



Generalitat de Catalunya  
**Institut d'Estadística de Catalunya**

## Métodos de nowcasting para la economía catalana

**Mònica Gasulla, Idescat**

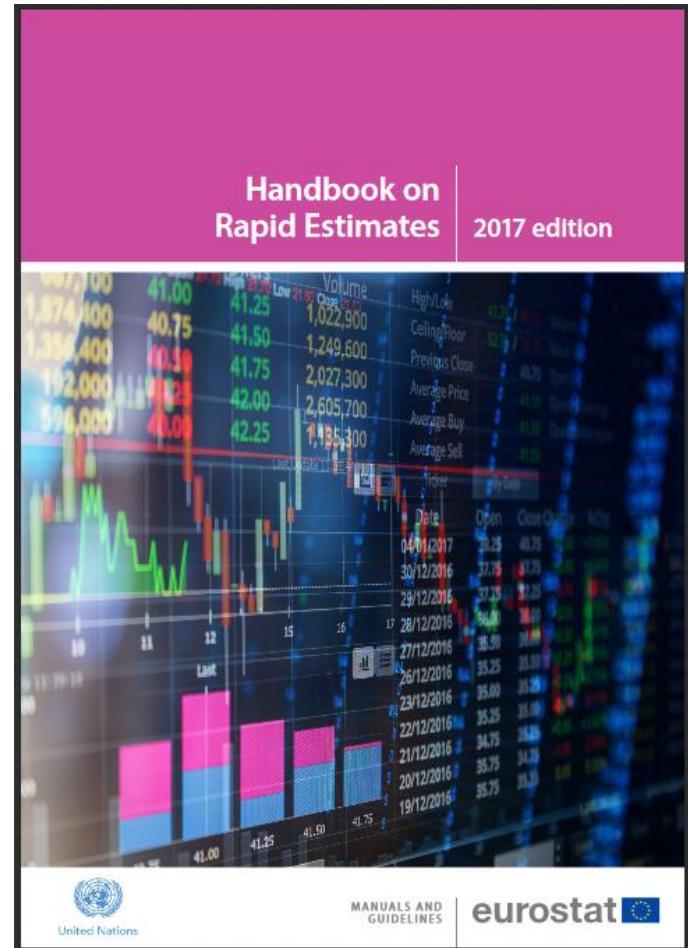
5 de setiembre de 2019

- 1. PIB en tiempo real (nowcasting)**
  - 1.1 Introducción**
  - 1.2 Tratamiento de datos**
  - 1.3 Selección de indicadores y análisis preliminar**
  - 1.4 Estimación de Modelos Factoriales Dinámicos (DFM)**
  - 1.5 Generación de predicciones**
  - 1.6 Combinación de predicciones**
- 2. Herramientas para el nowcasting**
- 3. Resultados: ejemplos**
- 4. Bibliografía y direcciones de interés**

# 1.1 Introducción

## Estimaciones NOWCASTING:

- Un **Nowcast** es una estimación en tiempo real de la evolución de una variable tanto **antes como después** de finalizar el trimestre de referencia.
- Se obtienen a partir de un conjunto de información incompleta formado por una mezcla de datos **cualitativos** y **cuantitativos**.
- Se basan en **métodos estadísticos** específicos que optimizan el uso de la información disponible.
- No permiten obtener una serie temporal sino una estimación específica para un **trimestre determinado**.



# 1.1 Introducción

## Estimaciones NOWCASTING:

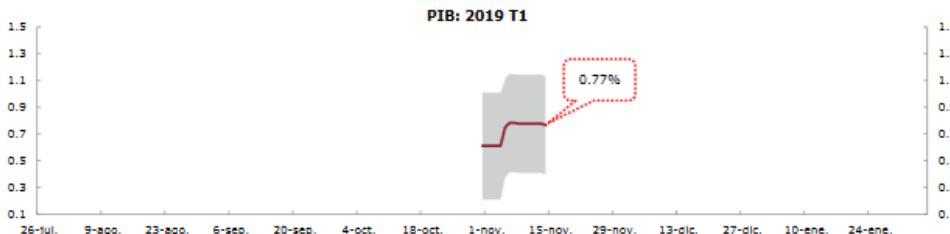
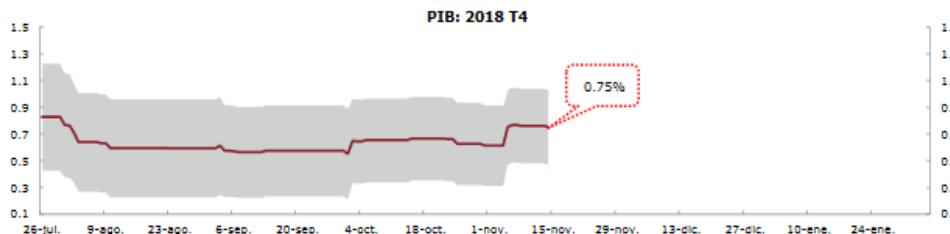
### AIReF – MIPred (Modelo Integrado de Predicciones)

#### Evolución de la Predicción del PIB en tiempo real

Período: 2018:TIV y 2019:TI

Datos de volumen CVEC<sup>(1)</sup> y Tasa intertrimestral

14 de noviembre de 2018



(1) Corregidos de variaciones estacionales y calendario

Fuente: AIReF



Autoridad Independiente  
de Responsabilidad Fiscal

Se ha actualizado la previsión con la  
información de los siguientes indicadores:

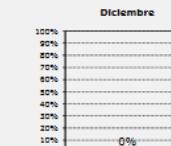
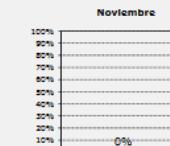
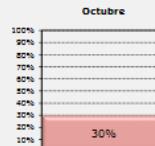
Indicador

Indicador	Periodo
VGE	Grandes empresas. Ventas totales deflactadas
RBT	Grandes empresas. Retribución bruta deflactada

Próximos indicadores:

Fecha	Indicador	Periodo
21-nov.	IMPB	septiembre
3-dic.	ELE	noviembre
3-dic.	FIN	octubre
4-dic.	AFI	noviembre
5-dic.	PMIS	noviembre

Proporción de los indicadores observados en el trimestre:



<http://www.airef.es/es/datalab/previsiones-del-pib-en-tiempo-real/>

# 1.1 Introducción

Eurostat ha elaborado un inventario **de modelos y métodos para el nowcasting**:

## Bridge models

- Ordinary least squares (OLS)
- Autoregressive Distributed Lags Model (ADL)
- Error Correction Model (ECM)
- Regression with ARIMA errors (RegARIMA)
- ARMA models with exogenous inputs (ARMAX)
- Transfer function models (TFM)

## Multivariate models

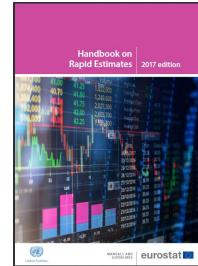
- Vector autoregressive models (VAR)
- Vector Error Correction Models (VECM)
- Global VAR (GVAR)
- Bayesian VAR (BVAR)

## Factor models

- Static Factor Models (SFM)
- Dynamic Factor Models (DFM)

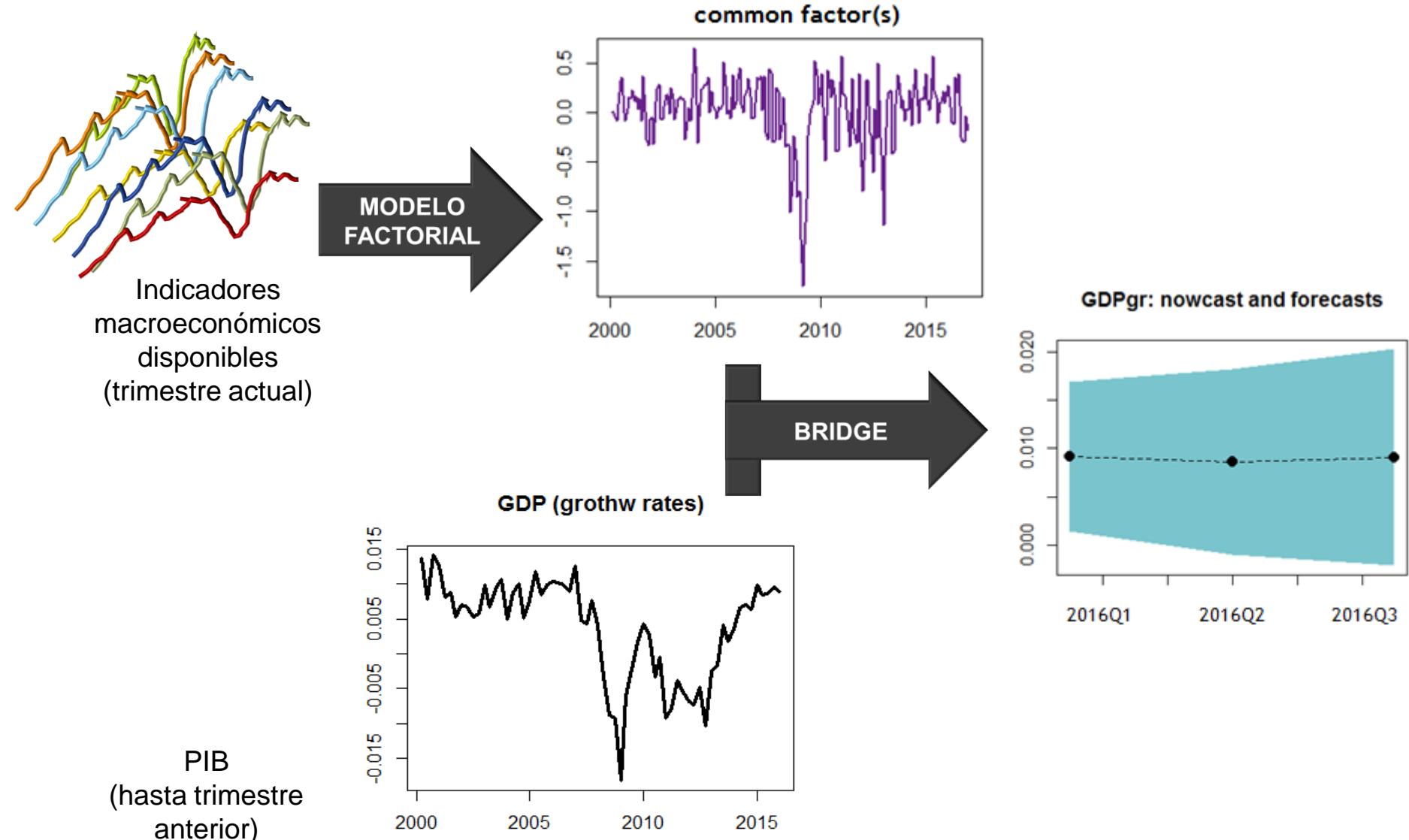
## Mixed-frequency methods

- Bridge regression
- Mixed-frequency VAR (MF-VAR)
- Bayesian mixed-frequency VAR (MF-BVAR)
- Mixed-frequency factor models (MF-FM)
- Mixed data sampling models (MIDAS)



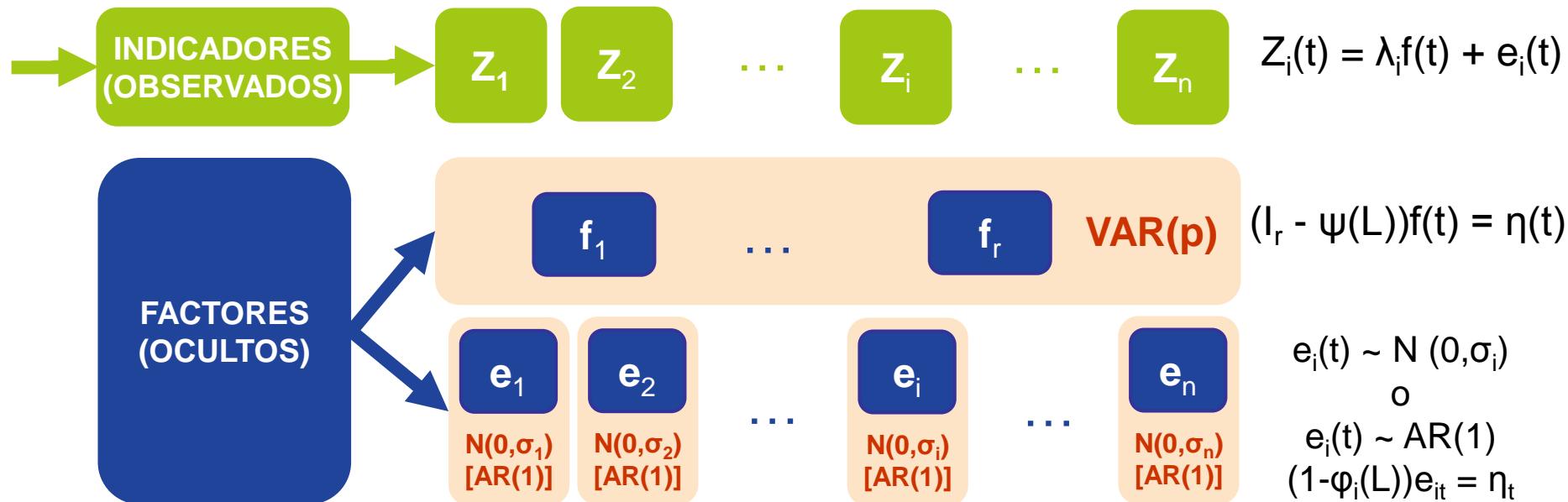
**Idescat**, a partir de la experiència de las Reservas Federales de Atlanta y de Nueva York y de la AIReF en España, **ha optado por desarrollar la metodología DFM** dada su solidez, consistencia y capacidad predictiva.

# 1.1 Introducción



# 1.1 Introducción

**Modelos factoriales:** dado un período  $t$ ,  $t=1, \dots, T$ , el valor del indicador estacionario  $Z_i(t)$  se obtiene como combinación lineal de un conjunto de factores comunes ( $f_1(t), \dots, f_r(t)$ ) distribuidos según un VAR(p) y un término idiosincrásico  $e_{it}$  generado por un ruido blanco o un proceso AR(1)

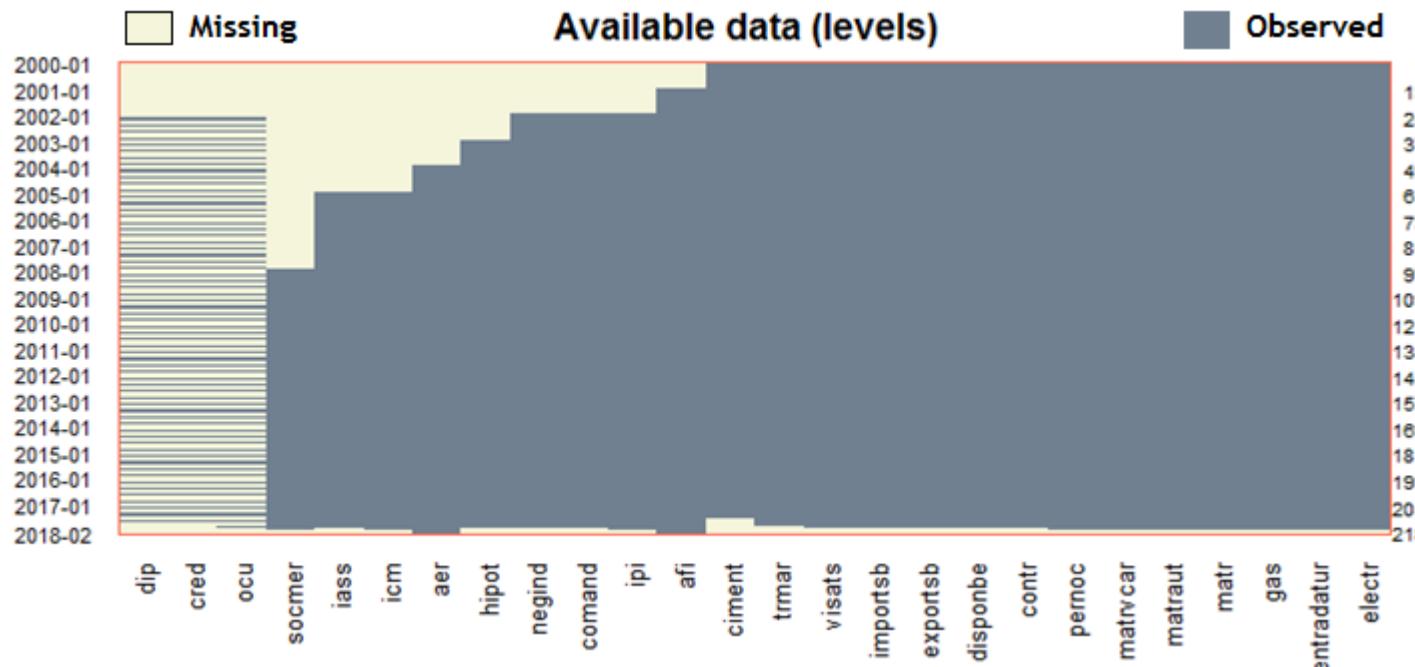


Modelo (uni)factorial:  $Z_{it} = \lambda_i f_t + e_{it}$   $i=1, \dots, n$   $t=1, \dots, T$   
Estático  $f \sim AR(p)$   $e_i \sim AR(1)$

# 1.1 Introducción

En la práctica, el reto de estimar el modelo factorial radica en la falta de estacionariedad y, sobre todo, en la naturaleza heterogénea de los indicadores:

- Diferencias en las fechas de inicio
- Diferencias en los calendarios de publicación
- Frecuencias mixtas



# 1.1 Introducción

**ALGORITMO EM\***: adecuado para la estimación de variables latentes con datos faltantes

- Dataset completo de indicadores mensuales estacionarios (tasas MoM)
- Estimación inicial de factores mediante un modelo factorial clásico
- Estimación inicial de la verosimilitud

**M step**

Imputación de datos mediante regresiones factores-indicador

**E step**

Estimación de los parámetros estáticos y dinámicos del modelo

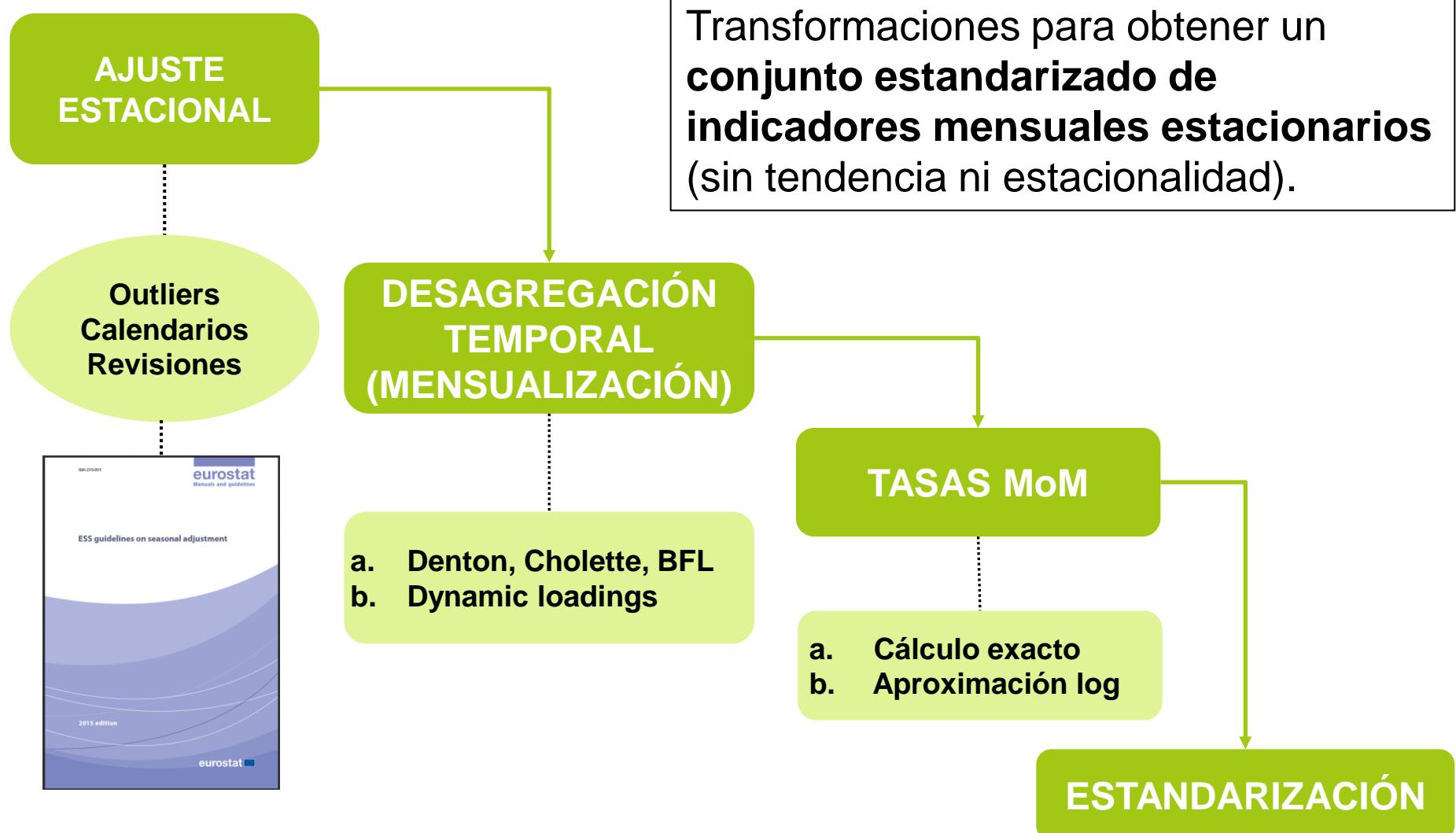
Formulación del modelo en un espacio de estados  
Filtro y suavizado de Kalman:

- Estimación de factores y generación de predicciones
- Evaluación de la verosimilitud

**ITERACIÓN HASTA ESTABILIZAR VERO SIMILITUD**

\* *Expectation-Maximization*

# 1.2 Tratamiento de datos



# 1.3 Selección de indicadores y análisis preliminar

Establecimiento de **criterios y parámetros** que pueden ser evaluados **antes** del algoritmo de estimación.

## SELECCIÓN DE INDICADORES

Correlaciones MoM, QoQ i YoY  
Capacidad predictiva

## IMPUTACIÓN VALORES

- a. Constantes: 0's, medias
- b. Modelos ARIMA
- c. Regresión ACP regularizada

## DETERMINACIÓN Nº FACTORES

Criterios visuales  
Criterios cuantitativos

## DATASET INICIAL

- a. Original
- b. Longitudinal
- c. Transversal
- d. Mixto

## DETERMINACIÓN ORDEN VAR (p)

Inspección correlograma  
Criterios BIC, AICC,...

# 1.4 Estimación de Modelos Factoriales Dinàmicos

Establecimiento de **criterios y parámetros** que sólo pueden ser evaluados **tras la ejecución** del algoritmo de estimación.

## MODELO FACTORIAL

- a. PCA
- b. ML

- 1. OLS
- 2. WLS
- 3. GLS

## TÉRMINOS IDIOSINCRÁSICOS (RESIDUOS)

- a.  $N(0, \sigma_i)$  independientes
- b. AR(1) independientes

## TRIMESTRALITZACIÓN FACTOR

- a. Mariano-Murasawa
- b. Medias trimestrales

# 1.4 Estimación de Modelos Factoriales Dinámicos

## MODELO FACTORIAL DINÁMICO

$$Z_{it} = \lambda_i f_t + e_{it}$$
$$(I_r - \Psi(L))f_t = \eta_t$$

$$(1 - \varphi_i L) e_{it} = u_t$$



## ESPACIO DE ESTADOS

$$Z_{it} = H X_t + W_t \quad \text{Eq de medida}$$
$$X_t = G X_{t-1} + V_t \quad \text{Eq de transición}$$

$$W_t \sim N(0, R) \quad V_t \sim N(0, Q)$$

## FACTORES IDIOSINCRÁSICOS $N(0, \sigma^2)$

$p$  estados

Matriz de transición  $p \times p$

## FACTORES IDIOSINCRÁSICOS AR(1)

$p+n$  estados

Matriz de transición  $(p+n) \times (p+n)$

$$X_t = (f_t, f_{t-1}, \dots, f_{t-p+1})$$

$$W_t = (e_{1t}, e_{2t}, \dots, e_{nt})$$

$$H = (\lambda_1, \lambda_2, \dots, \lambda_n) = \Lambda = \text{matriz de cargas}$$

$G$  = matriz basada en  $\Psi$

$V_t$  = shocks del modelo VAR( $p$ )

$$X_t = (f_t, f_{t-1}, \dots, f_{t-p+1}, e_1, e_2, \dots, e_n)$$

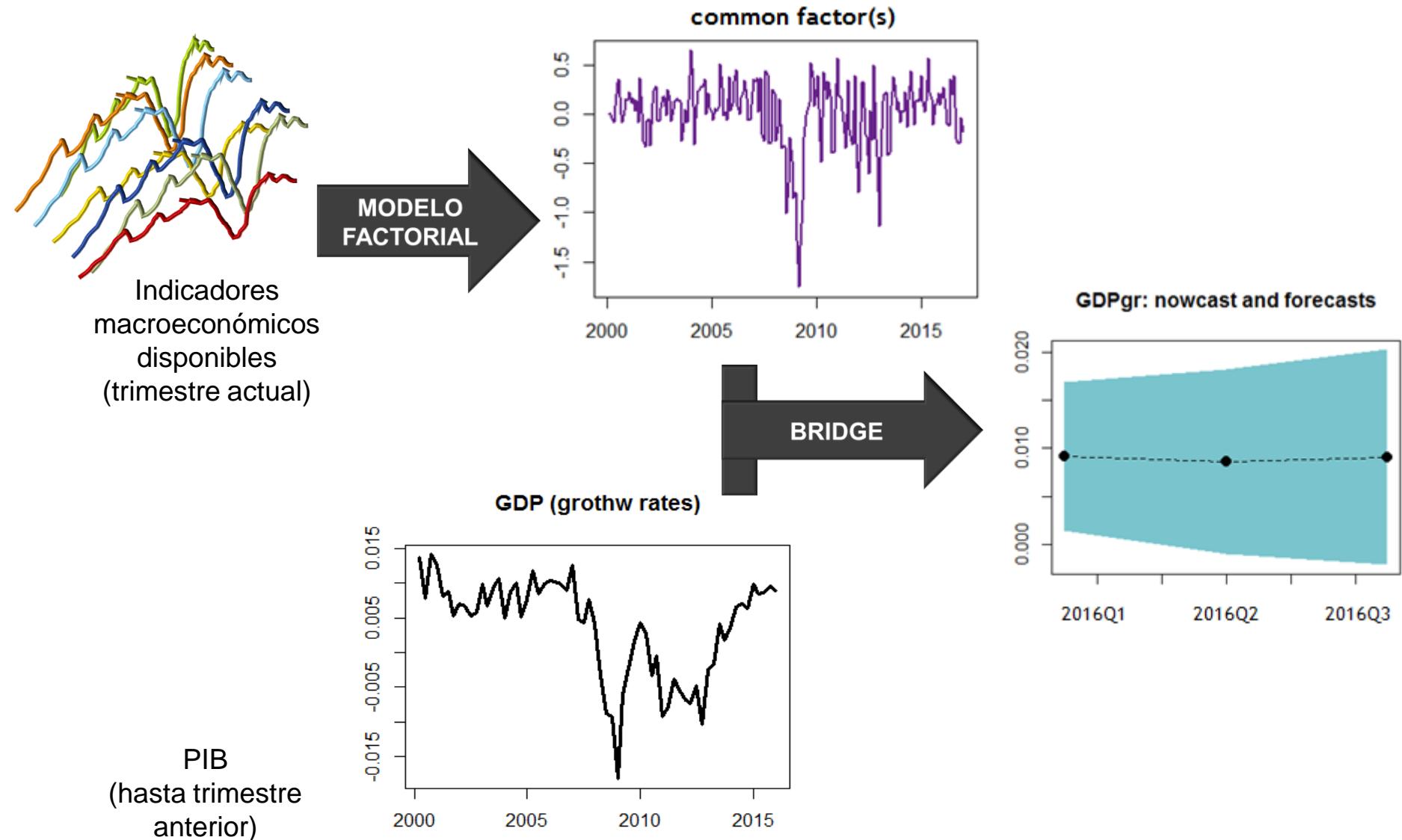
$$W_t = 0_n$$

$$H = (\Lambda \ 0_n \ I_n)$$

$G$  = matriz basada en  $\Psi$  y  $\varphi_i$

$V_t$  = shocks del VAR( $p$ ) y los AR(1)

# 1.5 Generación de predicciones (bridge)



# 1.5 Generación de predicciones (bridge)

**Aspectos posteriores a la estimación** del modelo factorial, relacionados con el acoplamiento de las tasas QoQ del PIB y el factor trimestralizado.

## CRITERIOS DE CALIDAD

Medidas de ajuste  
Diagnósticos residuos  
Estabilidad  
Medidas de precisión

## MODELOS PIB-FACTOR (BRIDGE)

- a. Integrado en el modelo factorial
- b. Reg-ARIMA
- c. Función de transferencia
- d. Reg-ARIMA con intervenciones
- e. Desagregación temporal
- f. ADL / VAR
- g. OLS

...

## OBTENCIÓN DE PREVISIONES FINALES

- a. Selección del “mejor” modelo
- b. Combinación de nowcasts

# 1.6 Combinación de predicciones

## COMBINACIÓN DE NOWCASTS

La estimación genera 12 predicciones en cada actualización del conjunto de indicadores, obtenidas mediante la ejecución de 4 modelos factoriales y 3 bridges.

Cuando no es fácil identificar un modelo óptimo, la agregación o combinación de predicciones permite generar un **único nowcast**. Para este proceso, se puede tener en cuenta la capacidad predictiva de los bridges en 1 año (expresada en términos del **RMSE** = **Error cuadrático medio de predicción**):

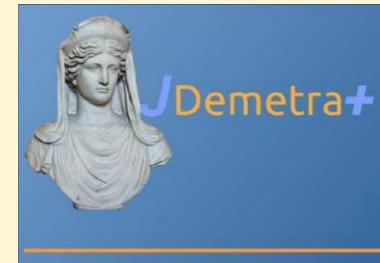
- **Media y media recortada basada en los 4 modelos más predictivos**
- **Mediana**
- **Media ponderada por la capacidad predictiva, en valor o ranking**

25-gen	OCU +				factor.model	gdp.model	2017Q4	2018Q1	RMSE
model	AICc	BIC.rARIMA	BIC.TSO	BIC.TFM					
FM1	5735.1	21.6	1.8	19.8	FM1	rARIMA	0.73	0.72	0.125
FM2	4831.0	19.3	14.8	23.2	FM1	TSO	0.72	0.62	0.117
FM3	4486.7	16.4	8.7	36.2	FM2	TFM	0.84	0.75	0.070
FM4	4501.9	28.9	28.9	14.3	FM2	rARIMA	0.81	0.85	0.137
					FM2	TSO	0.77	0.78	0.156
					FM2	TFM	0.84	0.88	0.061
method	2017Q4	2018Q1			FM3	rARIMA	0.78	0.87	0.182
mean	0.799	0.788			FM3	TSO	0.80	0.87	0.176
trmean.acc	0.790	0.776			FM3	TFM	0.80	0.85	0.074
median	0.807	0.775			FM4	rARIMA	0.85	0.75	0.359
wmean.rank	0.808	0.789			FM4	TSO	0.85	0.75	0.359
wmean.rmse	0.812	0.789			FM4	TFM	0.82	0.77	0.032

## 2. Herramientas para el nowcasting

**Idescat** ha desarrollado todo el proceso de nowcasting en **R**.

- **Descarga de datos**: paquetes RODBC, jsonlite y httr.
- **Ajuste estacional y procesamiento de indicadores**: extensiones propias de los paquetes RJDemetra, JDLight y otras R-interfaces de JDemetra+; tempdisagg para la desagregación temporal y missMDA para imputar valores faltantes.
- Suite de **modelos factoriales**: paquetes KFAS, MARSS, nowcasting, phtt y nlme.
- Suite de **modelos bridge**: paquetes forecast, TSO y extensiones propias de MTS y signal.



# 2. Herramientas para el nowcasting

## Proceso periódico de carga y actualización de indicadores

Una única interfaz de acceso a datos remotos formada por un conjunto de conectores:

- Acceso a bases de datos externas (HTTP i JSON)
- Acceso a base de datos internas (ODBC)

 PIB flashCAT20181113.xlsx
 PIB flashCAT20181107.xlsx
 PIB flashCAT20181108.xlsx
 PIB flashCAT20181106.xlsx
 PIB flashCAT20181105.xlsx
 PIB flashCAT20181029.xlsx
 PIB flashCAT20181026.xlsx
 PIB flashCAT20181024.xlsx
 PIB flashCAT20181002.xlsx
 PIB flashCAT20180921.xlsx
 PIB flashCAT20180904.xlsx
 PIB flashCAT20180903.xlsx
 PIB flashCAT20180828.xlsx
 PIB flashCAT20180813.xlsx
 PIB flashCAT20180719.xlsx
 PIB flashCAT20180717.xlsx
 PIB flashCAT20180713.xlsx
 PIB flashCAT20180709.xlsx
 PIB flashCAT20180705.xlsx
 PIB flashCAT20180627.xlsx
 PIB flashCAT20180621.xlsx

periode	gas_cat	trmar_cat	visats_rav	disponbe	disponbe	disponbe
1/1/2000	453176.4	4839833	1050308	111.3689	112.5507	127.4369
2/1/2000	451387.9	4428635	1428514	97.12452	105.3063	149.7087
3/1/2000	496193.9	4814580	1482708	104.9584	109.541	163.1508

1	A	B	C	D	E	F	G	H
periode	trimestre	cred_cat_pc	dip_cat_pc	ocu_cat	pib_cat_rate	cred_cat	dip_cat	
2	1/1/2000	2000 01	35307088	85132048		0.7		
3	4/1/2000	2000 02	102026305	85132048		0.7		
4	7/1/2000	2000 03	102865732	86252549		0.4		
5	10/1/2000	2000 04	1012838472	83962397		1.4		
6	1/1/2001	2001 01	109638702	89255417	2783.316119	0.6		
7	4/1/2001	2001 02	115676743	93728503	2840.27065	0.8		
8	7/1/2001	2001 03	11257350	93113683	2857.42857	0.8		
9	10/1/2001	2001 04	116338654	93521259	2822.37543	0.4		
10	1/1/2002	2002 01	121805442	93783153	2826.03938	0.6	1636803.111	1309011.486
J1	4/1/2002	2002 02	126517436	105123719	2903.8244	0.3	1730543.638	1433463.278
12	7/1/2002	2002 03	128787413	106352489	2954.3114	0.3	175856.06	1452221.563
13	10/1/2002	2002 04	131323449	106160788	2933.3142	0.5	176454.627	1426451.473
14	1/1/2003	2003 01	134843238	1053446534	2981.0764	1.3	1802782.158	1416423.411
15	4/1/2003	2003 02	144503145	108301913	3023.22934	0.7	1960319.018	1423193.653
16	7/1/2003	2003 03	149553417	107603935	3101.4474	0.7	1975628.793	1463144.297
17	10/1/2003	2003 04	156563702	115713939	320.6038	1.4	2038031.628	1436363.185
18	1/1/2004	2004 01	164766621	118768955	3105.1804	0.8	214653.331	1547550.634
19	4/1/2004	2004 02	171446659	125775740	3165.37938	1.2	225546.032	163726.3.027
20	7/1/2004	2004 03	182760813	123533128	3206.62884	0.9	232562.27	1648381.104
21	10/1/2004	2004 04	187026314	133444118	3205.26164	0.7	2345300.663	1673721.206
22	1/1/2005	2005 01	196579477	137286155	3291.1276	1.1	2472551.236	1723300.564
23	4/1/2005	2005 02	210771707	132293832	3363.6014	1	2585806.903	1632018.145
24	7/1/2005	2005 03	217857142	134382093	3401.13	1	2666027.375	1644501.615
25	10/1/2005	2005 04	231010231	141241848	3430.2388	1	278355.795	1701383.711
26	1/1/2006	2006 01	243242139	144155174	3454.5044	1.1	2927231.624	1730547.584
27	4/1/2006	2006 02	255938806	150560802	3487.873	1	3052371.132	177553.853
28	7/1/2006	2006 03	261972370	155728189	3505.1618	0.8	3177745.352	1838455.34
29	10/1/2006	2006 04	281673238	165029434	3515.364	1.4	3305533.151	1936617.262
30	1/1/2007	2007 01	294762977	170635893	3519.4028	0.8	3453442.801	2007666.097
31	4/1/2007	2007 02	307036342	172588287	3563.562	0.8	3525646.164	1381920.584
32	7/1/2007	2007 03	318957642	172826740	3585.3274	0.1	3671032.307	182882.431
33	10/1/2007	2007 04	325244507	175826575	3621.338	0.9	3665371.784	1826252.384
34	1/1/2008	2008 01	334382392	178543647	3608.7178	0.1	374182.372	2006142.873
35	4/1/2008	2008 02	341757552	180009685	3637.1724	-1	374587.214	2025152.234
36	7/1/2008	2008 03	348246194	183704734	3581.1208	-0.5	3823407.974	2082715.158
37	10/1/2008	2008 04	350226002	1933614569	3492.3524	-2.2	3848154.352	212793.23
38	1/1/2009	2009 01	352506181	194536140	3323.306	-1.4	39280395	1653323.026
39	4/1/2009	2009 02	362133551	200712848	3288.586	-1.3	388374.504	2202081.165
40	7/1/2009	2009 03	365539375	197827844	3304.8404	0.4	3864648.135	1846447.528
41	10/1/2009	2009 04	356464283	202365480	3243.3168	0.3	3870124.841	2124362.798
42	1/1/2010	2010 01	356061024	197170473	3210.9472	0.2	3818063.924	172319.282
43	4/1/2010	2010 02	360101923	207531332	3255.5568	0.7	383154.486	2242733.366
44	7/1/2010	2010 03	375755567	208719222	3287.6384	-0.4	3874193.851	1826703.201
45	10/1/2010	2010 04	37276815	206317228	3242.334	0.3	3860460.721	2136530.637
46	1/1/2011	2011 01	358350758	2132.374	-0.1	-0.1	3757043.493	2132348.178
47	4/1/2011	2011 02	347425661	202068571	3265.3702	-0.2	3626537.983	210436.362
48	7/1/2011	2011 03	34573279	196.94047	3213.53	-0.2	3576608.647	2062157.588
49	10/1/2011	2011 04	34473653	194332820	3136.583	-1.2	3586015.556	20124110.98
50	1/1/2012	2012 01	344436177	189338837	3072.495	-1.1	357745.207	182652.631
51	4/1/2012	2012 02	335515147	187374551	3053.526	-1.3	346545.621	181872.378
52	7/1/2012	2012 03	332543890	186861922	3032.6348	-0.4	3374193.851	1835934.536
53	10/1/2012	2012 04	316398603	181756627	2962.0524	-0.6	3167088.137	1812002.243
54	1/1/2013	2013 01	303147356	186705831	2933.185	-0.5	311125.503	193693.947
55	4/1/2013	2013 02	302162550	186288487	2940.6358	0.1	301760.103	187918.619
56	7/1/2013	2013 03	297005157	184376035	3005.7462	0.9	2970837.503	1843233.656
57	10/1/2013	2013 04	297061756	201003360	2998.7552	0.1	2958172.725	2001517.616
58	1/1/2014	2014 01	292224132	185505889	2960.6348	0.4	2935057.731	185913.777
59	4/1/2014	2014 02	283463939	187374283	3033.39882	0.7	2875523.775	1967056.089
60	7/1/2014	2014 03	284023044	186652394	3074.7526	0.8	284175.322	1966713.957
61	10/1/2014	2014 04	27715050	181701773	3048.0564	0.3	273186.637	1966500.277
62	1/1/2015	2015 01	276105452	184463755	3023.207	1	273181.343	1966353.668
63	4/1/2015	2015 02	271213334	186401666	3075.0348	1.1	270184.333	195364.289
64	7/1/2015	2015 03	267619776	185127173	310.858	0.9	2678039.21	1952604.139
65	10/1/2015	2015 04	264855913	181376552	310.8554	1	2642927.648	192678.198
66	1/1/2016	2016 01	247230396	180348495	3127.5612	0.9	2509428.681	183046.53
67	4/1/2016	2016 02	233804565	185207675	3181.2828	0.9	237478.247	181619.433

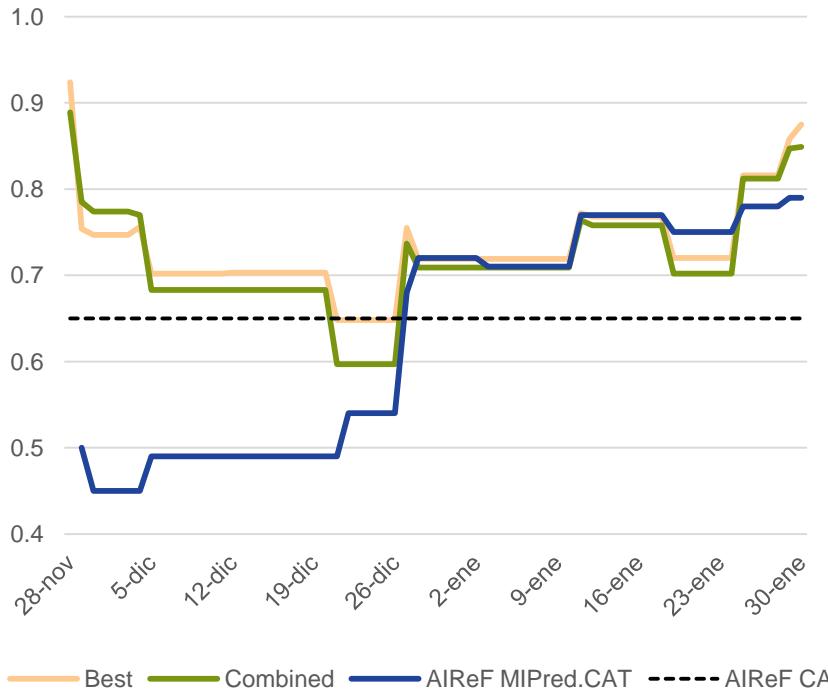
### 3. Resultados: ejemplos

#### EJEMPLO: 4º TRIMESTRE DEL 2017

**MIPred.CAT**: Modelo AIReF para Catalunya (estimador del PIB METCAP)

**NOW.PIB.CAT**: Modelo Idescat para Catalunya (estimador del PIB propio)

**MIPred.CAT - 2017Q4**



**NOW.PIB.CAT - 2017Q4**



### 3. Resultados: ejemplos

#### EJEMPLO : 4º TRIMESTRE DEL 2018

##### Impacto de la actualización del número de afiliados medios del mes de octubre:

29-oct				
EPA+	factor.model	bridge	2018Q4	RMSE
	FM1	rARIMA	0.45	0.11
	FM1	TSO	0.49	0.12
	FM1	TFM	0.52	0.15
	FM2	rARIMA	0.64	0.10
	FM2	TSO	0.64	0.10
	FM2	TFM	0.76	0.18
	FM3	rARIMA	0.53	0.12
	FM3	TSO	0.51	0.11
	FM3	TFM	0.64	0.18
	FM4	rARIMA	0.44	0.12
method	2018Q4	FM4	TSO	0.49
mean	0.548	FM4	TFM	0.50
trmean.acc	0.521			
median	0.514			
wmean.rank	0.570			
wmean.rmse	0.538			

05-nov				
AFI+	factor.model	bridge	2018Q4	RMSE
	FM1	rARIMA	0.66	0.11
	FM1	TSO	0.69	0.12
	FM1	TFM	0.75	0.14
	FM2	rARIMA	0.80	0.09
	FM2	TSO	0.77	0.08
	FM2	TFM	0.95	0.17
	FM3	rARIMA	0.55	0.12
	FM3	TSO	0.53	0.11
	FM3	TFM	0.66	0.18
	FM4	rARIMA	0.65	0.11
method	2018Q4	FM4	TSO	0.67
mean	0.698	FM4	TFM	0.71
trmean.acc	0.667			
median	0.680			
wmean.rank	0.720			
wmean.rmse	0.696			

	NEW	YoY	scaled YoY	PoP	scaled PoP	2018Q4
29-oct-2018	EPA	2.34	0.33	0.37	0.09	0.538
05-nov-2018	AFI	2.80	0.57	0.45	1.23	0.696

## 4. Bibliografía y enlaces de interés

- Eurostat (2017). Handbook on Rapid Estimates.
- Eurostat (2017). Handbook on Cyclical Composite Indicators.
- Eurostat (2016). Overview of GDP flash estimation methods.
- Cuevas A, Pérez-Quirós G, Quilis E (2015). Integrated model of short-term forecasting of the Spanish economy (MIPred model).
- Cuevas A, Quilis E (2009). A factor analysis for the Spanish economy (FASE). Mineco, Ministerio de Economía y Hacienda, Madrid.

## 4. Bibliografía y enlaces de interés

- JDemetra+ : <https://github.com/jdemetra/jdemetra-app/releases>
- JDLight package: <https://github.com/nbbrd/jdemetra-jdlight>
- rJDemetra package: <https://github.com/nbbrd/rjdemetra>
- nowcasting package: <https://cran.r-project.org/web/packages/nowcasting/index.html>
- Federal Reserve Bank of New York (2017). Macroeconomic Nowcasting and Forecasting with Big Data. MATLAB code available at <https://github.com/FRBNY-TimeSeriesAnalysis/Nowcasting>
- MATLAB - repositorio E. Quilis:  
<https://www.mathworks.com/matlabcentral/profile/authors/1008800-enrique-m-quilis>

# Gracias



[www.idescat.cat](http://www.idescat.cat)



<https://twitter.com/idescat>

