore long term efects
- Motivating ECM with stationary data
  - Long and short term effects
  - ECM and ADL are equivalent

Mixed Frequencies

Often time series with different frequency are used for forecasting. So one may want forecast quarterly GDP with a monthly composite indicator. The objective is to forecast a

lower-frequency variable, Y, sampled at periods denoted by time index t. Past realizations of the lower-frequency variable are denoted by the lag operator, L. For example, if Y_t is the quarterly GDP, then the GDP one quarter prior would be the first lag of Y_t, $LY_t = Y_{t-1}$, two months prior would be $L^2 Y_t = Y_{t-2}$, and so on.

Mixed Frequencies

In addition to lags of Y, we are interested in the information content of a higher-frequency variable, X, sampled m times between samples of Y (e.g., between $t-1$ and t). L_{HF} denotes the lag operator for the higher-frequency variable. If X_t is the monthly indicator, then $L_{HF}X_t$ denotes the indicator value of the last month of the previous quarter.

Time Aggregation

One solution to the problem of mixed sampling frequencies is to convert higher-frequency data to match the sampling rate of the lower-frequency data. One can for example calculate averages of X between samples of the low frequency variable:

$$\bar{X}_t = \frac{1}{m} \sum_{k=1}^{m} L_{LH}^{k} X_t.$$

With the two variables in the same time domain on can use a regression model for forecasting:

$$Y_t = \alpha + \sum_{i=1}^{p} \beta_i L^i Y_t + \sum_{j=1}^{n} \gamma_j L^j \bar{X}_t + \epsilon_t$$

Time Aggregation

The standard aggregation methods depend on the stock/flow nature of the variables and, typically, it is the average of the high-frequency variables over one low-frequency period for stocks, and the sum for flows.

Taking the latest available value of the higher frequency variable is another option for both stock and flow variables. The underlying assumption is that the information of the previous high-frequency periods is reflected in the latest value, representative of the whole low-frequency period.

Bridge Models

With time aggregation often the problem occurs, that at a specific time point not all of the high-frequency data is already available. The Bridge Model Approach consist of two steps:

1. Forecasting the high-frequency series with time series methods (e.g. ARMA) for all time points which are needed for time aggregation.
2. An ADL model is used at the low frequency to obtain forecasts for Y.

Step Functions

The time aggregation method assumes the slope coefficients on each of the individual observations of X are equal. Alternatively, one could assume that each of the slope coefficients for each k sampling of X are unique. This model, including one lag of the predictor X ($n = 1$), is

$$Y_t = \alpha + \sum_{i=1}^{p} \beta_i L^i Y_t + \sum_{k=1}^{m} \gamma_k L_{HF}^k X_t + \epsilon_t.$$

The model can be estimated with OLS.

Step Functions

Once the model is extended to multiple lags, the number of parameters could become quite large (especially, when the high frequency data is of higher frequency than monthly). A more general model is

$$Y_t = \alpha + \sum_{i=1}^{p} \beta_i L^i Y_t + \sum_{k=1}^{m \cdot n} \gamma_k L_{HF}^k X_t + \epsilon_t$$

which allows for up to n lower-frequency lags.So the problem with this type of models is, that the number of coefficient can get huge.

MIDAS (Mixed Data Sampling)

The time-averaging model is parsimonious but discards any information about the timing of innovations to higher-frequency data. The step-weighting model preserves the timing information but requires the user to estimate a potentially large number of parameters.

MIDAS (Mixed Data Sampling)

To solve the problem of parameter proliferation while preserving some timing information MIDAS models can be used:

$$Y_t = \alpha + \sum_{i=1}^{p} \beta_i L^i Y_t + \sum_{k=1}^{m} \Phi(k;\theta) L_{HF}^k X_t + \epsilon_t.$$

where the function $\Phi(k;\theta)$ is a polynomial that determines the weights for temporal aggregation. The weighting function, $\Phi(k;\theta)$, can have any number of functional forms; the desire here is to achieve flexibility while maintaining parsimony.

MIDAS (Mixed Data Sampling)

Suggestions for $\Phi(k;\theta)$ are often a beta function formulation or an exponential Almond specification. The later is

$$\Phi(k;\theta_1,\theta_2) = \frac{exp(\theta_1 k + \theta_2 k^2)}{\sum_{j=1}^{m} exp(\theta_1 j + \theta_2 j^2)}$$

In this case, simple time averaging is obtained when $\theta_1 = \theta_2 = 0$.

Outline

Turning Points in the Business Cycle

Generally, practitioners in business cycle analysis sometimes assume that economic cycles are constituted by an alternation of two conjonctural phases, namely a phase of high economic activity (or expansion) and a phase of low economic activity (or contraction). These phases can be defined in classical, growth or growth rate cycles. Sometimes also or than two phases are considered.

The objective of parametric models is to provide, at each date t, an estimated probability of being in a specific phase.

Binary response models

If there is a reference series and a if the phases (dating) of the reference series are available a binary variable can defined that takes the value 1 when the economy belongs to one phase and 0 when it belongs to the other phase. This $0-1$ variable can be used for logit or probit regressions.

Logistic Regression

Let Y be a binary variable with values 0 and 1 and X a predictor (e.g. a composite indicator), the the logistic regression model (logit) is

$$Log\left[\frac{prob(Y_t = 1)}{1 - prob(Y_t = 1)}\right] = a + bx_t. \tag{1}$$

The model can be extended to contain lags of X and lags of Y.

Markov Switching

Markov switching models consist to the class of nonlinear time series models. They base on the idea of probability switching between various states (e.g. upswing and downswing). In the following Markov switching autoregressive models are discussed. Markov switching regression models use also explanatory variables.

Markov Switching

Hamilton (1989) considers the Markov switching autoregressive (MSA) model. Here the transition is driven by a two-state Markov chain. A time series x_t follows an MSA model if it satisfies:

$$x_t = \begin{cases} c_1 + \sum_{i=1}^{p} \phi_{1,i} x_{t-i} + a_{1,t} & \text{if } s_t = 1, \\ c_2 + \sum_{i=1}^{p} \phi_{2,i} x_{t-i} + a_{2,t} & \text{if } s_t = 2, \end{cases} \tag{2}$$

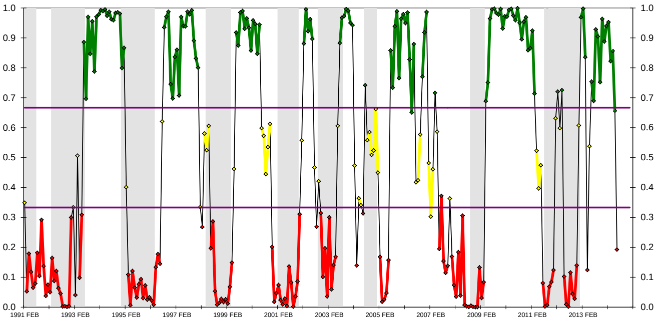
where s_t assumes values in $\{1, 2\}$ and is a first-order Markov chain with transition probabilities

$$P(s_t = 2 | s_{t-1} = 1) = w_1, \; P(s_t = 1 | s_{t-1} = 2) = w_2. \tag{3}$$

Markov Switching

The innovational series $\{a_{1,t}\}$ and $\{a_{2,t}\}$ are sequences of iid random variables with mean zero and finite variance and are independent of each other. A small w_i means that to model tends to stay longer in state i. In fact, $1/w_i$ is the expected duration of the process to stay in state i.

Ifo Business Cycle Traffic Lights



Source: Calculations of the Ifo Institute.

Outline

Notation

Let $\{y_t\}_{t=1}^{T}$ a time series being forecasted and $\{\hat{y}_{it}\}_{t=1}^{T}$ and $\{\hat{y}_{jt}\}_{t=1}^{T}$ two forecasts. Let the associated forecast errors be $\{e_{it}\}_{t=1}^{T}$ and $\{e_{jt}\}_{t=1}^{T}$.

Measures of prediction error

Mean prediction error:

$$MPE(\hat{y}_i) = \frac{1}{T}\sum_{t=1}^{T} e_{it}$$

Mean absolute prediction error:

$$MAPE(\hat{y}_i) = \frac{1}{T}\sum_{t=1}^{T} |e_{it}|$$

Mean squared prediction error:

$$MSPE(\hat{y}_i) = \frac{1}{T}\sum_{t=1}^{T} (e_{it})^2$$

GDP Forecasts fro Germany

Year	IMF	GD	EU	SVR	OECD	JWB	Realized
1991	3.3	3.0	3.1	3.0	3.0	3.0	3.4
1992	2.4	2.0	2.2	2.0	1.8	2.0	1.5
1993	2.6	0.5	-0.5	0.0	0.7	0.0	-1.9
1994	1.1	1.0	0.0	0.0	0.4	1.0	2.3
1995	2.7	2.5	3.0	3.0	2.8	3.0	1.9
1996	2.9	2.5	2.4	2.0	3.2	2.3	1.9
1997	2.3	2.5	2.2	2.5	2.2	2.5	2.2
1998	3.0	2.8	3.2	3.0	3.0	3.0	2.8
1999	2.6	2.3	2.2	2.0	2.2	2.8	1.5
2000	2.5	2.7	2.6	2.7	2.3	2.6	3.0
2001	3.2	2.7	2.8	2.8	2.7	2.8	0.6
2002	0.8	1.3	0.7	0.7	1.0	0.8	0.2
2003	2.0	1.4	1.4	1.0	1.5	1.1	-0.1
2004	2.0	1.7	1.6	1.5	1.4	1.8	1.5
2005	1.6	1.5	1.5	1.4	1.4	1.3	0.9
2006	1.0	1.2	1.2	1.0	1.6	1.6	2.7
2007	1.2	1.4	1.2	1.8	1.7	1.9	2.5
2008	2.3	2.2	2.1	1.9	2.1	1.7	1.3
2009	-0.1	0.2	0.0	0.0	-0.9	-2.0	-5.0
2010	0.4	1.2	1.2	1.6	1.6	1.4	3.6
2011	1.9	2.0	2.2	2.2	2.5	2.3	3.0
2012	1.2	0.8	0.8	0.9	0.4	0.7	0.7
2013	0.8	1.0	0.8	0.8	0.5	0.4	0.4
Average forecast horizon (days)	478	439	422	417	403	353	

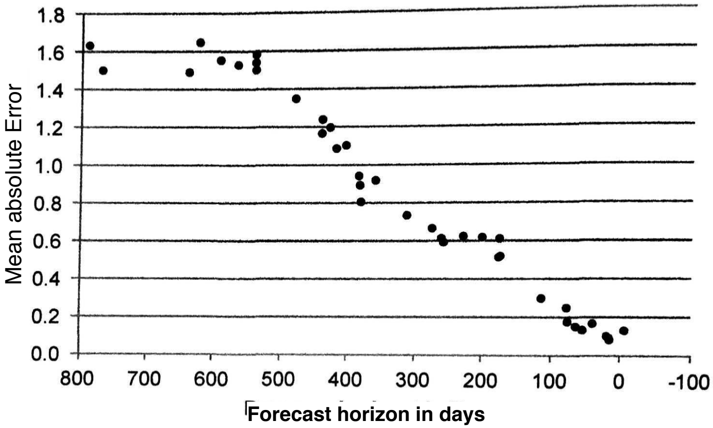
Source: Döhrn (2015)

GDP Forecasts fro Germany

	IMF	GD	EU	SVR	OECD	JWB
MPE	0.56	0.41	0.30	0.30	0.36	0.31
MAPE	1.37	1.13	1.11	1.03	1.06	0.91
MSPE	3.51	2.48	2.38	2.24	2.05	1.42

Source: Döhrn (2015)

GDP Forecasts fro Germany



Source: Döhrn (2015)

Measures of prediction error

Signal to noise ratio:

$$SNR(\hat{y}_i) = \frac{MSPE(\hat{y}_i)}{\sigma_y^2}$$

GDP Forecasts fro Germany

	Mean Squared Error	Noise-to-Signal-Ratio
Private Consumption	1.16	1.73
Public Consumption	0.91	0.84
Equipment Investments	47.70	0.77
Construction Investments	8.94	0.75
Imports	22.22	0.66
Exports	18.87	0.79
GDP	2.48	0.73

Source: Döhrn (2015)

Comparision of forecast errors

Theil Coefficient

$$Theil_{ij} = \frac{MSPE(\hat{y}_i)}{MSPE(\hat{y}_j)}$$

Denote the loss associated with forecast error e_t be $L(e_t)$; hence the time-t quadratic loss would be $L(e_t) = e_t^2$. The time-t loss differential between forecast i and j is then

$$d_{ijt} = L(e_{it}) - L(e_{jt})$$

and means of the loss differential can be calculated. Typical loss function are squared or absolute loss. But for example also asymmetric loss can be calculated.

Diebold Mariano Test

The Diebold Mariano test is an asymptotic test of the hypothesis that the mean of the loss differential is zero. In practice the test can be calculated by regression of the observed loss differential on an intercept, using heteroscedasticity and autocorrelation robust (HAC) standard errors for testing significance of the intercept.

One can potentially extend the regression to condition on additional variables that may explain the loss differential, thereby moving from an unconditional to a conditional expected loss perspective. For example, comparative predictive performance may differ by stage of the business cycle, in which case one might include a 0-1 business cycle chronology variable in the HAC regression.

Diebold Mariano Test

The common test takes the estimated the estimated intercept α and divide it by the HAC estimated standard deviation s_α. The test statistic is then

$$DM = \frac{\alpha}{s_\alpha}$$

This statistic is asymptotically standard normal distributed. In small small samples (few forecasts) the Null hypothesis is rejected to often, so an small sample correction can be used. With T the number of forecasts an h the forecast horizon the correction is

$$MDM = DM \cdot \sqrt{\frac{N + 1 - 2h + \frac{h(h-1)}{N}}{N}}$$

In addition it is proposed to use the t-distribution with $T - 1$ degrees of freedom to calculate critical values.

Measures of prediction error for a binary variable

When the target variable, r_t, is a binary indicator while the forecast is a probability of a state, p_t, similar techniques can be used as in the case of continuous variables.
The quadratic probability score is:

$$QPS(p) = \frac{1}{T}\sum_{t=1}^{T} 2(p_t - r_t)^2$$

QPS ranges between $[0, 2]$ with 0 perfect accuracy. A similar loss function that assigns more weight to larger forecast errors is the log probability score:

$$LPS(p) = \frac{1}{T}\sum_{t=1}^{T}((1 - r_t)\log(1 - p_t) + r_t \log p_t)$$

The range of LPS is $[0, \infty]$ with 0 perfect accuracy.

Measures of prediction error for a binary variable

Contingency tables can also be used for a descriptive evaluation of the methodology in the case of binary forecasts and outcomes. They provide a summary of the percentage of correct predictions, missed signals (e.g. no prediction of slowdown when it takes place), and false alarms.