

Using the state-space framework of JDemetra+ in R

Jean Palate
David de Antonio Liedo
Raf Baeyens

National Bank of Belgium, R&D Unit (Statistics)

Table of contents

1. Motivations

2. RJDSSF

- Overview
- Model definition
- Model estimation

3. Examples

- STS with time varying trading days
- Labor force survey with rotating panels
- Unobserved components model for inflation forecasting

4. Future extensions

5. Diffusion

- ▶ JD(emetra)+
 - ▶ Java libraries for seasonal adjustment (and other time series methods)
 - ▶ Huge use of state-space models $\implies$ Rich framework (OO design, fast algorithms ...)
 - ▶ Main statistical references: **Durbin and Koopman**, Anderson and Moore, De Jong, Harvey, Kailath et al., Proietti ...
- ▶ Reuse of the SSF by statisticians
 - ▶ Java too complex $\implies$ R solution
- ▶ Existing R packages on state-spaces
 - ▶ **KFAS**, MARSS, dlm ...
 - ▶ But ...
 - ▶ Not always simple (definition, estimation ...)
 - ▶ Need for a common Java/R implementation
 - ▶ $\implies$ RJDSSF

- ▶ Univariate models
 - ▶ Advanced SA decomposition: time varying regression, seasonal specific models
 - ▶ Temporal disaggregation (mixed frequencies)
- ▶ Multivariate models
 - ▶ SUTSE models
 - ▶ Labor force survey with rotating panels (ABS, ONS)
 - ▶ Forecasting inflation (R&D + research department of NBB)

▶ Main features

- ▶ Automatic creation of the model (no need to define the matrices of the model)
- ▶ Automatic identification and transformation of the parameters for ML estimation
- ▶ Rich results
- ▶ Common framework for a large set of (uni/multivariate) models

▶ Advantages

- ▶ Power and speed of the JD+ framework
- ▶ Most technical details are hidden

▶ Drawbacks

- ▶ Designed for a special class of models
- ▶ Limited to the pre-specified components (but extensible solution in Java)

RJDSSF: Model definition

The state-space models considered in RJDSSF are defined as the concatenation of independent state blocks, on which we can apply various measurements

- ▶ State array

$$\alpha_t = (\alpha'_{1,t} \quad \cdots \quad \alpha'_{n,t})' \quad (1)$$

- ▶ Dynamics

$$\alpha_{i,t+1} = T_{i,t}\alpha_{i,t} + \mu_{i,t}, \mu_{i,t} \sim N(0, [\sigma^2]V_{i,t}) \quad (2)$$

The transition matrices and the covariances of the innovations are block diagonal matrices

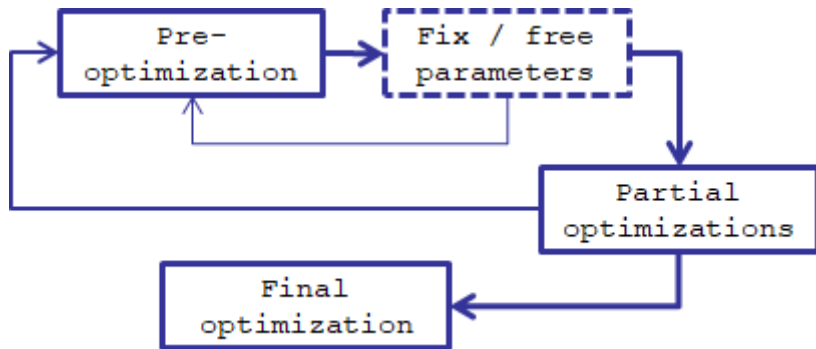
- ▶ Measurements

$$y_t = Z_t\alpha_t + \epsilon_t, \epsilon_t \sim N(0, [\sigma^2]D_t) \quad (3)$$

y_t univariate or multivariate, measurements across the blocks, independent measurement errors

RJDSSF: model estimation

- ▶ Estimation by ML
 - ▶ Transformation of the parameters $\implies$ better properties for (quasi-)Newton methods
- ▶ Complex iterative processing (taking into account the properties of the parameters and the limits of their domain)

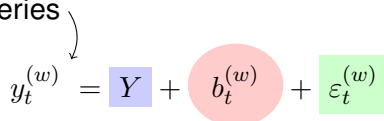


Example 1: STS with time varying trading days

```
# Usual BSM with time varying trading days
bsm<-function(s, seasonal="HarrisonStevens", tdgroups=c(1,2,3,4,5,6,0)){
  # create the model
  bsm<-jd3_ssf_model()
  # create the components and add them to the model
  add(bsm, jd3_ssf_locallineartrend("ll"))
  add(bsm, jd3_ssf_seasonal("s", frequency(s), type=seasonal))
  add(bsm, jd3_ssf_td("td", frequency(s), start(s), length(s), tdgroups))
  add(bsm, jd3_ssf_noise("n"))
  # create the equation
  eq<-jd3_ssf_equation("eq")
  add(eq, "ll")
  add(eq, "s")
  add(eq, "td")
  add(eq, "n")
  add(bsm, eq)
  #estimate the model
  rslt<-estimate(bsm, s, marginal=F, concentrated=T)
  return (rslt)
}
```


Example 2: Surveys with rotating panels

- ▶ Wave specific series

$$y_t^{(w)} = Y + b_t^{(w)} + \varepsilon_t^{(w)}$$
The equation $y_t^{(w)} = Y + b_t^{(w)} + \varepsilon_t^{(w)}$ is displayed. A curved arrow points from the text 'Wave specific series' to the term $y_t^{(w)}$. The term Y is enclosed in a light blue square. The term $b_t^{(w)}$ is enclosed in a light red circle. The term $\varepsilon_t^{(w)}$ is enclosed in a light green square.

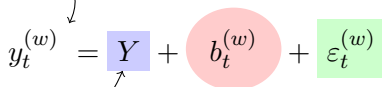
- ▶ Common process
- ▶ Wave bias
- ▶ Wave specific survey error

Y follows (for instance) a basic structural model (ARIMA modelling could also be used)

$\varepsilon_t^{(w)}$ follow a complex multi-variate auto-regressive process

Example 2: Surveys with rotating panels

- ▶ Wave specific series

$$y_t^{(w)} = Y + b_t^{(w)} + \varepsilon_t^{(w)}$$
The equation $y_t^{(w)} = Y + b_t^{(w)} + \varepsilon_t^{(w)}$ is displayed. The term Y is enclosed in a light blue square, $b_t^{(w)}$ is in a light red circle, and $\varepsilon_t^{(w)}$ is in a light green square. A curved arrow points from the text 'Wave specific series' to the $y_t^{(w)}$ term, and another curved arrow points from 'Common process' to the Y term.

- ▶ Common process
- ▶ Wave bias
- ▶ Wave specific survey error

Y follows (for instance) a basic structural model (ARIMA modelling could also be used)

$\varepsilon_t^{(w)}$ follow a complex multi-variate auto-regressive process

Example 2: Surveys with rotating panels

- ▶ Wave specific series

$$y_t^{(w)} = Y + b_t^{(w)} + \varepsilon_t^{(w)}$$

The diagram illustrates the decomposition of the wave-specific series $y_t^{(w)}$ into three components: a common process Y , a wave bias $b_t^{(w)}$, and a wave-specific survey error $\varepsilon_t^{(w)}$. The equation is $y_t^{(w)} = Y + b_t^{(w)} + \varepsilon_t^{(w)}$. The term Y is in a blue square, $b_t^{(w)}$ is in a red circle, and $\varepsilon_t^{(w)}$ is in a green square. Arrows indicate the relationships: a curved arrow from 'Wave specific series' points to $y_t^{(w)}$; a curved arrow from 'Common process' points to Y ; and a curved arrow from 'Wave bias' points to $b_t^{(w)}$.

- ▶ Common process
- ▶ Wave bias
- ▶ Wave specific survey error

Y follows (for instance) a basic structural model (ARIMA modelling could also be used)

$\varepsilon_t^{(w)}$ follow a complex multi-variate auto-regressive process

Example 2: Surveys with rotating panels

- ▶ Wave specific series

$$y_t^{(w)} = Y + b_t^{(w)} + \varepsilon_t^{(w)}$$

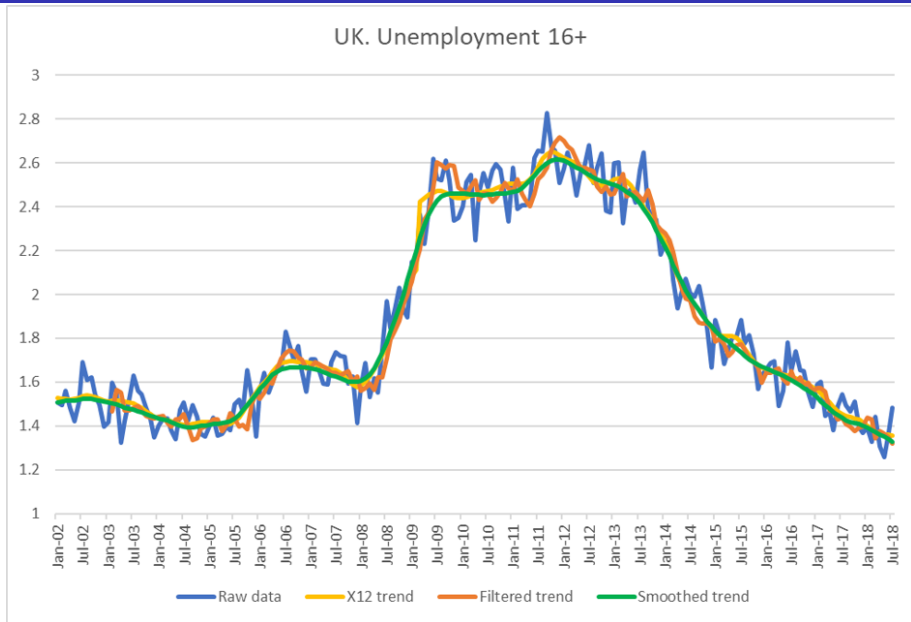
The diagram illustrates the decomposition of the wave-specific series $y_t^{(w)}$ into three components: a common process Y , a wave bias $b_t^{(w)}$, and a wave-specific survey error $\varepsilon_t^{(w)}$. The equation is $y_t^{(w)} = Y + b_t^{(w)} + \varepsilon_t^{(w)}$. The term Y is in a blue box, $b_t^{(w)}$ is in a red circle, and $\varepsilon_t^{(w)}$ is in a green box. Arrows indicate the relationships: a curved arrow from the text 'Wave specific series' points to $y_t^{(w)}$; a curved arrow from 'Common process' points to Y ; a curved arrow from 'Wave bias' points to $b_t^{(w)}$; and a curved arrow from 'Wave specific survey error' points to $\varepsilon_t^{(w)}$.

- ▶ Common process
- ▶ Wave bias
- ▶ Wave specific survey error

Y follows (for instance) a basic structural model (ARIMA modelling could also be used)

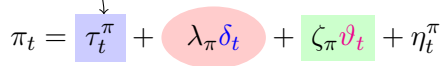
$\varepsilon_t^{(w)}$ follow a complex multi-variate auto-regressive process

Example 2: Labour force survey (ONS)



Example 3: Modelling inflation (UCM)

- ▶ Trend component



The equation $\pi_t = \tau_t^\pi + \lambda_\pi \delta_t + \zeta_\pi \vartheta_t + \eta_t^\pi$ is shown. An arrow points from the text 'Trend component' to the term τ_t^π , which is enclosed in a blue square. The term $\lambda_\pi \delta_t$ is enclosed in a red oval, and the term $\zeta_\pi \vartheta_t$ is enclosed in a green square.

$$\pi_t = \tau_t^\pi + \lambda_\pi \delta_t + \zeta_\pi \vartheta_t + \eta_t^\pi$$

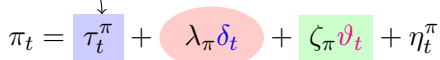
- ▶ Cyclical component
- ▶ Oil effects

$$u_t = \tau_t^u + \kappa_u \delta_t + \eta_t^u$$

$$P_t^{oil} = \tau_t^{oil} + \zeta_{oil} \vartheta_t + \eta_t^{oil}$$

Example 3: Modelling inflation (UCM)

- ▶ Trend component

$$\pi_t = \tau_t^\pi + \lambda_\pi \delta_t + \zeta_\pi \vartheta_t + \eta_t^\pi$$
The equation $\pi_t = \tau_t^\pi + \lambda_\pi \delta_t + \zeta_\pi \vartheta_t + \eta_t^\pi$ is shown. The term τ_t^π is enclosed in a light blue square, $\lambda_\pi \delta_t$ is enclosed in a light red oval, and $\zeta_\pi \vartheta_t$ is enclosed in a light green square. An arrow from the text 'Trend component' points to the blue square, and another arrow from the text 'Cyclical component' points to the red oval.

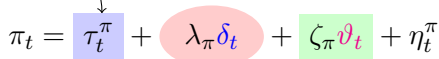
- ▶ Cyclical component
- ▶ Oil effects

$$u_t = \tau_t^u + \kappa_u \delta_t + \eta_t^u$$

$$P_t^{oil} = \tau_t^{oil} + \zeta_{oil} \vartheta_t + \eta_t^{oil}$$

Example 3: Modelling inflation (UCM)

- Trend component

$$\pi_t = \tau_t^\pi + \lambda_\pi \delta_t + \zeta_\pi \vartheta_t + \eta_t^\pi$$
The equation $\pi_t = \tau_t^\pi + \lambda_\pi \delta_t + \zeta_\pi \vartheta_t + \eta_t^\pi$ is shown. The term τ_t^π is enclosed in a light blue square, $\lambda_\pi \delta_t$ is in a light red oval, and $\zeta_\pi \vartheta_t$ is in a light green square. Arrows from the text labels 'Trend component', 'Cyclical component', and 'Oil effects' point to these respective terms.

- Cyclical component

- Oil effects

$$u_t = \tau_t^u + \kappa_u \delta_t + \eta_t^u$$

$$P_t^{oil} = \tau_t^{oil} + \zeta_{oil} \vartheta_t + \eta_t^{oil}$$

Example 3: Modelling inflation (UCM). Cont.

$$\pi_t = \tau_t^\pi + \lambda_\pi \delta_t + \zeta_\pi \vartheta_t + \eta_t^\pi \quad (4)$$

$$\pi_t^{core} = \tau_t^\pi + \lambda_{core} \delta_{t-4} + \eta_t^{core} \quad (5)$$

$$\pi_t^m = \tau_t^m + \zeta_m \vartheta_t + \eta_t^M \quad (6)$$

$$P_t^{oil} = \tau_t^{oil} + \zeta_{oil} \vartheta_t + \eta_t^{oil} \quad (7)$$

$$u_t = \tau_t^u + \kappa_u \delta_t + \eta_t^u \quad (8)$$

$$y_t = \tau_t^y + \kappa_y \delta_t + \eta_t^y \quad (9)$$

$$S_t^Y = m_S + \kappa_S (\delta_t - \delta_{t-4}) + \eta_t^S \quad (10)$$

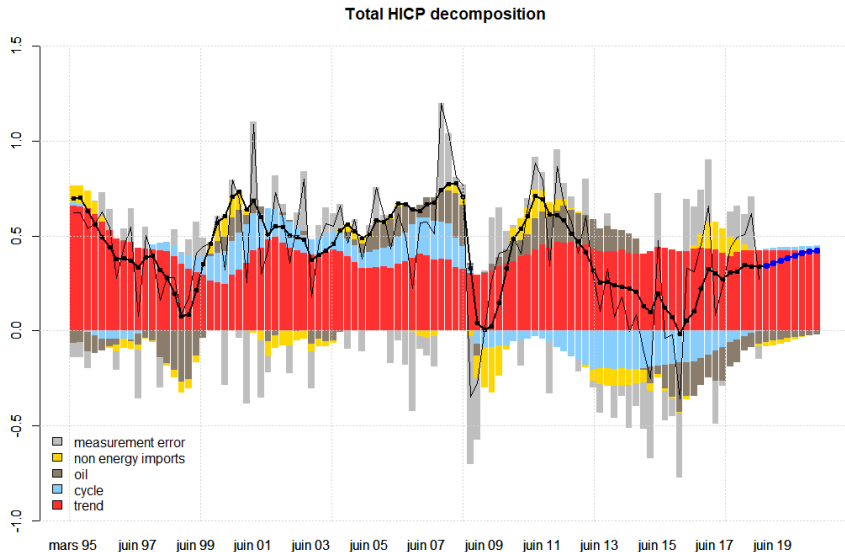
$$\mu_t = m_\mu + \kappa_\mu (\delta_t - \delta_{t-4}) + \eta_t^\mu \quad (11)$$

Surveys or variables reflecting expectations: [Basselier et. al \(2018\)](#), NBB

$$\underbrace{\pi_t^E}_{E[\pi_{t+h}|t]} = \tau_{t+h|t}^\pi + \lambda_\pi \delta_{t+h|t} + \zeta_\pi \vartheta_{t+h|t} + \eta_{t+h|t}^\pi \quad (12)$$

- ▶ Qualitative survey data (e.g. Business and Consumer expectations)
- ▶ Quantitative data (e.g. ECB Survey of Professional Forecasters)

Example 2: application for euro area data: Total HICP



- ▶ Additional components (on demand)
 - ▶ VAR
 - ▶ ...
- ▶ Constraints
 - ▶ On parameters
 - ▶ On state items
- ▶ Regression components
- ▶ Automatic scaling of the variables
- ▶ Automatic verification of identifiability

- ▶ Java code (JD+ 3.0):

<https://github.com/jdemetra/jdemetra-core>

- ▶ R package, examples and documentation (coming soon):

<https://github.com/nbbird>