

# Comparison of EMD and HP Filter for Cycle Extraction with Korean Macroeconomic Indices

Minjeong Park<sup>a</sup> · Byeongchan Seong<sup>b,1</sup>

<sup>a</sup>Statistical Research Institute, Statistics Korea

<sup>b</sup>Department of Applied Statistics, Chung-Ang University

(Received February 13, 2014; Revised March 28, 2014; Accepted April 5, 2014)

---

## Abstract

We introduce the empirical model decomposition (EMD) to decompose a time series into a set of components in the time-frequency domain. By using EMD, we also extract cycle and trend components from major Korean macroeconomic indices and forecast the indices with the components combined. In order to evaluate their efficiencies, we investigate volatility, autocorrelation, persistence, Granger causality, nonstationarity, and forecasting performance. They are then compared with those by Hodrick-Prescott filter which is the most commonly used method.

Keywords: Empirical model decomposition, Hodrick-Prescott filter, time-frequency analysis.

---

## 1. 서론

시계열 분석에서는 시계열의 모형화 및 예측을 위해서 전통적으로 분해법(decomposition)이 널리 사용되어 왔다. 경제 시계열은 추세, 순환변동, 계절 성분 및 불규칙 성분의 4가지 구성 요소로 이루어진다고 알려져 있는데 (Cho와 Son, 2009), 계절 성분은 1년 이내의 주기적인 변동으로 정의되며, 주로 계절 조정의 과정을 거쳐서 제거된다. 따라서 추세 성분 및 순환변동을 추출하는 방법에 대해서 더 많은 연구의 초점이 맞춰져 왔고, 관련된 대표적인 방법을 살펴보면, Beverage-Nelson 분해법, 비관측인자모형(unobserved component model), Hodrick-Prescott(HP) 필터, Baxter-King(BK) band-pass 필터 등이 있다 (Lee, 2009).

추세 성분과 순환변동을 추출하는 이러한 방법들은, 시간 영역(time domain)과 진동수 영역(frequency domain)의 분석법으로 구분될 수 있다. 대부분의 방법들이 시간 영역에 해당되며, BK 필터 방법은 진동수 영역에 속한다. 진동수 영역에서의 분석법은 각 성분을 푸리에 변환을 이용해 시간에 따라 변하지 않는 진폭과 진동수를 가지는 삼각 함수들의 합으로 표현하고 분해한다. 이러한 2가지 성분을 분리하는 기준이나 방법은 각 국가별로 상이하며 지속적인 연구의 필요성이 강조되고 있다 (Hodrick와 Prescott, 1997; Marcet와 Ravn, 2003). 우리나라의 경우에는 국내총생산(Gross Domestic Product; GDP) 자

---

This paper was submitted to Chung-Ang University as a result of Byeongchan Seong's sabbatical leave in 2013.

<sup>1</sup>Corresponding author: Department of Applied Statistics, Chung-Ang University, 221, Heukseok-dong, Dongjak-gu, Seoul 156-756, Korea. E-mail: [bcseong@cau.ac.kr](mailto:bcseong@cau.ac.kr)

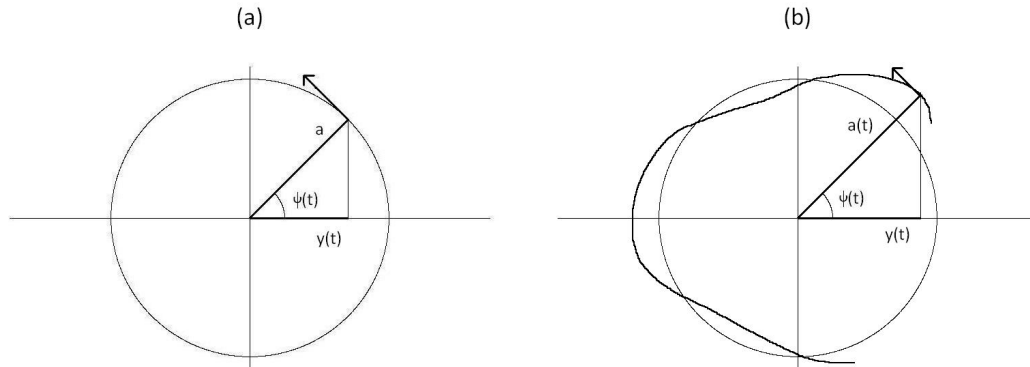


Figure 2.1. Simple harmonic motion vs. instantaneous phase and amplitude (Boashash, 1992)

료에서 추세와 순환변동을 추출하는 문제에 대하여 HP 필터보다는 band-pass 필터 사용이 실제 경기순환국면 식별에 적절함이 제시되기도 하였다 (Lee, 2009).

한편, 거시 경제 시계열에서의 추세 성분과 순환변동은 다양한 경제학적 의미를 가질 수 있다. 특히, 순환변동은 경제 안전화 정책의 합리적 수립 및 집행을 위해서 필요한 경기변동 지표에 대한 체계적인 분석을 위해서 자주 사용되고 있다. 예를들어, Ryu와 Lee (2007)은 잠재 국내총생산으로부터 도출된 GDP 갭률을 추출하고, 이를 우리나라 경기순환주기와 비교하고 있다. Nam (2007)는 실제 경기순환국면 식별을 위하여 우세한 추출법에 대하여 논의하였고 추출법에 따른 순환주기의 단기화 및 장기하향화 특징을 분석하였다. 특히, 순환변동의 스펙트럼을 활용하여 외환위기 전후의 순환변동을 비교하였다.

본 논문에서는 우리나라에 적합한 최적의 분해법을 모색하기 위한 일환으로서, 시간-진동수 영역에서 시계열을 여러 구성 성분으로 분해하는 방법인 경험적 모드 분해법(empirical mode decomposition; EMD)을 소개하고, 이를 이용하여 주요한 거시 경제 지표들에 대한 순환변동 및 추세 성분을 추출한다. 또한, 추출된 성분을 조합하여 예측에 활용한다. 특히, EMD는 시간에 따라 변하는 진동수를 가지는 하나의 삼각 함수로 각 성분을 표현하게 되어 푸리에 변환의 한계를 뛰어넘어 자료를 다룰 수 있으며, 비선형이거나 비정상적(nonstationary) 시계열을 다룰 수 있는 장점이 있다 (Huang 등, 1998).

본 논문의 구성은 다음과 같다. 2절에서 EMD 방법론을 소개하며, 3절에서 우리나라의 주요 거시 경제 지표를 대상으로 순환변동과 추세 성분을 추출하고 예측에 활용한다. 또한, 그 효율성을 살펴보기 위하여, 추출된 구성 성분들의 다양한 통계적 성질 및 예측력을 계산하고, 가장 보편적으로 널리 사용되고 있는 HP 필터에 의한 결과와 비교한다. 마지막으로 4절에서 결론을 제시한다.

## 2. EMD 방법론

전통적으로 시계열  $y(t)$ 의 진동수란, 주어진 시간동안  $y(t)$ 가 진동하는 횟수로 정의된다. 복소 평면에서  $y(t) = a(t) \cos \psi(t)$ 는 진폭  $a(t)$ 가 상수이고 움직이는 속도가 일정할 때 Figure 2.1의 (a)와 같은 원 위의 일정한 움직임으로 취급할 수 있다. 이때  $\psi'(t)$ 는 상수이며  $y(t)$ 의 진동수는 전통적인 개념으로 해석할 수 있다. 그러나  $y(t)$ 의 움직이는 속도가 일정하지 않고, 즉  $\psi'(t)$ 가 1차 이상의 함수이고, 진폭  $a(t)$ 도 시간에 따라 변하면서 움직인다면 이는 Figure 2.1의 (b)와 같이 표현되며  $y(t)$ 의 진동수는 전통적인 개념으로는 정의되지 않는다. 이러한 시간에 따라 변하는 진동수는 순시 진동수(instantaneous frequency; IF)로 정의된다 (Boashash, 1992).

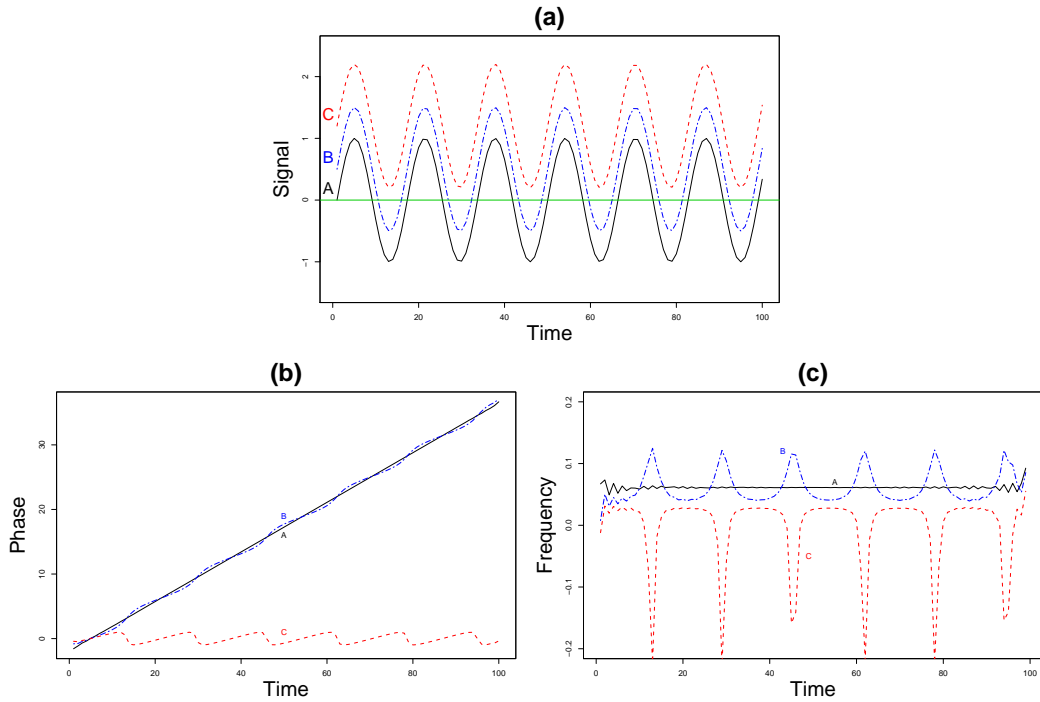


Figure 2.2. Applying Hilbert transform to different signals, their phases and IFs (Huang *et al.*, 1998)

순시 진동수, IF는 두가지 방법으로 정의될 수 있다. 첫 번째로는 시간-진동수 분포상의 스펙트럴 피크로 정의할 수 있다. 이는 국소적으로  $y(t)$ 에 가장 잘 들어맞는 일정한 속도로 움직이는 삼각 함수의 진동수로 이해되기도 한다. 두 번째로 IF는 위상(phase), 즉  $\psi(t)$ 의 미분값으로 정의되는데, 이를 위해 다음과 같은 힐버트 변환(Hilbert Transform; HT)이 이용된다,

$$H[y(t)] = p.v. \int \frac{y(t - \tau)}{\pi\tau} d\tau = (y * g)(t).$$

이때  $p.v.$ 는 코시 주요값(Cauchy principal value)이고  $g(t) = 1/t$ 이다. 힐버트 변환을 통해서  $y(t)$ 의 직각위상(quadrature)을 얻을 수 있다. 예를 들면,  $H[\cos(t)] = \sin(t)$  또는  $H[\sin(t)] = -\cos(t)$ 가 된다. 따라서  $y(t) = a(t) \cos \psi(t)$ 에 대해서 힐버트 변환을 통해 아래와 같은  $y(t)$ 의 해석 신호(analytic signal),  $z(t)$ 를 얻게 된다,

$$z(t) = y(t) + iH[y(t)] = a(t)e^{i\psi(t)}.$$

이러한  $z(t)$ 에 대해 역탄젠트 함수를 취하는 등의 위상 펼침(phase unwrapping) 과정을 통해 위상  $\psi(t)$ 를 얻고, 위상의 미분값으로 IF를 정의할 수 있다.

위의 첫 번째 정의를 이용한 IF의 표현을 위하여 시간-진동수 분포를 정교하게 얻고자 하는 많은 노력이 있었고, 구체적으로는 웨이블릿의 이용이나 위그너-바일같은 분포함수에 관한 연구들이 이루어져 왔다 (Mallat, 1998). 반면 두 번째 방법은 실제로는 힐버트 변환을 적용하기 어려운 시계열이 많기 때문에 EMD가 제시되기 전에는 거의 활용되지 않았다. 힐버트 변환은 시계열이 제로 대칭이어야만 올바

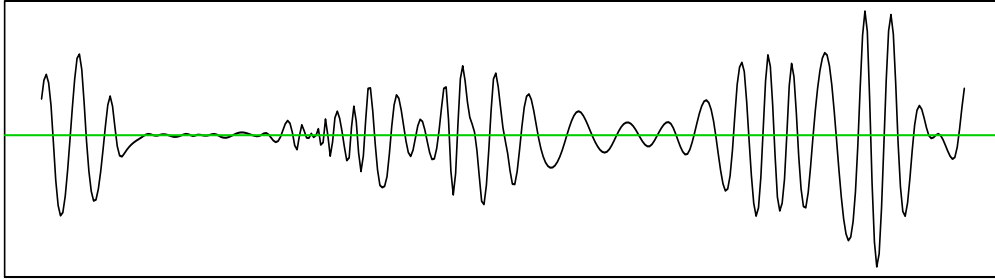


Figure 2.3. Zero-symmetric component, IMF

르게 적용되기 때문이다. 예를 들어, Figure 2.2 (a)의 시계열과 같은, 동일한 진동수를 가지는 세 개의 삼각 함수 A, B, C가 있다고 하자. 각 시계열의 위상과 진동수를 힐버트 변환을 이용하여 구하면 각각 Figure 2.2의 (b) 및 (c)와 같이 표현된다. 특히, 전혀 제로 대칭이 아닌 시계열 C의 경우, 힐버트 변환이 (c)에서 점선으로 나타나는 전혀 의미없는 음의 진동수를 제공하는 반면, 제로 대칭인 시계열 A의 경우, 실선으로 표현되는 정확한 진동수를 얻을 수 있다.

이러한 힐버트 변환을 일반적인 비선형 및 비정상 시계열에도 적용하여 IF를 적절하게 구하기 위하여 EMD가 제안되었다 (Huang 등, 1998). EMD는 시계열 자료를, 넓은 의미에서 제로 대칭으로 정의되는, 다음과 같은 내재 모드 함수(intrinsic mode decomposition; IMF)들로 분해하는 방법이다,

1. IMF의 국소 극값(국소값과 극대값)들의 개수와 영을 통과하는 점들의 개수의 차이는 한 개 이하이다.
2. 극값들을 이용해 만든 상위막(upper envelope)과 하위막(lower envelope)의 국소 평균값은 0이다.

전형적인 IMF의 예는 Figure 2.3와 같다. 단순한 삼각 함수와 달리 시간에 따라 진동수가 변하는 모습을 보이고 있으며, 진폭 역시 시간에 따라 변하고 있다. 그러나 극대점들을 연결하여 상위막을 만들고, 극소점들을 연결하여 하위막을 만들면, 두 막은 서로 대칭이고 그 합은 0이 되는 형태를 가지고 있다. 이렇게 제로 대칭인 IMF에 힐버트 변환을 적용하면 적절한 직각위상이 얻어지며, 수학적 의미를 가지는 두번째 정의에 의해 IF를 얻게 된다. 시계열을 이러한 IMF들로 분해하는 EMD의 과정은 아래와 같이 요약될 수 있다.

1. 시계열의 국소 극대값과 극소값을 찾는다.
2. 국소 극대값을 스플라인 보간법(spline interpolation)을 통해 연결하여 상위막을 만든다.
3. 마찬가지로 국소 극소값에 대해 하위막을 만든다.
4. 시그널에서 두 막의 평균값을 빼 주고 남은 성분( $h_1$ )이 IMF의 정의를 만족하는지 점검한다.
5. 성분  $h_1$ 이 IMF의 정의를 만족하지 않으면  $h_1$ 을 새로운 시계열로 취급하여 과정 1~4를 반복한다.

Figure 2.4의 (a)는 EMD의 과정 1번의 극값을 표현하고, (b)는 2~4번에서 설명하는 상위막, 하위막, 평균값을 나타내며, (c)는 4번의  $h_1$ , 그리고 (d)는 1~4번의 반복 작업 후, 최종적으로 추출된 IMF의 정의를 만족하는 성분이다. 이렇게 IMF를 하나씩 추출하여, 남은 성분이 상위막이나 하위막을 형성하지 않도록 극값의 수가 적은 추세로 판정이 날 때까지 전체 과정을 반복한다. 이상의 시계열 분해과정을 체

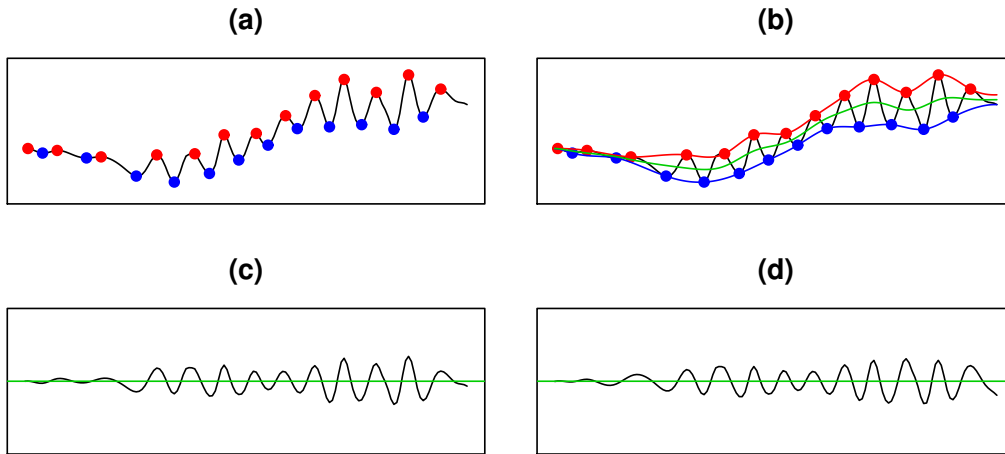


Figure 2.4. Example of the EMD sifting process

거름(sifting)이라 하며, 체거름 과정은 시계열에서 고진동수(high frequency)를 가지는 성분부터 순차적으로 추출하여 시계열을 분해한다. 각 IMF에 대하여 힐버트 변환을 적용하면 최종적으로 시계열은

$$y(t) = \sum_{j=1}^n a_j(t) \exp\left(i \int w_j(t) dt\right)$$

로 표현되며, 이때  $n$ 개 성분에 대하여  $a_j(t)$ 는 진폭을,  $w_j(t)$ 는 IF를 나타낸다. 이처럼 힐버트 변환과 EMD를 통해 시계열에 포함되어 있는 IF들을 구할 수 있게 된다.

진동수는 시계열을 나타내는 핵심적인 성질이므로, 하나의 진동수를 가지는 성분들로 시계열을 분해하는 것은 획기적인 업적이라 할 수 있다. 더불어 각 IMF들의 직교성을 포함한 여러 성질들이 논의되어 있고 (Huang 등, 1998), EMD의 원리에 대하여도 이론적인 접근이 이루어져 있다 (Rilling와 Flandrin, 2008; Park 등, 2011). 또한 잡음에 덜 민감하도록 스플라인 보간법 대신에 평활화 스플라인(smoothing spline)을 이용한 EMD가 제시되어 사용될 수 있다. 이를 통계적 경험모드분할법(statistical EMD; SEMD)이라 하며, 이에 관한 구체적인 내용은 관련 논문을 참고하기 바란다 (Kim과 Oh, 2009; Kim 등, 2012).

본 논문에서는 지금까지 기술한 EMD 과정을 이용해서 계절조정된 자료에서 순환변동 및 추세 성분을 추출한다. 세부적으로는 SEMD를 사용하여 잡음에 의한 불필요한 IMF의 생성을 제어하였고, 첫 번째 생성되는 IMF를 순환변동으로 간주하였음을 밝혀둔다.

### 3. 자료분석

본 절에서는, 추세와 순환변동을 분해하는 방법에 따라 각 성분의 특징이 어떻게 달라지는지 살펴본다. 분석을 위하여 사용된 변수는 총 8개로 국내총생산, 선행종합지수(Composite Leading Index; CLI), 전산업생산지수(Index of All Industry Production; IAIP), 광공업생산지수(Index of Industrial Production; IIP), 서비스업생산지수(Index of Services; IoS), 소매판매액지수(Retail Sales Index; RSI), 설비투자지수(Index of Equipment Investment; IEI), 및 소비자물가지수(Consumer Price Index; CPI)이다. 이 변수들은 현재 통계청에서 'KOSTAT 경제동향'을 통해 발표하고 있는 주요 실물 경제 지

표들 중 일부이다. 자료는 모두 로그변환을 취하였으며, 자료의 기간은 2000년 1분기부터 2013년 3분기까지이며 최종 1년의 자료는 예측에 활용하였다.

현재 순환변동 추출을 위해 널리 이용되고 있는 HP 필터는 평활화 과정을 통해 추세를 얻고, 자료에서 그 추세를 빼서 순환변동을 얻는다. 본 분석에서 사용한 HP의 평활화 계수는 1600으로 이는 미국 GDP에 적합하다고 제시되어 보통 사용되고 있는 값이다 (Hodrick와 Prescott, 1997; Marcet와 Ravn, 2003). 반면 EMD는 제거과정과정을 통해 순환변동들을 먼저 얻고 남은 성분이 추세가 되며, SEMD에서 사용된 평활화 계수는 10이고 바운더리는 대칭(symmetric) 옵션으로 처리하였다 (Kim 등, 2012).

HP 필터는 추세를 먼저 제거하므로 순환변동이 항상 하나만 얻어지지만, EMD는 순환변동을 먼저 추출하므로 자료에 따라 존재하는 여러 개의 순환변동이 얻어질 수 있다. 본 분석에서는 SEMD에 의해 설비투자지수를 제외하고는 모두 1개의 IMF를 얻었고 이를 순환변동으로 간주하였다. 2개의 IMF를 얻은 설비투자지수는 첫 IMF를 순환변동으로, 남은 성분들의 합을 추세로 간주하여 HP 필터와의 비교를 단순화하였다.

HP 필터와 SEMD의 각 평활화 계수는 다양할 수 있고, 그 값에 따라 얻어지는 순환변동들도 모두 다르다. 또한 시간에 따라 자료가 늘어나면서 순환변동의 모양이 달라질 수도 있으나, 본 논문에서는 연구의 범위를 고정된 평활화 계수와 자료의 길이에 대한 HP 필터와 EMD에 의한 순환변동 및 추세 성분의 비교에 한정한다.

자료분석을 위하여 공개된 R 패키지를 사용하였으며, HP 필터와 EMD를 위하여 mFilter와 EMD 패키지를 각각 이용하였고 추세성분의 단위근 검정을 위하여 CADFtest 패키지를, 추세성분의 예측을 위하여 vars와 forecast 패키지를 이용하였다.

### 3.1. 순환변동 특성 비교

총 8개의 시계열에 대하여, 두 방법에 의한 추출된 순환변동들을 살펴보면 Figure 3.1과 같다. 비교를 위하여 전년도기대비 차이값들을 배경에 막대로 표시하였다. HP 필터에 의한 순환변동은 파란 실선으로, EMD에 의한 순환변동은 붉은 점선으로 표시하였다. 전체적인 흐름은 유사하나, 순환변동 추출 방법에 따라 세부적으로 순환변동의 값들이 다름을 볼 수 있다.

이렇게 추세제거 방법에 따라 다르게 얻어진 순환변동의 값들을 이용해, 기존 문헌에서는 추세제거 방법에 따른 경기순환점을 비교하기도 했다 (Nam, 2007). 그러나 순환변동의 값들은 추세제거 방법 뿐만 아니라, 관측 시간이 길어짐에 따른 구조변화나 추세제거 방법의 바운더리 처리와 같은 세부 옵션에 의해서도 계속해서 달라질 수 있다. 이러한 순환변동 값들의 직접적인 비교에 관한 문제는 향후 심층 연구 주제로 다루어질 수 있을 것이며, 본 논문에서는 기존의 문헌에서 주로 사용된 통계량들을 이용하여 비교하기로 한다.

기존 문헌에서는 순환변동에 대해 주로 다음과 같은 4가지 통계량에 의하여 그 특성을 살펴보았다 (Lee, 2009); (1) 표준편차에 의한 변동성, (2) GDP와의 시차상관계수(cross-correlation coefficient)를 이용한 동행성, (3) AR(1) 계수를 통한 지속성, (4) 각 변수간의 인과성을 파악하기 위한 그랜저 인과성(Granger-causality). 이러한 통계량들은 주로 통계작성 관련 기관들이나 관련 문헌들에서 순환변동의 특징을 기술하기 위해 사용되어 왔다. 통계학적으로 볼 때, 'GDP와의 시차상관계수'와 'GDP와의 인과관계 검정'은 의미상 중복되는 측면이 있고, 특히 인과관계 검정이 시차구조와 가정에 따라서 유의성이 바뀌는 경향이 있을 수 있으므로, 이러한 통계량들을 순환변동 추출법들 간의 우월성을 판단하는데 절대적인 기준으로 삼을 수는 없다. 그러나, 기존의 경제·금융 관련 문헌들의 연구내용과 비교하거나 통계관련 기관들에서 활용하기에는 유용한 장점을 가진다.

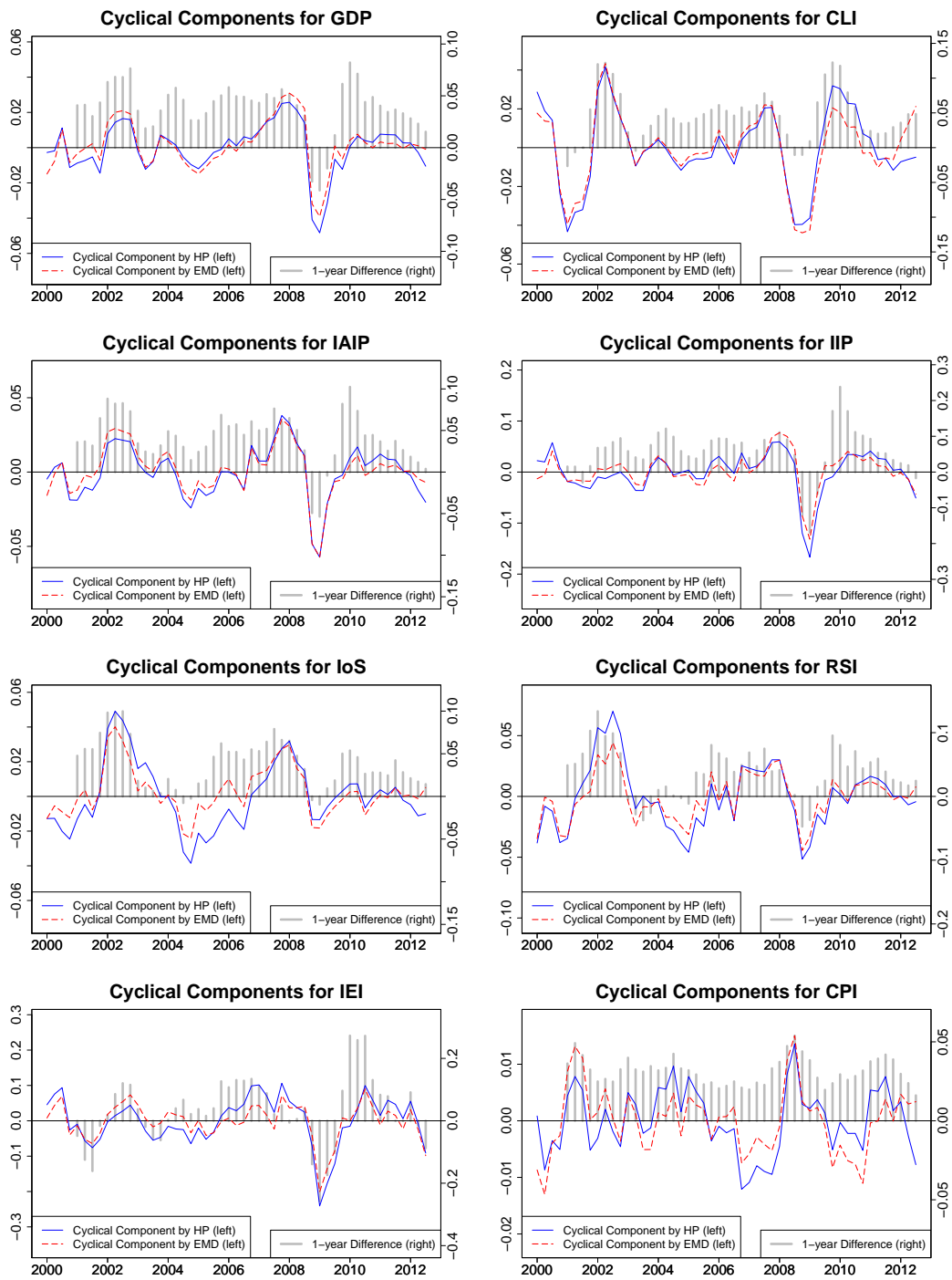


Figure 3.1. Cyclical components for Korean economic variables

**Table 3.1.** Properties for cyclical components by HP and EMD

Indices	Method	S.D.	Cross-correlation coefficient with GDP							AR(1)
			-3	-2	-1	0	1	2	3	
GDP	HP	0.014								0.659
	EMD	0.014								0.651
CLI	HP	0.020	0.269	0.549	0.630	0.479	0.100	-0.304	-0.496	0.765
	EMD	0.019	0.223	0.502	0.601	0.463	0.120	-0.247	-0.384	0.755
IAIP	HP	0.018	0.004	0.331	0.722	0.904	0.554	0.123	-0.127	0.671
	EMD	0.017	0.016	0.306	0.701	0.896	0.538	0.119	-0.090	0.658
IIP	HP	0.040	-0.051	0.216	0.583	0.849	0.562	0.090	-0.204	0.674
	EMD	0.035	-0.172	0.086	0.508	0.830	0.535	0.038	-0.203	0.591
IoS	HP	0.019	-0.065	0.217	0.466	0.536	0.364	0.191	0.053	0.829
	EMD	0.014	0.096	0.413	0.708	0.779	0.514	0.264	0.063	0.736
RSI	HP	0.027	0.208	0.409	0.660	0.637	0.371	0.079	-0.115	0.740
	EMD	0.020	0.197	0.411	0.731	0.720	0.368	0.042	-0.138	0.571
IEI	HP	0.069	0.084	0.245	0.523	0.783	0.715	0.373	-0.002	0.713
	EMD	0.053	-0.099	0.021	0.328	0.646	0.615	0.279	-0.055	0.570
CPI	HP	0.006	-0.451	-0.605	-0.531	-0.206	-0.099	0.056	0.190	0.538
	EMD	0.006	-0.218	-0.380	-0.288	0.044	0.129	0.211	0.282	0.566

Note: The names of the indices are abbreviated such as GDP for Gross Domestic Product, CLI for Composite Leading Index, IAIP for Index of All Industry Production, IIP for Index of Industrial Production, IoS for Index of Service, RSI for Retail Sales Index, IEI for Index of Equipment Investment, and CPI for Consumer Price Index.

먼저, Table 3.1은 HP 필터 및 EMD에 의하여 얻어진 각 순환변동을 비교하기 위하여 통계량 (1)~(3)을 계산한 것이다. 추출법에 따라 변수별 순환변동들의 변동성과 지속성은 유사하며, 동행성은 소비자물가지수를 제외하고 유사하게 나타난다.

Table 3.2은 순환변동들 간의 인과성을 살펴본 결과이다. 각 행( $X$ )에는 인과성을 주는 변수들이, 각 열( $Y$ )에는 인과성을 받는 변수들이 나열되어 있고, 괄호 안은 각 변수의 시차를 나타낸다. 귀무가설은 변수  $X$ 가 변수  $Y$ 에게 인과성을 가지지 않는다(not Granger cause)는 것이다. HP 필터와 EMD에 의한 두 결과가 서로 다른 부분들은 표에서 밑줄로 강조하였으며, 이를 유의수준 5%에서 정리하면 다음과 같다.

국내총생산(GDP)의 경우 EMD에 의해서 인과성을 주는 요인으로 서비스업생산지수(IoS)가 시차 1과 4에서 더욱 부각되고 있으며, 선행종합지수(CLI)의 경우 HP 필터에 의하여 전산업생산지수(IAIP, 시차 = 2)가 부각된다. 전산업생산지수(IAIP)의 경우 HP 필터의 결과가 소비자물가지수(CPI, 시차 = 2, 시차 = 4) 및 소매판매액지수(RSI, 시차 = 3, 시차 = 4)를, EMD의 결과는 설비투자지수(IEI, 시차 = 1) 및 광공업생산지수(IIP, 시차 = 2, 시차 = 4)를 인과성을 주는 요인으로 더욱 부각시키고 있다. 경험적인 전산업생산지수의 속성상 EMD의 결과가 더 타당한 인과성으로 보여진다. 광공업생산지수(IIP)의 경우 EMD의 결과가 HP 필터에 비해 소매판매액지수(RSI, 시차 = 1)를 인과성을 주는 요인으로 추가하고 있다. 다만, 광공업생산지수의 속성상 HP의 결과가 더 타당하다고 생각된다.

서비스업생산지수(IoS)의 경우 HP 필터의 결과가 소매판매액지수(RSI, 시차 = 1, 시차 = 2) 및 소비자물가지수(CPI, 시차 = 3)를 인과성을 주는 요인으로 판단하는 반면, EMD의 결과는 다른 7개의 변수 중에서는 인과성을 주는 요인이 없는 것으로 나타났다. 서비스업생산지수의 속성상 EMD의 결과가 좀 더 자연스러운 것으로 보여진다. 소매판매액지수(RSI)의 경우 EMD에 의한 결과가 선행종합지



**Table 3.2.** Granger causality for cyclical components by HP and EMD

Lag	HP ( $H_0 : X \nrightarrow Y$ )								EMD ( $H_0 : X \nrightarrow Y$ )							
	Y								Y							
$X(-1)$	GDP	CLI	IAIP	IIP	IoS	RSI	IEI	CPI	GDP	CLI	IAIP	IIP	IoS	RSI	IEI	CPI
		***					***		***						***	
	CLI	***		***	***		***		***		***	***		**	***	
	IAIP	***	**				***		***	**					***	
	IIP		***	*			***			***	***				***	
	IoS								***					*	***	
	RSI	***		**		**	*		***		**	**			***	
	IEI		***							***	***					
	CPI	***		***	***	***	**		***		**	***			***	
$X(-2)$	GDP	CLI	IAIP	IIP	IoS	RSI	IEI	CPI	GDP	CLI	IAIP	IIP	IoS	RSI	IEI	CPI
			*			**	***							*	***	
	CLI	***		**	**	**	***		**		**	***		***	***	
	IAIP	**	**			**	***		**					*	***	
	IIP		**				***			***	**			**	***	
	IoS	**				***	*		***					***	**	
	RSI	***		**		**			***		*				***	
	IEI									*				**		
	CPI	***	***	***	***	**	**		**	***		***			*	
$X(-3)$	GDP	CLI	IAIP	IIP	IoS	RSI	IEI	CPI	GDP	CLI	IAIP	IIP	IoS	RSI	IEI	CPI
							***								***	
	CLI	*		**	**		**		*		**	**		*	***	
	IAIP		**				***			**					***	
	IIP		*				***			*	*				***	
	IoS	**	***			***			***	***				**	*	
	RSI	**		**		*			***						***	
	IEI															
	CPI	***	*	***	**	**	**		*	**	*	***				
$X(-4)$	GDP	CLI	IAIP	IIP	IoS	RSI	IEI	CPI	GDP	CLI	IAIP	IIP	IoS	RSI	IEI	CPI
			*				***								***	
	CLI	*					*		*						***	
	IAIP	**	*				***		*					*	***	
	IIP						***			*	**				***	
	IoS		**			***	*		**	**				***	*	
	RSI	*		**		**	**		**						**	
	IEI	*														
	CPI	***	*	***	*	*	*		*	*		***				

Note: An asterik denotes significance at the 10% level, two at the 5% level and three at the 1% level, respectively.

수(CLI, 시차 = 1), 광공업생산지수(IIP, 시차 = 2) 및 설비투자지수(IEI, 시차 = 2)를 인과성을 주는 요인으로 더 포함하고 있어, EMD의 결과가 타당하다고 보인다. 마지막으로, 설비투자지수(IEI)의 경우 HP 필터에 의한 결과는 소비자물가지수(CPI, 시차 = 3)를, EMD에 의한 결과는 서비스업생산지

**Table 3.3.** ADF tests for original time series and trend components by HP and EMD

Indices	Original			HP			EMD		
	lag order	test	<i>p</i> -value	lag order	test	<i>p</i> -value	lag order	test	<i>p</i> -value
GDP	1	-3.07	0.1253	5	-3.26	0.0867	5	1.50	1.0000
CLI	2	-4.58*	0.0034	5	-3.56*	0.0453	4	-0.70	0.9665
IAIP	1	-2.42	0.3623	5	-1.35	0.8619	4	3.52	1.0000
IIP	1	-3.43	0.0607	5	0.20	0.9973	4	0.80	0.9996
IoS	1	-1.86	0.6560	10	1.01	0.9998	7	-0.81	0.9564
RSI	10	-4.24*	0.0086	3	-3.60*	0.0409	10	-3.06	0.1277
IEI	8	-3.82*	0.0249	4	-0.47	0.9813	3	-0.89	0.9484
CPI	4	-3.48	0.0544	8	1.62	1.0000	4	0.13	0.9967

Note: An asterik denotes significance at the 5% level.

수(IoS, 시차 = 1) 및 소매판매액지수(RSI, 시차 = 2, 시차 = 3)를 더 포함하는 것으로 나타났다. 전반적으로 EMD에 의한 결과가 좀 더 타당한 경우로 보여지는 시계열이 많다고 보이나 엄밀한 비교를 위한 모의실험 연구가 향후 수행되는 것이 바람직하다고 생각된다.

### 3.2. 추세 특성 비교

경제시계열 자료에서 흔히 발견되는 성질은 비정상성의 단위근(unit root)으로서 이는 임의보행성(random walk)이라고도 불리우며, 시계열 분석의 방향을 결정짓는 중요한 성질로 고려할 수 있다. 단위근 유무에 따라 차분 또는 확률적 추세 및 결정적 추세의 선택이 정해지기 때문이다 (Tsay, 2005). 따라서, EMD 및 HP 필터를 통하여 얻어진 추세 성분에 대하여 단위근 유무를 살펴보는 것은 큰 의미가 있다고 하겠다.

비교를 위하여, 8개 거시 경제 지표의 원시계열에 대해서도 단위근 검정을 시행하였으며, 단위근 검정은 ADF(Augmented Dickey-Fuller) 검정, 필요한 자기회귀모형(AR)의 차수는 AIC 정보량 기준을 사용하였다. GDP를 포함한 대부분의 거시경제지표들은 특성상 경제규모에 따라서 그 크기가 커지므로 원시계열의 단위근 검정시 일차의 시간변수를 모형에 추가하였으며, 각 필터에 의한 추세성분들 또한 그 정의에 의해서 시간변수를 모형에 추가하여 단위근 검정을 실시하였다. 이를 위하여 Hamilton (1994, p.501)을 참고하여야. Table 3.3은 원시계열과 EMD 및 HP 필터에 의하여 생성된 추세 성분에 대한 단위근 검정을 위한 차수(lag order), 검정통계량 및 *p*-값을 나타낸다.

유의수준 5%에서 살펴볼 때, EMD에 의한 결과로는 모든 추세 성분에서 단위근이 있다고 판단되었다. 반면, HP 필터에 의해서는 설비투자지수를 제외하고는 원시계열과 비슷한 단위근 검정 결과를 보여주었다. 즉, 선행종합지수와 소매판매액지수에 있어서 HP 필터에 의한 추세 성분은 원시계열과 동일하게 단위근이 없다는 결론을 주었다. 다음 절에서는 이러한 단위근 검정의 결과가 시계열 모형화에 의해서 예측에 어떤 영향을 미치는지 살펴볼 것이다.

### 3.3. 예측력 비교

본 절에서는 EMD와 HP 필터에 의해 분해된 성분들을 이용할 경우 예측력을 비교하고자 한다. 이를 위하여, EMD 및 HP 필터에 의하여 각각 생성된 순환변동 및 추세 성분을 2차원의 VAR(vector autoregressive) 모형으로 통합 모형화하였다. 추정된 모형은 두 성분의 예측을 위해서 사용되었으며, 예측된 두 성분의 결합을 통해서 원시계열을 예측하고 다음과 같은 평균 제곱근 오차(root mean squared

**Table 3.4.** RMSEs ( $\times 10^3$ ) for comparison of forecasting performances

Indices	Original	HP	EMD
GDP	3.59	4.15	<b>2.28</b>
CLI	<b>4.39</b>	5.66	15.78
IAIP	6.28	<b>4.68</b>	8.69
IIP	<b>24.21</b>	32.63	26.52
IoS	6.43	7.17	<b>5.15</b>
RSI	12.81	19.54	<b>6.17</b>
IEI	<b>30.43</b>	67.48	51.46
CPI	7.30	32.48	<b>3.69</b>

error; RMSE)를 계산하였다;

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2}, \quad \hat{y}_i = \hat{C}_i + \hat{T}_i.$$

단,  $y_i$ 는 원시계열의 실제값,  $\hat{y}_i$ 는 예측값,  $\hat{C}_i$  및  $\hat{T}_i$ 는 각각 예측된 순환변동 및 추세 성분을 나타내며,  $m$ 은 예측의 반복회수를 나타낸다. 이와 유사하게 EMD를 통하여 시계열자료를 예측하는 방법에 관하여서는 Kim 등 (2008)을 참고하여야.

예측력에 대한 비교를 위하여, 위의 계산된 결과는 원시계열을 일변량 ARIMA 모형으로 예측한 결과와 비교하였다. VAR 및 ARIMA 모형에 있어서 적절한 차수는 AIC 정보량 기준에 의해서 정하였으며, 각 분해법에 의해서 생성된 추세 성분은 단위근 검정을 통하여 비정상적으로 판단될 경우 차분을 행하였다. 순환변동은 모든 경우에 정상적으로 판단되었다.

RMSE의 계산을 위한 예측은 다음과 같은 방법을 통하여 반복적으로 이루어졌다. 먼저 2000년 1분기부터 2012년 3분기 자료를 이용하여 VAR 및 ARIMA 모형을 추정하였으며, 이를 이용하여 2012년 4분기 자료를 예측하고 예측오차를 구하였다. 다음은 2000년 1분기부터 2012년 4분기 자료를 이용하여 2013년 1분기 자료를 예측하고 예측오차를 구하였다. 이와 같이 1분기씩 추가적으로 시계열자료의 길이를 늘려가며 2013년 2분기 및 2013년 3분기 자료를 예측하였다. 즉, 총  $m = 4$ 번의 1-시점 미래 예측을 실시하였다.

Table 3.4은 총 8개의 경제지표에 대하여, 각 방법에 대한 RMSE의 값을 정리한 것이다; 단, 각 경제지표에 대하여 가장 작은 MSE 값들은 진하게 표시되었다. 모든 값들은 소수점 아래의 유효숫자를 위하여 원래의 값에  $10^3$ 을 곱하여 정리하였다.

EMD는 GDP를 포함하여 4가지 지표(서비스업생산지수, 소매판매액지수, 소비자물가지수)의 예측에서 가장 우수하였으며, HP는 전산업생산지수의 예측에서만 우위를 보였다. 원시계열의 ARIMA 또한 선행종합지수를 포함한 3가지 지표(광공업생산지수, 설비투자지수)의 예측에서 가장 우수한 결과를 보였다. 전반적으로 원시계열이 정상적으로 판단될 경우, 원시계열의 ARIMA 모형의 예측력이 더 우수한 것으로 보여진다; 단, 광공업생산지수는  $p$ -값이 0.0607로 원시계열이 비정상적으로 판단되었다. 원시계열이 비정상적일 경우, EMD 또는 HP 필터가 우수한 예측력을 보이는 것으로 판단된다. 또한, Figure 3.2에 의하며 살펴볼 때, 원시계열에 의한 예측의 RMSE는 비교적 안정된 결과를 주고 있으며, EMD는 선행종합지수, 광공업생산지수 및 설비투자지수를 제외한 5개 지표의 예측에서 RMSE의 값이 상대적으로 낮게 나타났다.

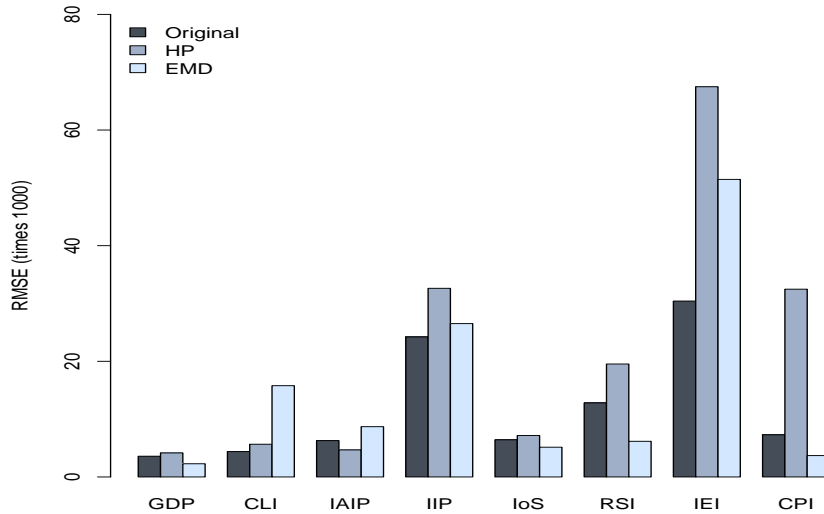


Figure 3.2. Bar Chart of RMSEs for comparison of forecasting performances

#### 4. 결론

본 논문에서는 순환변동과 추세 성분의 분리를 위하여 시간-진동수 영역에서의 시계열 분해법인 EMD를 소개하고, 이를 주요한 한국의 거시 경제 지표들에 적용하여 HP 필터와 비교하였다. 이를 위하여 각 분해법에 따라 얻어진 성분들의 통계적 특징들이나 통계적 예측 결과 등을 비교하였다. 이는 시간-진동수 영역에서의 분해법인 EMD를 기존의 방법들과 이론적으로 엄밀히 비교하거나 순환변동과 추세 성분을 추출하는 최적의 방법을 논하기 이전에, 각 방법들에 의하여 얻어진 성분들의 특성을 살펴 보고자 하는 것이다. 두 방법에 의하여 추출된 순환변동들은 전반적으로 유사한 흐름과 비슷한 값의 변동성, 동행성 및 지속성을 보여주었으나, 그랜저 인과성 측면에서는 EMD가 상대적으로 더 타당한 결론을 보여주었다. 추세 성분의 단위근 검정에서는 두 방법이 서로 다른 결론을 보여주었으며, 특히 모든 시계열들에 대하여 EMD에 의한 추세에는 단위근이 있다고 판단되었다. 순환변동 및 추세 성분을 조합한 예측력에 있어서도, EMD가 HP 필터보다 우수한 RMSE 값을 보여주었다. 이와 같은 연구결과를 바탕으로 향후 보다 심도 깊은 연구주제들-예를 들면, 금융위기와 같은 구조변화 및 바운더리의 처리 문제, 각 성분들의 스펙트럼의 분석 및 해석, EMD에 변동성 개념의 적용, 비정상성을 가지는 성분들의 예측 및 공통추세(common trend)의 모형화-과 우리나라의 실정에 적합한 순환변동과 추세 성분의 추출에 관한 이론적 연구의 발판을 마련하고자 한다.

#### References

- Boashash, B. (1992). Estimating and interpreting the instantaneous frequency of a signal - part 1: fundamentals, *Proceedings of the IEEE*, **80**, 519–538.
- Cho, S. and Son, Y. (2009). *Time Series Analysis by Using SAS/ETS*, Yulgok Books, Seoul.
- Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University Press, Princeton.
- Hodrick, R. J. and Prescott, E. C. (1997). Postwar U.S. business cycles: An empirical investigation, *Journal of Money, Credit, and Banking*, **29**, 1–16.

- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C. and Liu, H. H. (1998). The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis, *Proceeding of the Royal Society London A*, **454**, 903–995.
- Kim, D. and Oh, H.-S. (2009). EMD: A Package for Empirical Mode Decomposition and Hilbert Spectrum, *The R Journal*, **1**, 40–46.
- Kim, D., Paek, S.-H. and Oh, H.-S. (2008). A Hilbert-Huang transform approach for predicting cyber-attacks, *Journal of the Korean Statistical Society*, **37**, 277–283.
- Kim, D., Kim, K. and Oh, H.-S. (2012). Extending the scope of empirical mode decomposition by smoothing, *EURASIP Journal on Advances in Signal Processing*, **1**, 168.
- Lee, J. (2009). Changes in the business cycle of the Korean economy: Evidence and explanations, *KDI Journal of Economic Policy*, **31**, 47–85.
- Mallat, S. (1998). *A Wavelet Tour of Signal Processing*, Academic press.
- Marcet, A. and Ravn, M. O. (2003). *The HP-Filter in Cross-Country Comparisons*, CEPR DP 4244.
- Nam, S.-H. (2007). An analysis on the characteristics of recent business cycles, *Economic Analysis*, **13**, 79–109.
- Park, M., Kim, D. and Oh, H.-S. (2011). A reinterpretation of EMD by cubic spline interpolation, *Advanced in Adaptive Data Analysis*, **3**, 527–540.
- Rilling, G. and Flandrin, P. (2008). One or two frequencies? The empirical mode decomposition answers, *IEEE Transactions on Signal Processing*, **56**, 85–95.
- Ryu, M. and Lee, H. S. (2007). Measuring potential output and forecasting future inflation, *Sogang Economic Papers*, **36**, 61–83.
- Tsay, R. S. (2005). *Analysis of Financial Time Series*, 2nd Edition, Wiley.

# 순환성분 추출을 위한 EMD와 HP 필터의 비교분석: 한국의 거시 경제 지표에의 응용

박민정<sup>a</sup> · 성병찬<sup>b,1</sup>

<sup>a</sup>통계청 통계개발원, <sup>b</sup>중앙대학교 응용통계학과

(2014년 2월 13일 접수, 2014년 3월 28일 수정, 2014년 4월 5일 채택)

---

## 요약

본 논문에서는 시간-진동수 영역에서 시계열을 여러 구성 성분으로 분해하는 방법인 경험적모드분해법(Empirical Mode Decomposition)을 소개하고, 이를 이용하여 한국의 주요 거시 경제 지표를 대상으로 순환변동과 추세 성분을 추출하고 예측에 활용한다. 그 효율성을 살펴보기 위하여, 추출된 구성 성분들의 변동성, 동행성, 지속성, 인과성, 비정상성 및 예측력을 계산하고, 가장 보편적으로 널리 사용되고 있는 Hodrick-Prescott 필터에 의한 결과와 비교한다.

주요용어: 경험적모드분해법, Hodrick-Prescott 필터, 시간-빈도 분석법.

---

---

이 논문은 2013년도 중앙대학교 성병찬의 연구년 결과물로 제출됨.

<sup>1</sup>교신저자: (156-756) 서울시 동작구 흑석동 221번지, 중앙대학교 경영경제대학 응용통계학과.

E-mail: bcseong@cau.ac.kr