



Original Article

금강하구역 DPSIR 지표 간 동태적 관계에 대한 분석

김대환 · 민동기[†]
건국대학교 경제학과

Dynamic Relationship among DPSIR Indicators for Geumgang Estuary

Dae Hwan Kim and Dong Ki Min[†]

Department of Economics Konkuk University, Seoul 05029, Korea

요 약

구동·압력·상태·영향·반응(DPSIR) 지표체계는 (1) 구동·압력요인으로부터 상태·영향요인으로, (2) 상태·영향요인으로부터 반응요인으로, (3) 다시 반응요인으로부터 상태·영향요인으로의 인과관계를 가정하고 있다. 그러나 이러한 동태적 관계에 대한 정량적 분석은 찾아보기 힘든데, 이는 DPSIR 지표 체계에 포함되는 환경·생태 변수들의 경우 긴 시계열을 확보하기 어려워서 이 변수들과 관련된 동태관계를 추정하는 것이 매우 어렵기 때문이다. 본 연구의 첫번째 목적은 DPSIR에서 가정하고 있는 동태관계에 대한 실증근거를 제시하는 것이다. 이를 위해 DPSIR에서 가정하고 있는 동태관계를 대립가설로 설정한 후, 가설검정을 실시하였다. 짧은 시계열로 인해 발생하는 어려움을 극복하기 위해 연속변수를 이산화(discretization)시키는 방법도 적용해 보았다. 본 연구의 두번째 목적은 DPSIR 변수들 간의 동태관계를 간결하게 표현해 주는 모형을 제시하는 것이다. 이 모형을 이용하여 기존의 시나리오 분석을 개선시킬 수 있을 것으로 기대된다. 본 연구의 실증분석은 금강하구역 데이터를 사용하였다.

Abstract – The Driver-Pressure-State-Impact-Response (DPSIR) framework assume causal relationship from driver and pressure indicators to state and impact indicators, from state and impact indicators to response indicators, and from response indicators to state and impact indicators. However, quantitative analysis of such dynamic relationship is rare because only short times series are available for ecological variables included in the DPSIR framework, which makes it is difficult to estimate dynamic relationship. The first goal of this research is to present empirical evidence for the dynamic relationship that is assumed in the DPSIR framework. To do so, we set up the assumed dynamic relationship as the alternative hypothesis and carry out hypothesis testing. To overcome the difficulty arising from short time series, we have tried the approach of discretizing continuous variables. The second goal of this research is to present a simple dynamic model of DPSIR variables. This model will help improve the DPSIR-based scenario analysis. The empirical analysis of this research is based on Geumgang estuary data.

Keywords: Geumgang estuary(금강하구역), DPSIR framework(DPSIR 지표 체계), dynamic relationship(동태적 관계), discretization of continuous variables(연속변수의 이산화), scenario analysis(시나리오 분석)

1. 서 론

해안지역에서의 인간활동과 생태계 간의 상호작용을 종합적으로 이해하기 위해 지속가능성 지표체계가 빈번히 사용되는데, 특히 구동·압력·상태·영향·반응(DPSIR) 프레임워크를 기반으로 한 지표 체계가 많이 사용된다. ABP Marine Environmental Research[2008], Vlaams-Nederlandse Schelde Commissie[2011], Gang *et al.*[2005],

Gang and No[2016], Korea Marine Institute[2007], Korea Environment Institute[2014] 등이 DPSIR 프레임워크를 사용하여 해안지역을 대상으로 한 지표체계를 개발한 사례이다. Gari *et al.*[2015]는 하구역을 포함한 해안지역에 DPSIR 지표체계를 적용한 문헌 41개를 검토하였음을 언급한 바 있다. Lewison *et al.*[2016]은 보다 체계적인 문헌 검색을 통해서 해안지역에 DPSIR 지표체계가 적용된 연구를 231개 확인하였음을 언급한 바 있다.

DPSIR 지표체계를 적용한 연구의 수는 매우 많지만, 자료의 제

[†]Corresponding author: dkm2@konkuk.ac.kr

약으로 인해 통계 분석을 진행한 연구는 매우 드물다. DPSIR 지표 체계는 경제·사회 변수와 환경·생태 변수를 모두 포함하게 마련인데, 이 중 환경·생태 변수의 경우, 구할 수 있는 시계열의 길이가 매우 짧아 제대로 된 통계분석을 진행하는 것이 매우 어렵다는 점이 주된 이유다. 관련된 문헌을 체계적으로 검토한 Gari *et al.*[2015]과 Lewison *et al.*[2016]는 공통적으로 이 점을 지적하고 있다.

본 연구의 첫번째 목적은 DPSIR에서 가정하고 있는 동태관계에 대한 실증근거를 제시하는 것이다. DPSIR의 가장 기본적 가정은 구동·압력요인과 상태·영향요인 간, 그리고 상태·영향요인과 반응요인 간 인과관계가 존재한다는 것인데, 이를 실증적으로 검토한 연구는 존재하지 않는다. 본 연구에서는 DPSIR에서 가정하고 있는 동태관계를 대립가설로 설정한 후, 가설검정을 실시하였다. 짧은 시계열로 인해 발생하는 어려움을 극복하기 위해 우선 추정해야 하는 파라미터의 수를 최소화하였다. 자료생성과정(data generating process)을 명시적으로 보여주는 모형을 설정하여 추정하는 대신, 동태적 상관관계만으로 가설검정을 실시하였다. 검정통계량으로는 t 통계량의 최대치 혹은 최소치를 사용하였다. 어떤 시차(lag)에 대해서도 DPSIR 지표체계에서 가정하는 상관관계가 존재하지 않는다는 것을 귀무가설로 설정하였기 때문에, t 통계량의 최대치 혹은 최소치가 자연스럽게 검정통계량으로 선택되었다. 이러한 순서 통계치(order statistics)의 경우, p 값 등을 계산하기 위해서는 부트스트랩(bootstrap) 기법을 사용하는 것이 보통이어서 본 연구에서도 이를 따랐다.

본 연구에서는 또한 연속변수를 이산화(discretization)하는 방법을 통해 추정모형을 단순화하는 방법도 고려해 보았다. 연속변수를 이산화하는 것은 자료에 포함되어 있는 정보의 대부분을 무시하는 것이므로 추정의 효율성을 떨어뜨릴 수 있다는 점은 잘 알려져 있다. 의학 및 보건학에서는 Morgan and Elashoff[1986], Faraggi and Simon[1996], Owen and Froman[2005], Naggara *et al.*[2011] 등이, 마케팅에서는 Irwin and McClelland[2003] 등이, 전산학에서는 Mariooryad and Busso[2017] 등이, 통계학에서는 Kuss[2013] 등이 이런 관점을 강조한 바 있다. 반면 연속변수의 이산화를 통해 잘 드러나지 않는 동태관계를 확인할 수 있다는 관점도 존재한다. 특히 금융분야에서 이런 관점이 많이 논의되었는데, Wu and Zhang[1997], Leung *et al.*[2000], Anatolyev and Gospodinov[2010], Nyberg[2011] 등의 연구를 예로 들 수 있다. DPSIR 지표체계를 지수화 하는 방식을 논한 Perry and Masson[2013]도 이러한 관점을 취하고 있다. Perry and Masson[2013]은 연속변수로부터 단절점(break point)을 찾아내고, 이를 이산변수로 전환한 후 이를 지수화 하는 방식을 채택한 바 있다. 따라서 연속변수의 이산화가 긍정적 역할을 할 가능성에 대한 추가적 분석 필요성이 존재하고, 환경·생태 변수가 사용되는 분석의 경우에는 특히 그러하다 할 수 있다. 본 연구에서는 보다 단순한 방식으로 연속변수를 이산화한 후 변수들 간의 동태관계에 대한 가설검정을 진행한다.

본 연구의 두번째 목적은 DPSIR변수들 간의 동태적 관계를 간결하게 표현해 주는 모형을 제시하는 것이다. 이러한 모형은 특히

Table 1. Geumgang Estuary Indicator System

Indicators	
Driver	Population
	Agricultural production
	Livestock production
	Industrial production
	Port traffic
Rainfall volatility	
Pressure	Wastewater
	Organic waste
	Untreated wastewater
State	Untreated organic waste
	Geumgang biochemical oxygen demand(GBOD)
	Geumgang chemical oxygen demand(GCOD)
	Geumgang dissolved oxygen(GDO)
	Geumgang total phosphorus(GTP)
	Geumgang total nitrogen(GTN)
	Coastal chemical oxygen demand(CCOD)
Coastal total phosphorus(CTP)	
Impact	Coastal dissolved oxygen(CDO)
	Shellfish production(SFP)
	Seaweed production(SWP)
Response	Access to sanitation(ATS)
	Wastewater inspection(WWI)

Nobre[2009] 등이 시도한 DPSIR 기반 시나리오 분석을 개선시킬 수 있을 것으로 기대된다. 기존의 시나리오 분석에서 변수 선정은 DPSIR 체계를 따르지만, 실증분석을 위해서 가치평가방법 등의 추가적인 분석을 병행하는 방식을 취하고 있다. 본 연구에서는 DPSIR에서 가정하고 있는 동태관계를 반영하는 단순한 모형을 제시함으로써, 보다 일관성 있게 시나리오 분석을 진행할 수 있는 방법을 제안한다.

본 연구의 분석에는 금강하구역 데이터를 사용하였다. 총 23개의 DPSIR 지표가 분석에 포함되었는데 포함된 지표는 <Table 1>에 나열한 바와 같다.¹⁾

본 논문의 구성은 다음과 같다. 제2절에서는 DPSIR에서 가정하고 있는 동태관계를 검증하기 위한 가설을 설정하고, 가설검정 결과를 소개한다. 제3절에서는 가설검정 결과를 반영하는 간결한 동태 모형을 제시한다. 제4절은 결론을 담고 있다.

2. 가설의 설정 및 검정

2.1 가설 설정

DPSIR의 기본적 가정은 (1) 구동·압력요인으로부터 상태·영향요인으로의 음의 인과관계가 존재하고, (2) 상태·영향요인으로부터 반응요인으로의 음의 인과관계가 존재하며, (3) 다시 반응요인으로부터 상태·영향요인으로의 양의 인과관계가 존재한다는 것이다. 구

1)금강하구역을 대상으로 한 DPSIR 지표체계의 개발 과정은 Kim and Min[2017]에서 자세히 설명한 바 있으므로 본고에서는 이를 생략하겠다.

동·압력요인은 생태계의 지속가능성에 간접적, 직접적으로 부정적 영향을 끼칠 수 있는 인간의 사회경제적 활동을 나타내는 요인들이다. 상태·영향요인은 인간에게 유익한 가치를 창출하는 생태계의 기능이 원활히 작동하고 있는지, 즉 생태계 서비스(eco-system service)가 원활히 제공되고 있는지를 보여주는 요인들이다. 마지막으로 반응요인은 생태계 서비스의 원활한 제공을 위한 인간의 적극적 활동을 보여주는 요인이다. 따라서 구동·압력요인의 상승(인간활동의 증가)은 상태·영향요인의 하락(생태계 서비스의 감소)을 초래하고, 상태·영향요인의 하락은 반응요인의 상승(생태계 서비스 회복을 위한 노력의 증가)을 초래하며, 반응요인의 상승은 상태·영향요인의 상승을 초래할 것으로 예상할 수 있다.

우선 구동·압력요인으로부터 상태·영향요인으로서의 음의 인과관계를 어떻게 분석할 지 생각해 보자. 구동·압력요인을 $\{x_{1,t}\}, \dots, \{x_{M,t}\}$ 로 표시하고, 상태·영향요인을 $\{y_{1,t}\}, \dots, \{y_{M,t}\}$ 으로 표시하자. 구동·압력요인 중 하나를 임의로 선택하여 이를 $x_{i,t}$ 라 하고, 상태·영향요인 중 하나를 임의로 선택하여 이를 $y_{j,t}$ 라 하자. $x_{i,t}$ 로부터 $y_{j,t}$ 로의 음의 인과관계가 존재한다면 $x_{i,t-s}$ 와 $y_{j,t}$ 간의 상관관계가 음이 되도록 하는 $s \in \{1, 2, 3, \dots\}$ 가 존재할 것이다. 반면 $x_{i,t}$ 로부터 $y_{j,t}$ 로의 음의 인과관계가 존재하지 않는다면 모든 $s \in \{1, 2, 3, \dots\}$ 에 대해서 $x_{i,t-s}$ 와 $y_{j,t}$ 간의 상관관계는 0이거나 0보다 클 것이다. 이를 다시 표현하면 모든 $s \in \{1, 2, 3, \dots\}$ 에 대해서 $x_{i,t-s}$ 와 $y_{j,t}$ 간의 상관관계를 계산한 후 최소값을 구하면 0이거나 0보다 크다고 할 수 있다. 따라서 “ $x_{i,t}$ 로부터 $y_{j,t}$ 로의 음의 인과관계가 존재하지 않음”을 귀무가설로 설정하면 다음과 같이 표현할 수 있다:

$$H_0: \min[\text{cov}(y_{j,t}, x_{i,t-1}), \text{cov}(y_{j,t}, x_{i,t-2}), \dots] \geq 0 \quad (1)$$

대립가설은 “ $x_{i,t}$ 로부터 $y_{j,t}$ 로의 음의 인과관계가 존재함”이고 다음과 같이 표현할 수 있다:

$$H_1: \min[\text{cov}(y_{j,t}, x_{i,t-1}), \text{cov}(y_{j,t}, x_{i,t-2}), \dots] < 0 \quad (2)$$

$\{(x_{i,t}, y_{j,t}): i=1, \dots, M; j=1, \dots, N\}$ 의 각 원소에 대해 개별적으로 가설 검정을 진행할 수 있다. 상태·영향요인으로부터 반응요인으로 음의 인과관계와 반응요인으로부터 상태·영향요인으로의 양의 인과관계도 같은 방식으로 분석할 수 있다.

2.2 가설 검정 방법

두 시계열 $\{x_t\}$, $\{y_t\}$ 이 주어지고, x_t 로부터 y_t 로의 음의 인과관계가 존재하는지 확인하려는 경우를 생각해 보자. 우선 $l \in \{1, 2, \dots, L\}$ 에 대해 다음 회귀식을 추정할 수 있다:

$$y_t = \alpha_l \beta_l x_{t-l} + \varepsilon_{t,l}, \varepsilon_{t,l} \sim N(0, \sigma_l^2), t = l+1, \dots, T \quad (3)$$

“ β_l 가 가장 유의한 음수일 때의 l 값”을 l^* 로 표시하자. l^* 는 다음과 같이 표현할 수 있다:

$$l^* = \arg \max_{l \in \{0, 1, \dots, L\}} \tau_l \quad (4)$$

τ_l 는 t 통계량으로 다음과 같이 정의된다:

$$\tau_l = \frac{\text{cov}(y_t, x_{t-l})}{\sigma_l \sqrt{\text{var}(x_t)}} \quad (5)$$

l^* 의 추정치를 구하는 것은 단순하다.

앞서 제시한 귀무가설과 대립가설은 다음과 등가이다.

$$H_0: \min\{\beta_0, \dots, \beta_L\} \leq 0 \quad (6)$$

$$H_1: \min\{\beta_0, \dots, \beta_L\} > 0 \quad (7)$$

또한 다음과도 등가이다.

$$H_0: \min\{\tau_0, \dots, \tau_L\} \leq 0 \quad (8)$$

$$H_1: \min\{\tau_0, \dots, \tau_L\} > 0 \quad (9)$$

τ_l 에 대한 추정량을 $\hat{\tau}_l$ 이라 하고, $\min\{\hat{\tau}_0, \dots, \hat{\tau}_L\}$ 를 $\hat{\tau}$ 이라 부르자. $\hat{\tau}$ 에 대한 분포를 특정할 수 있으면 위 귀무가설과 대립가설을 검정하는 문제는 단순해진다.

가설검정은 부트스트랩을 통해서 진행한다. $\hat{\tau}$ 의 분포를 정하기 위해서는 $\hat{\tau}_0, \dots, \hat{\tau}_L$ 의 결합 분포를 정하여야 하고, 이를 위해서는 $\widehat{\text{var}}(x_t)$, $\hat{\beta}_0, \dots, \hat{\beta}_L$, $\hat{\sigma}_0, \dots, \hat{\sigma}_L$ 의 결합 분포를 정해야 한다. 문제의 복잡도를 고려할 때 $\widehat{\text{var}}(x_t)$, $\hat{\beta}_0, \dots, \hat{\beta}_L$, $\hat{\sigma}_0, \dots, \hat{\sigma}_L$ 의 결합 분포를 정하는 가장 현실적 방법은 부트스트랩이다. 부트스트랩 방법을 논하기 위해서는 추정량과 추정치를 구별하는 것이 필요하다. 추정량은 캐럿(^)을 이용하여 표시하고, 추정치는 틸더(~)를 이용하여 표시하기로 하자. 첫 번째 유사표본(pseudo sample)은 다음 식을 이용하여 정한다:

$$y_t^{(1)} = \hat{\alpha}_0 + \hat{\beta}_0 x_t + \varepsilon_{t,0}^{(1)}, \varepsilon_{t,0}^{(1)} \sim N(0, \hat{\sigma}_0^2), t = 1, \dots, T \quad (10)$$

l^* 의 회귀식으로부터 유사표본을 만들어 내는 것이 더 적절하다고 볼 수도 있지만 계산을 단순화하기 위해 시차(lag)가 없는 경우를 이용하겠다. 첫 번째 유사표본으로부터 얻어진 추정치를 $\widehat{\text{var}}(x_t)$, $\hat{\beta}_0^{(1)}, \dots, \hat{\beta}_L^{(1)}$, $\hat{\sigma}_0^{(1)}, \dots, \hat{\sigma}_L^{(1)}$ 라 부르자. $\widehat{\text{var}}(x_t)$ 는 모든 유사표본에서 동일한 값이 얻어진다. 이로부터 구한 $\min\{\tau_0, \dots, \tau_L\}$ 의 추정치를 $\hat{\tau}^{(1)}$ 이라 부르자. 유사표본을 반복적으로 만들어서 $\hat{\tau}^{(1)}, \dots, \hat{\tau}^{(K)}$ 를 구한 후 대립가설에 대한 p 값은 다음과 같이 계산한다:

$$\Pr(\hat{\tau}^{(k)} > 0 | k \in \{1, \dots, K\}) \quad (11)$$

x_t 로부터 y_t 로의 양의 인과관계가 존재하는 경우 부등호의 방향을 반대로 적용하면 된다.

2.3 결과

DPSIR 지표체계에 포함된 변수들 대부분은 단위근을 포함하고 있음을 Phillips and Perron[1988] 검정을 통해 확인할 수 있다. 단위근을 제거하기 위해 각 변수들을 1계차분한 후 이후의 분석을 진행하였다.

<Table 2>는 구동·압력요인과 상태·영향요인 간의 동태관계에 대한 추정 및 가설검정 결과의 일부를 보여준다. 가장 유의도가 높은 시차 l^* 는 1개월에서 48개월 사이에 고르게 분포되어 있는 것을 확

Table 2. Dynamic Relation between Driver, Pressure Indicators (X) and State, Impact Indicators (Y)

		Y			
		GBOD	GDO	GTP	GTN
X	Population	48 ***	48 ***	35 **	46 **
	Port traffic	20 ***	25 ***	25 ***	24 ***
	Agricultural Prod.	4 ***	33 ***	4 ***	3 ***
	Industrial Prod.	28 ***	28 ***	33 ***	27 ***
	Livestock Prod.	13 ***	19 ***	7 ***	8 ***
	Rainfall Vola.	20 ***	44 ***	2 ***	7 ***
	Waterwaste	16 ***	18 ***	18 ***	2 ***
	Organic waste	26 **	18 ***	18 ***	2 ***
	Untreated waterwaste	16 ***	18 ***	18 ***	2 ***
	Untreated organic waste	26 *	18 ***	18 ***	2 ***

The figures reported are the estimates of l^* , the most significant lags. *, **, *** shown next to the figures indicate statistical significance at 1%, 5%, and 10%, respectively.

Table 3. Dynamic Relation between State, Impact Indicators (X) and Response Indicators (Y)

		Y	
		ATS	WWI
X	GBOD	43 ***	8 ***
	GDO	22 **	8 ***
	GTP	41 ***	18 ***
	GTN	34 ***	34 ***
	GCOD	41 ***	41 ***
	CDO	47	36 ***
	CTP	34	34
	CTN	12 ***	12 ***
	CCOD	48 ***	48 ***
	SFP	11	47 ***
	SWP	44 ***	44 ***

The figures reported are the estimates of l^* , the most significant lags. *, **, *** shown next to the figures indicate statistical significance at 1%, 5%, and 10%, respectively.

인할 수 있다. l^* 의 최대값 L 을 48개월로 설정하였기 때문에 l^* 의 추정치가 48개월로 나온 경우는 해석에 주의가 필요하다. 즉 가장 유

Table 4. Dynamic Relation between Response Indicators (X) and State, Impact Indicators (Y)

		Y					
		GBOD	GDO	GTP	GTN	GCOD	CDO
X	ATS	48 *	48 ***	7 ***	2 **	7 **	47
	WWI	48	48 ***	7 ***	2 ***	24 **	36 ***

		Y				
		CTP	CTN	CCOD	SFP	SWP
X	ATS	12 ***	48 ***	48 ***	37 ***	15 ***
	WWI	24 ***	36 ***	24 ***	39 ***	48 ***

The figures reported are the estimates of l^* , the most significant lags. *, **, *** shown next to the figures indicate statistical significance at 1%, 5%, and 10%, respectively.

의도가 높은 시차는 48개월보다 클 가능성이 있다. 대부분의 경우 l^* 값은 48개월보다 작기 때문에, 가장 유의도가 높은 시차는 48개월보다 작다는 해석이 가능하다. 즉 구동·압력요인이 상태·영향요인에 가장 큰 영향을 미치는 시차는 48개월 미만인 것으로 해석할 수 있다. 통계적 유의성을 보면, 거의 모든 경우 동태관계가 매우 유의한 것으로 나온다. 즉 구동·압력요인과 상태·영향요인 간 음의 인과관계가 존재하지 않는다는 귀무가설을 기각할 수 있다. <Table 3>과 <Table 4>는 상태·영향요인과 반응요인 간 동태관계에 대한 추정 및 가설검정 결과를 보여준다. <Table 2>에서와 마찬가지로 l^* 추정치는 1개월에서 48개월 사이에 고르게 분포되어 있고, 요인 간 인과관계가 존재하지 않는다는 귀무가설은 대부분의 경우 기각된다.

2.4 이산변수를 바탕으로 한 분석

우선 각 시계열에 대해 전기 값 대비 증가(+1), 감소(-1), 혹은 변화 없음(0)을 나타내는 이산변수를 만든다. 이렇게 만들어진 이산변수 중 두 개를 $\{x_t\}$, $\{y_t\}$ 로 표시한다. 두 시계열 간의 동적관계를 살펴보기 위해 다음 전이행렬을 추정할 수 있다.

$$\Pi_t = \begin{pmatrix} \Pr(y_t = 1|x_t = 1) & \Pr(y_t = 1|x_t = 0) & \Pr(y_t = 1|x_t = -1) \\ \Pr(y_t = 0|x_t = 1) & \Pr(y_t = 0|x_t = 0) & \Pr(y_t = 0|x_t = -1) \\ \Pr(y_t = -1|x_t = 1) & \Pr(y_t = -1|x_t = 0) & \Pr(y_t = -1|x_t = 1) \end{pmatrix} \quad (12)$$

두 시계열 간 음의 상관관계가 있는 경우, 승산비(odd ratio, hit ratio)를 다음과 같이 정의할 수 있다:

$$\tau_t = \Pr(y_t = 1, x_t = -1) + \Pr(y_t = -1, x_t = 1) - \Pr(y_t = -1, x_t = -1) - \Pr(y_t = 1, x_t = 1) \quad (13)$$

τ_t 에 대한 추정량을 $\hat{\tau}_t$ 이라 하고, “승산비가 가장 유의할 때의 l^* 값”을 l^* 는 식 (4)에서 정의한 바와 같다. $\max\{\hat{\tau}_0, \dots, \hat{\tau}_L\}$ 를 $\hat{\tau}$ 으로 표시하면 가설 검정은 $\hat{\tau}$ 을 이용하여 진행할 수 있다. 앞서와 마찬가지로 $\hat{\tau}$ 의 분포는 부트스트랩을 이용하여 진행한다. k 번째 유사표본은 다음

Table 5. Dynamic Relation between Driver, Pressure Indicators (X) and State, Impact Indicators (Y): Discretization

X		Y			
		GBOD	GDO	GTP	GTN
Population		3	1	4	3
		***	***	***	***
Port traffic		1	7	29	24
		***	***	***	***
Agricultural Prod.		2	35	4	3
		***	***	***	***
Industrial Prod.		44	10	10	3
		***	***	***	*
Livestock Prod.		27	46	13	22
		***	***	***	***
Rainfall Vola.		10	43	2	38
		***	***	***	***
Waterwaste		8	23	34	48
		***	***	***	***
Organic waste		31	23	34	16
		***	***	***	***
Untreated waterwaste		8	16	34	6
		***	***	***	***
Untreated organic waste		42	14	1	22
		***	***	***	***

The figures reported are the estimates of l^* , the most significant lags. *, **, *** shown next to the figures indicate statistical significance at 1%, 5%, and 10%, respectively.

식을 이용하여 정한다:

$$y_i^{(k)} \text{ is } +1, 0, -1 \text{ with the probabilities given by } \hat{\Pi}_0 \begin{pmatrix} (x_i = 1) \\ (x_i = 1) \\ (x_i = -1) \end{pmatrix} \quad (14)$$

유사표본으로부터 $\tilde{\Pi}_0^{(k)}, \dots, \tilde{\Pi}_L^{(k)}$ 를 계산하고, 이로부터 $\tilde{\tau}^{(k)}$ 를 구할 수 있다. 이후의 분석과정은 앞서 설명한 대로이다.

<Table 5>는 이산변수를 바탕으로 한 구동·압력요인과 상태·영향요인 간의 동태관계에 대한 추정 및 가설검정 결과의 일부를 보여준다. <Table 2>의 결과와 비교해 볼 때 추정치가 다소 바뀌기는 하지만 전체적인 결론은 크게 다르지 않다. 즉 l^* 추정치는 1개월에서 48개월 사이에 고르게 분포되어 있고, 요인 간 인과관계가 존재하지 않는다는 귀무가설은 대부분의 경우 기각된다. 상태·영향요인과 반응요인 간의 동태관계에 대해서도 동일한 분석을 진행하였고, <Table 3>, <Table 4>에 보고한 것과 비슷한 결과를 얻었다. 지면을 고려하여 자세한 추정결과는 보고하지 않겠다.

3. 동태 모형

3.1 모형의 구성 및 추정

DPSIR 지표체계를 바탕으로 시나리오 분석을 진행하기 위해서는 단순한 동태모형을 설정하는 것이 필요하다. 시나리오 분석에서는 상태·영향·반응 요인간의 관계에 집중하는 경우가 많기 때문에 여

Table 6. Dynamic Model of State, Impact, and Response Indicators

GBOD(t) = 1.07*ATS(t-48)
GDO(t) = 0.31*ATS(t-48) + 0.02*WWI(t-48)
GTP(t) = 0.48*ATS(t-7) + 0.04*WWI(t-7)
GTN(t) = 0.23*ATS(t-2) + 0.02*WWI(t-2)
GCOD(t) = 0.27*ATS(t-7) + 0.02*WWI(t-24)
CDO(t) = 0.04*WWI(t-36)
CTP(t) = 2.47*ATS(t-12) + 0.11*WWI(t-24)
CTN(t) = 1.88*ATS(t-48) + 0.24*WWI(t-36)
CCOD(t) = 1.28*ATS(t-48) + 0.30*WWI(t-24)
SFP(t) = 4.60*ATS(t-37) + 0.49*WWI(t-39)
SWP(t) = 3.24*ATS(t-15) + 0.33*WWI(t-48)
ATS(t) = - 0.07*GBOD(t-43) - 0.17*GDO(t-22) - 0.09*GTP(t-41)
- 0.13*GTN(t-34) - 0.10*GCOD(t-41) - 0.04*CTN(t-12)
- 0.03*CCOD(t-48) - 0.04*SFP(t-44)
WWI(t) = - 0.69*GBOD(t-8) - 2.83*GDO(t-8) - 1.25*GTP(t-18)
- 3.35*GTN(t-34) - 1.71*GCOD(t-41) - 1.80*CDO(t-36)
- 1.80*CDO(t-36) - 1.01*CTN(t-12) - 1.48*CCOD(t-48)
- 0.20*SFP(t-47) - 0.40*SWP(t-44)

기에서는 <Table 3>과 <Table 4>에 제시된 추정결과를 바탕으로 상태·영향·반응 요인에 대한 단순화된 예측식을 제시하고자 한다.

<Table 6>은 상태·영향·반응 요인 각각에 대한 예측식을 보여준다. 선형회귀모형에서 설명변수들이 서로 독립적일 경우, 하나의 설명변수 만을 포함하여 회귀분석을 진행하더라도 계수의 추정치가 영향을 받지 않음이 알려져 있다. 따라서 설명변수들이 서로 독립적이라고 가정할 경우 <Table 3>과 <Table 4>의 분석에서 얻은 계수 값을 그대로 사용하여 예측식을 설정할 수 있다. 가령 <Table 3>과 <Table 4>의 분석에서

$$\begin{aligned} y_{i,t} &= a_1 + b_1x_{j_1,t-l_1} + e_{i,1,t} \\ y_{i,t} &= a_2 + b_2x_{j_2,t-l_2} + e_{i,2,t} \\ y_{i,t} &= a_3 + b_3x_{j_3,t-l_3} + e_{i,3,t} \end{aligned} \quad (15)$$

을 얻었다면 $y_{i,t}$ 에 대한 예측식은 다음과 같다:

$$y_{i,t} = b_1x_{j_1,t-l_1} + b_2x_{j_2,t-l_2} + b_3x_{j_3,t-l_3} + e_{i,t} \quad (16)$$

위 식에서 상수항은 제외시켰는데 각 변수가 1계 차분을 통해서 얻어진 점을 고려하면 상수항의 값은 0에 가까울 것으로 볼 수 있기 때문이다.

3.2 이산변수의 경우

<Table 7>은 이산화된 상태·영향·반응 요인 각각에 대한 예측식을 보여준다. 각 설명변수의 계수는 식 (13)에서 정의한 τ_i 의 추정치이다. 예측식을 설명하기 위해 $y_{i,t}$ 와 $x_{j_1,t-l_1}$, $x_{j_2,t-l_2}$ 간 음의 상관관계가 있는 경우를 고려해 보자. 조건부분포 $y_{i,t}|x_{j_1,t-l_1}$ 와 $x_{j_1,t-l_1}$ 가 독립이라고 가정하면

$$\Pr(y_{i,t} = 1 | x_{j_1,t-l_1}, x_{j_2,t-l_2}) \approx \frac{1}{2} [\Pr(y_{i,t} = 1 | x_{j_1,t-l_1}) + \Pr(y_{i,t} = 1 | x_{j_2,t-l_2})] \quad (17)$$

Table 7. Dynamic Model of State, Impact, and Response Indicators: Discretization

$GBOD(t) = (1/2) * [0.05*ATS(t-6) + 0.04*WWI(t-46)]$
$GDO(t) = (1/2) * [0.03*ATS(t-2) + 0.05*WWI(t-13)]$
$GTP(t) = (1/2) * [0.05*ATS(t-5) + 0.04*WWI(t-8)]$
$GTN(t) = (1/2) * [0.05*ATS(t-5) + 0.03*WWI(t-15)]$
$GCOD(t) = (1/2) * [0.04*ATS(t-12) + 0.03*WWI(t-21)]$
$CDO(t) = 0.02*WWI(t-48)$
$CTN(t) = 0.04*WWI(t-36)$
$CCOD(t) = 0.05*WWI(t-24)$
$SFP(t) = (1/2) * [0.04*ATS(t-21) + 0.05*WWI(t-20)]$
$SWP(t) = (1/2) * [0.04*ATS(t-40) + 0.04*WWI(t-12)]$
$ATS(t) = (1/5) * [-0.03*GBOD(t-16) - 0.04*GDO(t-9)$ $- 0.02*GTP(t-25) - 0.04*GTN(t-30) - 0.02*SWP(t-44)]$
$WWI(t) = (1/7) * [-0.03*GBOD(t-9) - 0.05*GDO(t-12)$ $- 0.03*GTP(t-25) - 0.04*GTN(t-42) - 0.02*CTN(t-36)$ $- 0.04*SFP(t-20) - 0.03*SWP(t-23)]$

$Pr(y_{i,t} = 1 | x_{j,t-t_1} = -1) = Pr(y_{i,t} = -1 | x_{j,t-t_1} = 1)$ 임을 가정하면,
 $Pr(y_{i,t} = 1 | x_{j_1,t-t_1}, x_{j_2,t-t_2}) = \tau_{j_1,t_1} \cdot x_{j_1,t-t_1}, Pr(y_{i,t} = 1 | x_{j_2,t-t_2}) = \tau_{j_2,t_2} \cdot x_{j_2,t-t_2}$
 이므로,

$$Pr(y_{i,t} = 1 | x_{j_1,t-t_1}, x_{j_2,t-t_2}) \approx \frac{1}{2} [\tau_{j_1,t_1} \cdot x_{j_1,t-t_1} + \tau_{j_2,t_2} \cdot x_{j_2,t-t_2}] \quad (18)$$

설명변수가 여러 개이고, 설명변수와 피설명변수 간의 관계가 다른 방향인 경우로 위 식을 확장하면 <Table 7>에 제시된 식을 얻게 된다.

4. 결 론

본고에서는 금강하구역 DPSIR 지표체계에 포함된 변수들 간의 동태적 관계를 추정하고, DPSIR 프레임워크에서 가정하고 있는 관계가 실제하는지에 대한 가설검정을 시행하였다. 가설검정 결과 대부분의 변수에 대해 인과관계가 존재하지 않는다는 귀무가설을 기각할 수 있었다. 짧은 시계열로 인해 발생하는 문제를 완화시키기 위해 추정해야 하는 파라미터의 수를 최소화하였고, 연속변수를 이산화(discretization)시키는 방법도 적용해 보았다. 또한 추정결과를 바탕으로 DPSIR변수들 간의 동태관계를 간결하게 표현해 주는 모형을 제시하였다.

이 모형을 이용하면 기존의 시나리오 분석을 개선시킬 수 있을 것으로 기대된다. Nobre[2009] 등이 시도한 시나리오 분석에서는 가치평가모형 등 추가적 분석을 병행하는 것이 필요하였지만, 본고에서 제시한 모형을 이용하면 DPSIR 프레임워크만을 이용한 시나리오 분석이 가능하다. 물론 제한된 데이터를 바탕으로 진행된 분석의 한계점은 언급할 필요가 있다. 최대한 단순화된 모형을 제시하려 시도하였으나 자료에서 확인할 수 없는 여러가지 가정을 필요한 점도 한계점이다. 이 같은 한계점에도 불구하고 가치평가모형 등의 추정이 불가능한 경우에도 시나리오 분석을 진행할 수 있다는 점에서 유용성이 있을 것으로 평가한다.

후 기

본 논문은 2018년 해양수산부 재원으로 한국해양과학기술진흥원의 지원을 받아 수행된 연구임(하구역 종합관리시스템 개발 연구).

References

- [1] ABP Marine Environmental Research, 2008, Understanding and Managing Morphological Change in Estuaries.
- [2] Anatolyev, S., and Gospodinov, N., 2010, Modeling financial return dynamics via decomposition, Journal of Business and Economic Statistics, 28(2), 232-245.
- [3] Faraggi, D., and Simon, R., 1996, A simulation study of cross-validation for selecting an optimal cutpoint in univariate survival analysis, Statistics in Medicine, 15, 2203-2213.
- [4] Gang, D.S., Nam, J.H., and Jeong, Y.H., 2005, Development and application of sustainable development indicators for Gyeonggiman Hangang estuary, Fishing and Marine Education Research, 17(2), 155-170.
- [5] Gang, H.C., and No, S.H., 2016, Incentives to promote green stores, Environmental Policy, 24(1), 163-189.
- [6] Gari, S.R., Newton, A., and Icely, J.D., 2015, A review of the application and evolution of the DPSIR framework with an emphasis on coastal social-ecological systems, Ocean and Coastal Management, 103, 63-77.
- [7] Irwin, J.R., and McClelland, G.H., 2003, Negative consequences of dichotomizing continuous predictor variables, Journal of Marketing Research, 40, 366-371.
- [8] Korea Environmental Institute, 2014, Development and Application of Sustainable Water Usage Indicators.
- [9] Korea Marine Institute, 2007, Development of Coastal Survey Indicators.
- [10] Kuss, O., 2013, The danger of dichotomizing continuous variables: A visualization, Teaching Statistics, 35(2), 78-79.
- [11] Leung, M.T., Daouk, H., and Chen, A., 2000, Forecasting stock indices: a comparison of classification and level estimation models, International Journal of Forecasting, 16, 173, 190.
- [12] Lewison, et al., 2016, How the DPSIR framework can be used for structuring problems and facilitating empirical research in coastal systems, Environmental Science and Policy, 56, 110-119.
- [13] Mariooryad, S., and Busso, C., 2017, The cost of dichotomizing continuous labels for binary classification problems: deriving a bayesian-optimal classifier, IEEE Transactions on Affective Computing, 8(1), 119-130.
- [14] Morgan, T.M., and Elashoff, R.M., 1986, Effect of categorizing a continuous covariate on the comparison of survival time, Journal of American Statistical Association, 81(396), 917-921.
- [15] Naggara, O., Raymond, J., Guilbert, F., Roy, D., Weill, A., and Altman, D.G., 2011, Analysis by categorizing or dichotomizing

- continuous variables is inadvisable: An example from the natural history of unruptured aneurysms, *AJNR American Journal of Neuroradiology*, 32, 437-440.
- [16] Nobre, A.M., 2009, An ecological and economic assessment methodology for coastal ecosystem management, *Environmental Management*, 44, 185-204.
- [17] Nyberg, H., 2011, Forecasting the direction of the US stock market with dynamic binary probit models, *International Journal of Forecasting*, 27, 561-578.
- [18] Owen, S.V., and Froman, R.D., 2005, Why carve up your continuous data?, *Research in Nursing and Health*, 28, 496-503.
- [19] Perry, R.I., and Masson, D., 2013, An Integrated analysis of the marine social-ecological system of the Strait of Georgia, Canada, over the past four decades, and development of a regime shift index, *Progress in Oceanography*, 115, 14-27.
- [20] Phillips, P.C.B., and Perron, P., 1988, Testing for a unit root in time series regression, *Biometrika*, 75(2), 335-346
- [21] Vlaams-Nederlandse Schelde Commissie, 2011, Indicators of the Long-term Vision 2030 (LTV) for the Scheldt Estuary: Reader's Guide.
- [22] Wu, Y., and Zhang, H., 1997, Forward premiums as unbiased predictors of future currency depreciation: a non-parametric analysis, *Journal of International Money and Finance*, 16(4), 609-623.

Received 28 June 2018

Revised 9 August 2018

Accepted 13 August 2018