
Paula Marta Bruno, Fernando Duarte Pereira
Technical University of Lisbon

Renato Fernandes
Instituto Politecnico de Santarem

Gonçalo Vilhena de Mendonça
Technical University of Lisbon

The responses to supramaximal exercise testing have been traditionally analyzed by means of standard parametric and nonparametric statistics. Unfortunately, these statistical approaches do not allow insight into the pattern of variation of a given parameter over time. The purpose of this study was to determine if the application of dynamic factor analysis (DFA) allowed discriminating different patterns of power output (PO), during supramaximal exercise, in two groups of children engaged in competitive sports: swimmers and soccer players. Data derived from Wingate testing were used in this study. Analyses were performed on epochs (30 s) of upper and lower body PO obtained from twenty two healthy boys (11 swimmers and 11 soccer players) age 11–12 years old. DFA revealed two distinct patterns of PO during Wingate. Swimmers tended to attain their peak PO (upper and lower body) earlier than soccer players. As importantly, DFA showed that children with a given pattern of upper body PO tend to perform similarly during lower body exercise.

Time series analysis is a powerful approach to understanding several biological systems. In the context of exercise testing, it defines an ordered sequence of values of a given variable (i.e., power output—PO) being measured repeatedly over time. The calculations, however, usually require very long data sets that can be difficult or impossible to obtain. Pediatric exercise testing provides useful information which
can be used for determining the need of a treatment or the presence of a training effect (15). Since children have a lower capacity of anaerobic glycolysis than adults, their response to supramaximal exercise performance has been viewed as an important marker of child development (5,10). However, in these settings, data are typically characterized by short, irregular and noisy time series. For this reason, physiological responses to exercise have been traditionally analyzed by means of standard parametric and nonparametric statistics. Unfortunately, these statistical approaches do not allow further insight into the pattern of variation of a given parameter over time. This is particularly relevant for specific pediatric disorders, in which peak muscle power and local muscle endurance are reduced (i.e., severe muscle atrophy and muscle dystrophy).

Since epochs of consecutive data are intrinsically dependent and correlated, the order of observations is particularly important when time is involved. There are several statistical approaches that do not consider the a priori assumption of independence, being therefore, more suited for this type of analyses. Nonconventional statistical techniques (i.e., nonlinear analyses) have proved useful for determination of atypical patterns of motor behavior during treadmill and handgrip exercise (2,8). Zuur et al. (16) presented a detailed mathematical derivation of the Dynamic Factor Analysis (DFA). DFA corresponds to a dimension reduction technique that allows modeling of a multivariate time series as a function of a few trends (common patterns). Despite the apparent potential of DFA to discriminate between specific patterns of motor behavior (i.e., muscle PO during exercise) in different populations, to our knowledge, this has never been explored. Therefore, the purpose of this study was to determine if the application of DFA contributes for distinguishing different patterns of PO, in response to supramaximal exercise, in two groups of children engaged in competitive sports: 1) swimmers—Sw, 2) soccer players—Sc. We hypothesized that this nonconventional statistical approach would add relevant information to that commonly derived from standard statistics.

Material and Methods

Participants

A total of 22 healthy boys (11 Sw, 11 Sc), aged 11–12 years, were included in the current study. All these participants had been engaged in competitive sports (swimming or soccer) for at least 2 years. Descriptive statistics are presented in Table 1.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Physical Characteristics of the Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Soccer</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>148.6 ± 3.0</td>
</tr>
<tr>
<td>Body mass (kg)</td>
<td>38.0 ± 4.7 *</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>17.2 ± 1.8 *</td>
</tr>
</tbody>
</table>

Values are mean ± SD. * p < 0.05. Welch’s t test was performed for Height, Mann-Whitney U test was performed for body mass, and independent t test was performed for BMI.
Study Design and Data to be Analyzed

Participants were evaluated, while performing upper and lower body supramaximal exercise (Wingate tests—WnT) on a randomized order, over the course of two visits on separate days and within a 7-day period. Testing was separated by at least 48 hr and, to minimize the effects of circadian and other similarly induced variations in performance, was performed at approximately the same time of day (between 3 and 5 p.m.). All subjects abstained from vigorous exercise for 24 hr before testing and were at least 3-hr postprandial upon arrival for testing. During the first visit, standing height and weight measurements were taken with participants wearing light-weight clothing and no shoes. Height was obtained using a stadiometer with measures obtained to the nearest 0.5 cm. Weight was measured on a balance-beam scale. Body mass index (BMI) was calculated by dividing the participants’ mass in kilograms by the square of their height in meters.

Lower Body Supramaximal Exercise Testing

Participants underwent a 5-min lower body warm-up on a mechanically braked cycle ergometer (Monark Ergomedic 894E) at a pedaling rate of 50 rpm at a constant PO of 50 W. The warm-up included 2–3 brief (3–5 s) sprinting bouts to maximal cycling speed (120 rpm) at higher power output. The intensity of warm-up was chosen to increase heart rate to approximately 140–150 bpm (6). Following the warm-up, participants rested on the cycle ergometer for 1 min, and they were instructed to pedal at full speed with the cycle ergometer unloaded for 5–8 s. At this stage, the full braking force [7.5% of subject’s body mass (kg)] was applied and a 30-s count was started (11).

Upper Body Supramaximal Exercise Testing

Upper body exercise testing was performed with the aid of handlebars which were attached to the same ergometer used during lower body supramaximal exercise testing. The ergometer was tightly secured to an elevated platform by means of screw pins and testing was conducted with the participants in the seated position. The height of the chair used in this study was varied to allow proper alignment between the shoulders of each participant and the ergometer axis of rotation. Before testing, care was undertaken to ensure that all participants attained full contact between their lower extremities and a stable surface. Even though warm-up procedures were similar to of lower body exercise, no external load was used during upper body cranking and sprinting bouts were performed to a maximal cycling speed of 50–60 rpm (3). Afterward, participants were also instructed to pedal at full speed with the cycle ergometer unloaded for 5–8 s. At this stage, the full braking force [5% of subject’s body mass (kg)] was applied and a 30-s count was started (11).

Testing was carried out in the laboratory with an environmental temperature between 21–24 °C and a relative humidity between 44–56%. In an attempt to control for possible circadian variations in PO performance, the measurements were performed between 3 and 5 p.m. at approximately the same time of day for all individuals. The raw PO signal during WnT was recorded online, at a frequency of 1 Hz, and measured variables were mean power and peak power. External resistance
was controlled, the power output was measured and mean power and peak power were calculated (offline) from the exercise results using the manufacturer software (Monark Anaerobic Wingate Software, version 1.0, Monark, Vansbro, Sweden). Mean power was calculated as the average PO over the 30-s sprint and peak power was defined as the highest power achieved at any 5-s stage of the test. The participation was voluntary and informed consent was obtained from the parents of all subjects before the study. The study protocol was approved by the university’s internal review board.

**Standard Statistics**

Descriptive statistics are reported as mean ± SD. Before comparing both groups, data were tested for normality and homogeneity of variance with the Shapiro-Wilk and Levene’s tests, respectively. For variables with Gaussian distribution, between-group (Sw and Sc) comparisons were computed using independent t test or Welch’s t test. In the cases of skewed distribution, Mann-Whitney U tests were performed. Statistical significance was set at p < .05. All data analysis was carried out using Statistical Package for the Social Sciences (SPSS Statistics 17.0 for Windows, SPSS Inc, Chicago, USA).

**Dynamic Factor Analysis**

DFA is a statistical approach that intends to model N observed time series in terms of M common trends (where M should be lower than N). Importantly, it allows modeling of short and nonstationary time series such as those derived from supramaximal exercise testing (30-s WnT). In structural models of time series, observations are described in terms of trends, seasonal effects, cycles, explanatory variables and noise (7). For PO epochs, such as those derived from WnT, the only critical components are trends and noise. Thus, in this study, a simpler version of the model is applicable (data = trends + constant + noise). Accordingly, the mathematical formulation for the dynamic factor model (DFM) with M common trends corresponds to:

\[
y_t = \Gamma \alpha_t + \mu + \varepsilon_t \quad (1),
\]

\[
\alpha_t = \alpha_{t-1} + \eta_t \quad (2),
\]

where \( y_t \) is the response vector at time t, \( \Gamma \) is the factor loadings matrix, \( \alpha_t \) is the common trends vector at time t, \( \mu \) is the level vector, \( \varepsilon_t \) and \( \eta_t \) are noise vectors at time t. In the context of WnT, the vector \( y_t \) contains the PO values at each second \( t = 0, \ldots, 30 \).

The primary purpose of DFA is to estimate underlying common trends. If there would be only one common trend (i.e., single pattern of PO production over individuals), M would be equal to one. On the contrary, if different patterns of PO would emerge, more common trends would be required and, therefore, M would be greater than one (2, 3, 4, or even more—depending on the diversity of trends). Each of the N PO epochs is related with each of the M common trends by means
of a loading. Factor loadings are the coefficients of linear combinations associating time series with the common trends. The magnitude and sign of the factor loadings determine how the trends are related to the series.

The error component $\varepsilon_t$ is assumed to be white noise (i.e., normally distributed with zero mean and a covariance matrix $H$). The easiest approach to model the covariance error matrix is to use a diagonal matrix. However a positive-definite, symmetric, nondiagonal matrix could be preferable in some situations, since off-diagonal elements of $H$ represent the joint information in the response variables, which cannot be explained by the common trends. The disadvantage of using such a covariance matrix is that the number of parameters increases drastically.

All unknown parameters in the model are estimated with the method of maximum likelihood estimation. For further details on DFA theory see Appendix.

Model Fit

In shortly, DFA procedure can be described as follows. Every analysis should always start with a simple plot of the series versus time. As interpretation of common trends and factor loadings is generally easier if the response variables have approximately the same scale, it is convenient to standardize each time series by subtracting its mean and dividing by its standard deviation (as a result, trends, loadings, and fitted values are unitless).

In structural models with trends and noise components, two different types of models should be estimated:

\[
\text{data} = M \text{ common trends} + \text{constant} + \text{noise}, \quad \text{where } H \text{ is diagonal},
\]

\[
\text{data} = M \text{ common trends} + \text{constant} + \text{noise}, \quad \text{where } H \text{ is symmetric},
\]

for each type of models, different values of $M$ can be used ($M = 1, 2, 3, \ldots$). Since several competing models are performed, a model selection tool is necessary. Akaike’s information criterion (AIC; 1), defined as twice the difference between the log likelihood function (measure of fit) and the number of parameters (penalty), is usually used. The DFM with the lowest AIC value is taken to be the ‘best’ candidate model, however special care should be taken as AIC tends to select overparameterized models. Small differences between AIC values are regarded as negligible; and a simpler model may be preferable (17, page 307).

Once selection has been made, a model validation is required. Useful tools are graphs in which fitted lines and the observed data are plotted (they identify the extent to which the model can capture the patterns in the time series). Residual analysis should also be considered.

Finally, common trends and factor loadings obtained with the chosen model must be interpreted. The loadings indicate which common trend is related to which time series, so a cutoff level should be taken (for example, 0.2) to define whether a series is associated with a trend.

Estimated models in this study were obtained with the statistics software package Brodgar (version 2.6.5, Highland Statistics Ltd, Newburgh, UK). For further details see Zuur et al. (16).
Results

Standard Statistics

Descriptive data are presented in Table 1. No differences were presented between groups for height; however Sw were heavier and had a larger body mass index (BMI). In comparison with Sc, Sw attained higher values of relative peak power during supramaximal upper body exercise ($p < .05$; Table 2). Conversely, no between group differences were found for any other PO variable included in Table 2.

Table 2  Power Output (W.kg$^{-1}$) Parameters by Body Part and by Sport Modality

<table>
<thead>
<tr>
<th></th>
<th>Peak power</th>
<th>Mean power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Soccer</td>
<td>Swimming</td>
</tr>
<tr>
<td>Upper</td>
<td>4.3 ± 0.6 *</td>
<td>5.4 ± 1.1</td>
</tr>
<tr>
<td>Lower</td>
<td>7.5 ± 0.7</td>
<td>7.9 ± 1.1</td>
</tr>
</tbody>
</table>

Values are mean ± SD. * $p < 0.05$. Independent $t$ test was performed for both upper and lower body peak power, and for upper body mean power, Welch’s $t$ test was performed for lower body mean power.

Dynamic Factor Analysis

Figure 1 shows the time series plots of PO through time for both upper and lower body exercise. Each line represents a PO epoch from each child. Both plots indicate some variability between subjects and, apparently, different patterns of PO during supramaximal exercise. As depicted in Figure 1A, there are two PO epochs (marked with line and dot) with very low and almost constant values over time (Sw2, Sw10). Their corresponding sequences revealed a PO of approximately 3.0 W.kg$^{-1}$ along 30 s, which in the context of WnT is incompatible with acceptable results. For this reason, both Sw2 and Sw10 were excluded from further analyses. In addition, as both plots confirmed the existence of missing values (at the end of the sequences); DFA was only computed on the initial 27 s of WnT.

Figure 1 — Power output (W.kg$^{-1}$): (A) upper body series; (B) lower body series of all participants.
DFA was conducted on the standardized relative PO (W.kg\(^{-1}\)) series and 16 models were estimated; 8 for each upper and lower body tests. The number of common trends (M) varied between one and four and both a diagonal and a symmetric error covariance matrix (\(H\)) were explored for each M.

Table 3 shows the AIC and number of parameters for each estimated model and indicates similarities between upper and lower body tests. Accordingly, a model with a symmetric \(H\) was, in this case, preferable. While models with a single common trend (AIC: -426.2 and -256.8) provided less information than others, those with four common trends (AIC: -476.8 and -315.5) revealed that more parameters was redundant. Thus, it was required to select between models with two or three trends.

### Table 3: AIC Values Obtained by Applying Dynamic Factor Model on the Standardized Relative Power Output Series by Body Part

<table>
<thead>
<tr>
<th>M</th>
<th>Upper body</th>
<th>Lower body</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H diagonal ((\hat{\beta}))</td>
<td>H symmetric ((\hat{\beta}))</td>
</tr>
<tr>
<td>1</td>
<td>928.7 (60)</td>
<td>-426.2 (250)</td>
</tr>
<tr>
<td>2</td>
<td>-114.8 (79)</td>
<td>-459.1 (269)</td>
</tr>
<tr>
<td>3</td>
<td>-252.8 (97)</td>
<td>-476.0 (287)</td>
</tr>
<tr>
<td>4</td>
<td>-358.3 (114)</td>
<td>-476.8 (304)</td>
</tr>
</tbody>
</table>

The values shown are AIC-Akaike Information Criteria (number of estimated parameters). M is the number of common trends. H is the error covariance matrix.

Table 4 presents the factor loadings for each estimated model. Loadings indicate the degree of association between each trend and each child. As can be seen, in models with three common trends, the third trend displayed very low values for factor loadings (in general, loadings were below 0.1 and third trend was never the most important one for each participant). This indicated that the third trend did not add relevant information to the model. The ratio between the sum squared residuals and the sum squared measured was also calculated for each child. These ratios (not shown) were very similar when comparing both models (M = 2 and M = 3). While models with three common trends had to estimate 287 parameters, models with two common trends only required 269 (Table 3). As simpler models are preferable, we selected those with two common trends and a symmetric \(H\) (AIC: -459.1 and -290.2) for subsequent analysis (upper and lower body tests).

The two estimated common trends for upper body tests are depicted in Figure 2. There was one common trend showing a gradual increase in PO followed by attainment of peak power at approximately 10 s. PO was then sustained for a period of 5 s, after which it decreased continuously until test termination (late peak power and drop—LPPD). The other common trend initiated with the production of peak power, followed by a sharp decrease in PO lasting 15 s and a smoother power drop from there to the end of exercise (early peak power and drop—EPPD).

Figure 3A shows the factor loadings of LPPD versus EPPD. An arbitrary cutoff level for loadings of 0.2 was chosen to define whether a series was associated with
Table 4  Estimated Factor Loadings Obtained by Dynamic Factor Model with a Symmetric Error Covariance Matrix

<table>
<thead>
<tr>
<th></th>
<th>Upper body</th>
<th></th>
<th>Lower body</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M = 2</td>
<td>M = 3</td>
<td>M = 2</td>
<td>M = 3</td>
</tr>
<tr>
<td>F1</td>
<td>0.319</td>
<td>0.288</td>
<td>0.107</td>
<td>-0.034</td>
</tr>
<tr>
<td>F2</td>
<td>0.268</td>
<td>0.282</td>
<td>0.158</td>
<td>0.079</td>
</tr>
<tr>
<td>F3</td>
<td>0.123</td>
<td>0.310</td>
<td>0.003</td>
<td>0.105</td>
</tr>
<tr>
<td>F4</td>
<td>0.240</td>
<td>0.245</td>
<td>0.171</td>
<td>0.068</td>
</tr>
<tr>
<td>F5</td>
<td>0.169</td>
<td>0.289</td>
<td>0.310</td>
<td>-0.014</td>
</tr>
<tr>
<td>F6</td>
<td>0.285</td>
<td>0.294</td>
<td>0.032</td>
<td>0.080</td>
</tr>
<tr>
<td>F7</td>
<td>0.111</td>
<td>0.315</td>
<td>0.322</td>
<td>0.042</td>
</tr>
<tr>
<td>F8</td>
<td>0.274</td>
<td>0.261</td>
<td>0.211</td>
<td>0.005</td>
</tr>
<tr>
<td>F9</td>
<td>0.315</td>
<td>0.285</td>
<td>0.124</td>
<td>-0.033</td>
</tr>
<tr>
<td>F10</td>
<td>0.334</td>
<td>0.298</td>
<td>0.086</td>
<td>-0.062</td>
</tr>
<tr>
<td>F11</td>
<td>0.303</td>
<td>0.283</td>
<td>0.157</td>
<td>-0.010</td>
</tr>
<tr>
<td>S1</td>
<td>0.128</td>
<td>0.301</td>
<td>0.322</td>
<td>-0.021</td>
</tr>
<tr>
<td>S2</td>
<td>0.312</td>
<td>0.304</td>
<td>-0.037</td>
<td>0.023</td>
</tr>
<tr>
<td>S3</td>
<td>0.309</td>
<td>0.275</td>
<td>-0.079</td>
<td>-0.045</td>
</tr>
<tr>
<td>S4</td>
<td>-0.022</td>
<td>0.324</td>
<td>-0.003</td>
<td>0.325</td>
</tr>
<tr>
<td>S5</td>
<td>0.048</td>
<td>0.320</td>
<td>0.052</td>
<td>0.333</td>
</tr>
<tr>
<td>S6</td>
<td>0.041</td>
<td>0.323</td>
<td>0.061</td>
<td>0.319</td>
</tr>
<tr>
<td>S7</td>
<td>0.012</td>
<td>0.239</td>
<td>-0.022</td>
<td>0.286</td>
</tr>
<tr>
<td>S8</td>
<td>0.030</td>
<td>0.320</td>
<td>0.032</td>
<td>0.336</td>
</tr>
<tr>
<td>S9</td>
<td>0.078</td>
<td>0.290</td>
<td>0.060</td>
<td>0.322</td>
</tr>
</tbody>
</table>

M is the number of common trends.
Factor loadings in bold are above the cutoff of 0.2 in absolute value. Factor loadings are unitless.
Figure 2 — Estimated common trends for power output series: (A) late peak power and drop upper body; (B) early peak power and drop upper body; (C) late peak power and drop lower body; (D) early peak power and drop lower body, obtained by dynamic factor model with two common trends and a symmetric error covariance matrix. Common trends are unitless.

Figure 3 — Factor loadings corresponding to the two common trends related to power output: (A) upper body series; (B) lower body series, obtained by dynamic factor model with two common trends and a symmetric error covariance matrix. Factor loadings are unitless.

a trend (Table 4). Accordingly, Sc and Sw had greater factor loadings for LPPD and EPPD, respectively. Therefore, generally, while Sc were associated with LPPD, Sw presented a dominance of EPPD trends. It could also be noticed that, in opposition to the Sc EPPD factor loadings, those of Sw3 and Sw4 were small but negative.
On the other hand, despite showing small LPPD factor loadings, Sc3, Sc5 and Sc7 presented higher values than those obtained for Sw. Figure 4 depicts the upper body PO data of the participants and their respective DFM estimated values with two common trends and a symmetric $H$. The plots indicate that, in general, the series were quite well fitted (only some times of the series Sw8 has not good adjustment). Low ratios between the sum of squares residuals and the sum of squares (not shown) measured indicate the adequacy of the model. In support of this, the residuals of all upper body PO epochs obtained by DFA with two common trends and a symmetric $H$, were randomly distributed around zero.

Figure 2 shows the estimated common trends for lower body tests. The underlying structure of each common trend was highly similar in shape to that obtained for upper body exercise (LPPD and EPPD). Factor loadings demonstrate that LPPD was associated with Sc, while Sw demonstrated an EPPD dominant trend (Table 4 and Figure 3B). As for upper body tests, the factor loadings indicate that while the pattern of Sw3 and Sw4 was driven by the LPPD common trend that of Sc3, Sc5 and Sc7 corresponded to EPPD. The lower body PO data of the participants and their respective DFM estimated values with two common trends and a symmetric $H$ are similar to that described for upper body tests, the plots and the ratios (not shown) indicate that the adjustment suggests the adequacy of the model.

![Figure 4](image-url) — Observed power output upper body series (dot) and fitted values (line) obtained by dynamic factor model with two common trends and a symmetric error covariance matrix. Observed and fitted values are unitless.
Discussion

The main finding of the current study is that DFA provides additional information to that of standard statistics for the analysis of PO time series derived from children performing supramaximal exercise (WnT—upper and lower body). The use of DFA allowed discriminating between two different common patterns of PO during upper and lower body WnT. Children with a given pattern of upper body PO tended to show similar performance during lower body exercise. We also found that, even though both groups showed comparable average power production during WnT; their pattern of PO was quite different over time. Specifically, Sw tended to attain their peak power earlier than Sc during both upper and lower body WnT. Therefore, these findings suggest that DFA is sensitive to the effects of sport training specificity on the performance of children during supramaximal exercise, and this is novel. The use of standard statistics for exploring differences in time series (between or within individuals) provides no information about the pattern underlying a sequence of values being measured repeatedly over time. This study intended to determine the suitability of a specific statistical approach for the analysis of time series derived from pediatric supramaximal exercise testing. Despite it has been previously used for other purposes (9,13,16), to the best of our knowledge, DFA has not been used in the context of exercise testing.

Physiological Implications

Although the current study did not include any physiological measurement, it is interesting to note that two distinct patterns of power output resulted from supramaximal exercise testing of children engaged in different sport modalities. To the best of our knowledge, the relative contribution of neuromuscular activation for power production during WnT in children has not been previously explored. According to our findings, DFA allows to discriminate different patterns of power output between children adapted to different sports. It could, therefore, be hypothesized that this approach provides some insight into the study specialization in young athletes. Since we also found similarities between the patterns of upper and lower body power output in each group, it could also be speculated that some degree of neuromuscular specialization may be involved (i.e., differences in motor unit recruitment).

Limitations

There are three main limitations to this study. First, two children from the Sw group had to be excluded due to poor performance during WnT. This reduced our sample size, thus limiting our conclusions by means of DFA. Second, as there were incomplete epochs of power output during the 30 s of WnT, we only included 27 data points in our analyses. Nevertheless, since the main differences between groups corresponded to the time of peak power occurrence during exercise, we do not believe that our findings are substantially influenced by having reduced the time series duration from 30 to 27 s. Third, WnT exercise performance is effort dependent; thus it is possible that children from different groups may have produced dissimilar
efforts. However, the uniformity of patterns of power output generated within each group further corroborates the assumption of valid supramaximal effort in both.

**Appendix**

Dynamic factor model (DFM) with M common trends corresponds to:

\[
\begin{align*}
\mathbf{y}_t &= \mathbf{\Gamma} \alpha_t + \mu + \mathbf{\varepsilon}_t \quad (1), \\
\alpha_t &= \alpha_{t-1} + \eta_t \quad (2),
\end{align*}
\]

where \( \mathbf{y}_t \) is of dimension N and represents the response vector at time t; \( \mathbf{\Gamma} \) is of dimension NxM and represents the factor loadings matrix; \( \alpha_t \) is of dimension M and represents the common trends vector at time t; \( \mu \) is of dimension N and represents the level vector; \( \mathbf{\varepsilon}_t \) is a noise vector at time t of dimension N; and \( \eta_t \) is also a noise vector at time t, but of dimension M.

In the model, common trends at time t are represented by the vector \( \alpha_t \) of dimension M (\( M = \) number of common trends). Each of the N time series (represented by \( \mathbf{y}_t \)) is related with each of the M common trends (represented by \( \alpha_t \)) by means of a loading. Loadings (coefficients of linear combinations associating time series with the common trends) are represented in \( \mathbf{\Gamma} \) matrix. The number of rows of \( \mathbf{\Gamma} \) is N, and the number of columns is M. Both an intercept and noise are also required for the model. They are represented by vectors \( \mu \) and \( \mathbf{\varepsilon}_t \), respectively, both with dimension N (the same as the \( \mathbf{y}_t \)—each epoch has its own intercept and its own error component). Therefore, the equation (1) could be expressed as

\[
\begin{bmatrix}
\mathbf{y}_{1,t} \\
\mathbf{y}_{2,t} \\
\vdots \\
\mathbf{y}_{N,t}
\end{bmatrix}
= \begin{bmatrix} 1 \\ 2 \\ \vdots \\ N \end{bmatrix} \begin{bmatrix} 1,1 \\ 2,1 \\ \vdots \\ N,1 \end{bmatrix} \begin{bmatrix} \mathbf{\gamma}_{1,1} \\ \mathbf{\gamma}_{2,1} \\ \vdots \\ \mathbf{\gamma}_{N,1} \end{bmatrix}
+ \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{bmatrix}
+ \begin{bmatrix} \mathbf{\varepsilon}_{1,t} \\ \mathbf{\varepsilon}_{2,t} \\ \vdots \\ \mathbf{\varepsilon}_{N,t} \end{bmatrix}
\quad (3),
\]

if there would be only one common trend; or as

\[
\begin{bmatrix}
\mathbf{y}_{1,t} \\
\mathbf{y}_{2,t} \\
\vdots \\
\mathbf{y}_{N,t}
\end{bmatrix}
= \begin{bmatrix} 1,1 \\ 2,1 \\ \vdots \\ N,1 \\
1,2 \\ 2,2 \\ \vdots \\ N,2 \\
1,t \\ 2,t \\ \vdots \\ N,t
\end{bmatrix} \begin{bmatrix} \mathbf{\gamma}_{1,1} \\ \mathbf{\gamma}_{2,1} \\ \vdots \\ \mathbf{\gamma}_{N,1} \\ \mathbf{\gamma}_{1,2} \\ \mathbf{\gamma}_{2,2} \\ \vdots \\ \mathbf{\gamma}_{N,2} \end{bmatrix}
+ \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\
1 \\ 1 \\ \vdots \\ 1 \\
1,t \\ 2,t \\ \vdots \\ N,t
\end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{bmatrix}
+ \begin{bmatrix} \mathbf{\varepsilon}_{1,t} \\ \mathbf{\varepsilon}_{2,t} \\ \vdots \\ \mathbf{\varepsilon}_{N,t} \end{bmatrix}
\quad (4),
\]

if there would be two common trends (loadings in the first column of \( \mathbf{\Gamma} \) associate with \( \alpha_{1,t} \) common trend; and loadings on second column associate with \( \alpha_{2,t} \) common trend). For more than two common trends, equation (4) should appear similarly but with matrices \( \mathbf{\Gamma} \) and \( \alpha_t \) of greater dimension. The \( \gamma_{i,j} \) loading is used to determine the degree of relationship between the \( \gamma_{i,t} \) epoch and \( \alpha_{j,t} \) trend (\( i = 1, \ldots , N, \ j = 1, \ldots , M \)). Large loading values (in absolute sense) imply that \( y_{i,t} \) follows the same
or the opposite pattern of $\alpha_{j,t}$, depending on $\gamma_{i}$, positive or negative, respectively. A loading value approaching zero suggests that $y_{i,t}$ does not follow the pattern of $\alpha_{j,t}$.

The progression of the trend component ($\alpha_{t}$) is driven by equation (2), and is based on a random walk process (i.e., the trend at each given time equals the previous trend plus some random error). The error component is given by $\eta_{t}$, which corresponds to a vector of dimension $M$ (same dimension of $\alpha_{t}$). In general, it is assumed that both error vectors are white noise, i.e., $\varepsilon_{t} \sim N(0,H)$, $\eta_{t} \sim N(0,Q)$, and that $\varepsilon_{t}$ and $\eta_{t}$ are independent of each other. It is also assumed that the starting vector $\alpha_{0} \sim N(a_{0},V_{0})$, in which $\alpha_{0}$ is independent of both $\varepsilon_{t}$, and $\eta_{t}$. The covariance matrix $H$ of the error term $\varepsilon_{t}$ is of dimension $N$, and several structures can be chosen for this matrix (12). The simplest approach is to use a diagonal matrix. However a symmetric nondiagonal matrix (where each pair of time series has its own component of covariance error) should be preferable in some situations. For the covariance matrix $Q$ of the error vector $\eta_{t}$, Harvey (7) suggest the utilization of the identity matrix (i.e., $Q = I$).

The parameters of matrices $\Gamma$, $H$, $\mu$, $a_{0}$, and $V_{0}$ are unknown and should be determined as accurately as possible. For this purpose, the method of maximum likelihood estimation, using the expectation maximization algorithm, is recommended. This algorithm was first introduced by Dempster et al. (4), and then adapted by Shumway and Stoffer (14), and it corresponds to a forward backward Kalman filter and smoother.

References