Validation of Stable Isotope Analysis in R (SIAR) Model for Source Appointment of Agricultural Water Pollution

Hyun-Jin Park¹, Sang-Sun Lim^{2*}, and Woo-Jung Choi^{3*}

¹Lecture and Postdoctoral Research Fellow, Department of Rural and Bio-systems Engineering & Education and Research Unit for Climate-Smart Reclaimed-Tideland Agriculture, Chonnam National University, Gwangju 61186, Korea

²Senior Researcher, Bio R&D Center, CJ Cheiljedang, Suwon, Gyeonggi-do 16495, Korea

³Professor, Department of Rural and Bio-systems Engineering & AgriBio Institute of Climate Change Management, Chonnam National University, Gwangju 61186, Korea

*Corresponding author: Woo-Jung Choi, wjchoi@chonnam.ac.kr

*Co-corresponding author: Sang-Sun Lim, ss.lim1@cj.net

ABSTRACT

Received: October 22, 2020

Revised: October 23, 2020

Accepted: October 25, 2020

ORCID

Hyun-Jin Park

<http://orcid.org/0000-0002-2642-5723>

Sang-Sun Lim

<https://orcid.org/0000-0002-1875-9707>

Woo-Jung Choi

<http://orcid.org/0000-0002-2009-8207>

The 'Stable Isotope Analysis in R (SIAR)' is useful in estimating the contribution of non-point pollution source to water pollution in intensive land-use areas with divergent pollution sources. However, the accuracy of the SIAR model has not been validated. In this study, the accuracy of the SIAR model was validated using simulation data set, and the applicability of the SIAR model was tested using literature data set. For source appointment of water pollution using SIAR, chemical variables including stable carbon ($\delta^{13}\text{C}$) and nitrogen ($\delta^{15}\text{N}$) isotope ratios as well as stoichiometry of C, N, and phosphorus (P) such as C/N, N/P, and C/P were used as end-members, which are the representative values of water chemistry variables for each potential source. In the simulation study, it was found that at least (n-1) end-members are required for accurate ($r^2 > 0.90$ for regression between given and estimated contribution of pollution sources) estimation of pollution source when the number of sources is n. However, it was also suggested that using correct tracer that represents the source characteristics is critical for the accurate estimation of pollution source. Therefore, it is recommended to use end-members > (n-1) for more accurate estimation of pollution sources. In the study of SIAR model application to a case study with 5 sources and 2 or 3 end-members (C/N, C/P, and $\delta^{13}\text{C}$ for total C; C/N, N/P, and $\delta^{15}\text{N}$ for total N; and C/P and N/P for total P), it was possible to appoint source contribution for total C, N, and P in water. However, the accuracy of the estimation is expected to be low due to less number of end-member used in SIAR than the recommendation (i.e., n-1). Therefore, it is highly necessary to develop end-members to represent the sources of total C, N, and P for improvement of the accuracy of source appointment using SIAR model.

Keywords: SIAR, Accuracy, Non-point sources, Isotope mixing model, Water pollution

Diagrams of model output of SIAR: (a) bi-plot, (b) histogram, (c) matrix plot, and (d) box plot.

Introduction

기후변화에 수반되는 이상 기상 등에 의해 토지가 집약적으로 이용되고 있는 우리나라에서 다양한 오염원에 기인한 수질 오염 문제가 더욱 심각해지고 있다 (Jeon, 2015; Kang and Lee, 2019). 특히, 농경지, 축산, 산림, 주거가 복합적으로 구성되어 있는 농촌지역은 오염원이 다양하고 오염물질의 유출 경로가 복잡하여 수질 오염 관리가 더욱 어렵다 (Sun et al., 2012; Jeon, 2015; Nam et al., 2019). 따라서, 농촌지역과 같은 복합토지이용 구역의 수질 관리를 위해서는 수질 오염원을 평가할 수 있는 기술 개발이 요구되는데, 특히 각 오염원 별 특성을 반영하는 추적자 (Tracer)를 활용하여 수질 오염원의 기여도를 평가하는 기술이 주목받고 있다 (Murphy and Morrison, 2014).

다양한 추적자 중 안정성동위원소를 활용하는 연구가 여러 국가에서 진행되어왔는데, 우리나라에서도 질소(N) 동위원소비 ($^{15}\text{N}/^{14}\text{N}$, $\delta^{15}\text{N}$)를 이용하여 다양한 토지이용형태별로 지표수와 지하수의 오염원을 추정하는 바 있다(Yoo et al., 1999; Hong et al., 2001; Lee et al., 2003; Choi et al., 2007, 2011). 하지만, $\delta^{15}\text{N}$ 만을 오염 추적자로 활용할 경우에 오염원 별 $\delta^{15}\text{N}$ 값이 중첩하는 문제와 세 가지 이상의 오염원 구명은 불가능한 문제가 있어, 농촌지역과 같은 복합토지이용 구역의 오염원을 구명하는데 적용하기 어렵다(Choi et al., 2017). 이러한 한계점을 극복하여 농촌지역과 같은 복합토지이용 구역의 수질 오염원 규명을 위해서는 $\delta^{15}\text{N}$ 과 탄소 동위원소비 ($^{13}\text{C}/^{12}\text{C}$, $\delta^{13}\text{C}$), 그리고 탄소와 질소를 포함한 다종의 원소 함량 등 다양한 수질 화학 지표를 추적자로 사용할 필요가 있다(Jeon, 2015). 수질 오염원 인자로 작용하는 다양한 오염원들은 물질의 구성 및 생성, 순환 과정의 차이에 의해 동위원소 및 원소 함량이 상이할 수 있기 때문에 오염원과 수질 시료의 동위원소 및 원소 함량을 비교 분석함으로써 복합토지이용 구역의 수질 오염원 추적에 활용할 수 있다(Choi et al., 2017; Park et al., 2020; Park et al., 2021). Jeon (2015)은 농촌지역 구역의 오염원 별 동위원소비 및 원소 함량비를 Tree model에 적용하여 하천수의 오염원 별 기여도를 평가한 바 있다. 오염원 분류를 위한 Tree model에서는 여러 오염원을 대상으로 다양한 화학 추적자를 분석하여, 추적자별 분포 범위가 유사한 오염원을 순차적으로 두 가지 집단으로 구분하고, 각 집단의 추적자 평균값을 두 가지 집단을 구분하는 end-member (각 오염원 또는 오염원 집단별 추적자의 대표 값)로 사용하여 순차적으로 오염원 집단의 기여도를 평가하여 최종적으로 다양한 개별 오염원의 기여도를 평가한다(Spruill et al., 2002). 하지만, Tree model은 두 집단 간의 상대적 오염 기여도를 단계별로 평가하기 때문에, 동일 집단내에 포함된 오염원 별 기여도를 직접적으로 평가할 수 없고 통계적 신뢰도 확보가 어렵다는 한계가 있다(Spruill et al., 2002).

이에 반해, 동위원소 혼합 모델(Isotope mixing model)을 이용하여 다중 오염원의 기여도를 동시에 평가 가능한 R 기반의 SIAR (Stable Isotope Analysis in R) 모델은 산출된 잠재 요인의 기여도에 대한 통계적 신뢰구간 확보가 가능하여 오염원 평가 결과의 신뢰도가 높은 장점이 있다(Parnell et al., 2013; Wang et al., 2019). SIAR 모델은 애초 동위원소를 이용하여 먹이 그물의 역학을 구명하기 위하여 개발되었지만, 최근 여러 연구에서 복합토지이용 구역의 지하수나 지표수의 질소오염원을 구명하거나(Xue et al., 2012; Matiatos, 2016; Yu et al., 2020; Yuan et al., 2020) 구역의 유기물 오염원을 구명하는 연구(Lu et al., 2016) 등에 활용된 바가 있어 비점오염원 구명 연구에도 적용이 가능한 것으로 확인되었다. SIAR 모델은 동위원소비뿐만 아니라 원소 함량비를 함께 입력 변수로 활용하면 잠재적 오염원의 기여도를 보다 더 정확하게 정량화할 수 있도록 개발되었기 때문에(Parnell, 2008), 비점오염원의 원소 함량비도 추적자로 활용할 수 있다. 예를 들면, Lu et al. (2016)은 구역의 유기물 오염원을 구명하기 위해 C/N을 이용하였으며, Garzon-Garcia et al. (2017)은 저수지 침전물의 오염원 구명을 위해 탄소 및 질소 함량(총탄소와 총유기탄소, 총질소 함량)등을 추적자로 이용하였다.

이와 같이 SIAR 모델은 비점오염원의 동위원소비와 원소 함량비를 이용하여 오염원을 추적하고 기여도를 평가하는데 활용될 수 있지만, 현재까지 진행된 연구에서는 SIAR 모형의 결과를 그대로 사용하여 복합토지이용 구역의 오염원을 추적하였고, SIAR 모형의 정확도가 이론적으로 검증된 바가 없다. 따라서, 비점오염원 기여도 평가 연구에서 SIAR 모형의 활용도 확대를 위해서는 다양한 시나리오별로 SIAR 모형의 정확도에 대한 이론적 검증이 필요하다. 본 연구는 복합토지이용 구역의 비점오염원 기여도 평가를 위한 SIAR 모형의 정확도를 검증하고 적합한 SIAR 모형 구동 방안을 제시하고자 하였다. 또한, 실제 농촌구역에서 조사한 결과를 활용하여 SIAR 모형을 구동함으로써 복합토지이용 구역의 비점오염원 기여도 평가를 위한 SIAR 모델의 적용 가능성을 평가하였다.

Materials and Methods

SIAR 모형 SIAR 모형은 2008 년 Bayesian 이론을 기반으로 하여 생태환경 연구 분야에서 물질 Sink (예, 수질)에 대한 Source (오염원)의 잠재적인 기여도를 평가하기 위해 개발되었다 (Parnell, 2008). 모형을 통해 정량화된 잠재 요인의 기여도는 다음과 같이 표현된다 (Parnell et al., 2010).

$$X_{ij} = \sum_{k=1}^K P_k (S_{jk} + C_{jk}) + \varepsilon_{ij}$$

$$S_{jk} \sim N(\mu_{jk}, \omega_{jk}^2)$$

$$C_{jk} \sim N(\lambda_{jk}, \tau_{jk}^2)$$

$$\varepsilon_{jk} \sim N(0, \sigma_j^2)$$

여기서, X_{ij} 는 Sink i ($=1, 2, 3, \dots, I$)의 j ($=1, 2, 3, \dots, J$) 동위원소비 또는 원소 함량비이다. S_{jk} 와 C_{jk} 는 각각 Source k ($=1, 2, 3, \dots, K$)의 j 동위원소비 또는 원소 함량비와 동위원소 분할 (fractionation) 지수이며, 이들은 일반적으로 각각 평균 μ_{jk} 와 표준편차 ω_{jk}^2 , 평균 λ_{jk} 와 표준편차 τ_{jk}^2 의 정규분포를 따른다. P_k 는 SIAR 모형에서 추정하는 오염원 k 의 기여도이다. ε_{ij} 는 오차이며, 이는 평균 0과 표준편차 σ_j^2 정규분포를 따른다. 따라서, SIAR 모델링을 위해 입력되는 변수는 총 3 종류인데, 첫째, 조사 지역의 수질 시료 (Sink)에 대한 동위원소비 또는 원소 함량비와 같은 추적자의 평균, 둘째, 오염원 추적자의 end-member의 평균과 표준편차, 셋째, 동위원소분할지수의 평균값과 표준편차로 구성된다. 이 때, 동위원소분할 작용을 고려하지 않을 경우, 분할 지수를 “0”으로 기입하여도 모형은 구동할 수 있다 (Parnell, 2008).

Table 1. List of R command for SIAR modeling.

Contents	Command in R [†]
Install SIAR	Install.package (“siar”)
Start SIAR package	library (siar)
Specify folder path	setwd (“folder path”)
Input mixture data	<i>sink</i> ← read.table (‘file’s name.txt’,header=TRUE)
Input source data	<i>source</i> ← read.table (‘file’s name.txt’,header=TRUE)
Input fractionation factor data	<i>fractionation</i> ← read.table (‘file’s name.txt’,header=TRUE)
Running SIAR model	<i>model</i> ← siarmcmcdirichletv4 (<i>mixture</i> , <i>source</i> , <i>fractionation</i> , concdep=0,500000, 50000)
Plotting biplot	siarplotdata (<i>model</i>)
Plotting histogram	siarhistograms (<i>model</i>)
Plotting matrix plot	siarmatrixplot (<i>model</i>)
Plotting boxplot	siarproportionbygroupplot (<i>model</i>)
Quantifying results	<i>sink</i> (“result.txt”) out ← <i>model</i> \$output out sink ()

[†]The italicized words (*sink*, *source*, *fractionation*, *model*, and *result*) are arbitrarily specified, and thus subject to change.

SIAR 모델의 기반인 R 프로그램은 무료 공용프로그램으로 <http://www.r-project.org/>에서 설치 가능하다. R 프로그램은 스크립트를 통해 명령문을 이용하여 모델을 구동하거나 결과를 도식화 하기 때문에 명령문에 대한 학습이 필요하며, 학습을 위한 자료는 프로그램 설치 홈페이지에서 얻을 수 있다.

R 프로그램에서 SIAR 패키지 설치 및 데이터 입력, 모델 구동 등에 필요한 명령어는 Table 1에 정리하였다. SIAR 패키지 외에도 엑셀 패키지를 설치하여 엑셀에 입력된 데이터를 SIAR 패키지의 입력 값으로 전환할 수도 있지만 (Nam et al., 2019), 본 연구에서는 엑셀 파일을 텍스트 형식의 파일로 변환하여 SIAR 패키지에 입력하였다. R 프로그램을 이용하여 SIAR 모형의 분석결과를 행렬도 (bi-plot), 히스토그램, 산점도 (matrix plot), 상자그림 (box plot)으로 도식화할 수 있으며 (Fig. 1), 수치화하여 텍스트 형식의 파일로 저장할 수도 있다. 저장한 텍스트 형식의 파일은 파일 형식을 변환하여 엑셀로 데이터를 추출 및 가공할 수 있다.

Fig. 1. Examples of diagrams of model output of SIAR: (a) bi-plot, (b) histogram, (c) matrix plot, and (d) box plot.

가상 데이터를 이용한 SIAR 모형의 정확도 검증 모형의 정확도 평가를 위해서 논, 밭, 과수원, 생활하수, 산림 등 5 개의 가상의 오염원을 설정하고 $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, C/N, N/P, Na/Ca 를 추적자로 사용하였다 (Table 2a). 오염원과 추적자의 종류는 SIAR 모형의 실제 유역 적용에 활용된 Jeon (2015)의 자료에 근거하여 선정하였다. Jeon (2015)에 의하면 논-밭-과수원-산림은 탄소와 질소의 공급원과 각각의 투입량과 손실량 차이에 의해 $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, C/N, 그리고 N/P 가 상이 하였으며, 생활하수는 특이적으로 Na/Ca 이 높았다. SIAR 모형은 두 가지 시나리오로 검증하였는데, 시나리오 1에서는 모형에 활용되는 추적자 end-member 의 종류와 개수에 따른 정확도를 검증하였으며, 시나리오 2에서는 시나리오 1의 결과를 확장하여 오염원 개수가 상이한 조건에서 각 오염원 별 기여도가 달라졌을 때의 정확도를 검증하였다.

Table 2. Simulation conditions of SIAR model: (a) mean values and standard deviation ($\mu \pm \sigma^2$) of end-members of five tracers for each source, (b) $\mu \pm \sigma^2$ of tracers for three sink (water samples) of scenario 1, and (c) $\mu \pm \sigma^2$ of tracers for three sink (water samples) of scenario 2.

(a) End-member of tracers of five sources

Source	End-members of tracers				
	$\delta^{13}\text{C}$ (‰)	$\delta^{15}\text{N}$ (‰)	C/N	N/P	Na/Ca
Paddy	-27.70 (0.21)	4.80 (0.06)	13.63 (0.09)	2.27 (0.12)	0.12 (0.00)
Upland	-25.17 (0.50)	9.43 (0.27)	13.47 (0.23)	1.13 (0.12)	0.12 (0.00)
Orchard	-23.83 (0.52)	10.37 (0.15)	12.13 (0.15)	1.47 (0.23)	0.09 (0.00)
Residence	-26.07 (0.12)	6.30 (0.12)	0.80 (0.06)	6.93 (0.15)	3.21 (0.00)
Forest	-28.27 (0.15)	-2.40 (0.21)	1.80 (0.17)	49.60 (0.42)	1.84 (0.00)

(b) Scenario 1

Sink	End-members of tracers				
	$\delta^{13}\text{C}$ (‰)	$\delta^{15}\text{N}$ (‰)	C/N	N/P	Na/Ca
1	-26.17 (0.05)	5.97 (0.18)	6.28 (0.28)	9.76 (0.74)	1.69 (0.07)
2	-26.22 (0.03)	6.13 (0.05)	11.05 (0.16)	9.62 (0.43)	0.46 (0.03)
3	-26.66 (0.06)	3.85 (0.17)	7.87 (0.09)	21.93 (0.78)	0.97 (0.01)

(c) Scenario 2

Sink	End-members of tracers				
	$\delta^{13}\text{C}$ (‰)	$\delta^{15}\text{N}$ (‰)	C/N	N/P	Na/Ca
1	-24.14 (0.01)	10.05 (0.04)	12.35 (0.01)	1.46 (0.01)	0.09 (0.00)
2	-24.90 (0.05)	9.12 (0.08)	12.79 (0.01)	1.51 (0.01)	0.10 (0.00)
3	-25.44 (0.04)	8.35 (0.04)	13.01 (0.03)	1.61 (0.00)	0.11 (0.00)
4	-24.55 (0.01)	9.82 (0.02)	12.81 (0.00)	1.32 (0.00)	0.11 (0.00)
5	-23.92 (0.02)	10.26 (0.02)	12.18 (0.01)	1.48 (0.00)	0.09 (0.00)
6	-25.80 (0.03)	6.97 (0.06)	3.57 (0.21)	5.67 (0.10)	2.49 (0.05)
7	-25.11 (0.01)	8.58 (0.02)	9.06 (0.04)	3.11 (0.01)	1.08 (0.01)
8	-25.22 (0.02)	8.10 (0.04)	7.70 (0.19)	3.77 (0.08)	1.38 (0.04)
9	-25.31 (0.03)	8.02 (0.05)	7.47 (0.09)	3.88 (0.03)	1.47 (0.02)
10	-25.65 (0.05)	7.76 (0.09)	9.90 (0.22)	2.98 (0.11)	0.90 (0.06)
11	-27.71 (0.03)	-0.15 (0.13)	3.63 (0.19)	39.05 (0.59)	1.63 (0.04)
12	-26.96 (0.06)	2.77 (0.16)	3.11 (0.09)	24.76 (0.54)	2.14 (0.03)
13	-26.09 (0.02)	5.06 (0.07)	5.90 (0.13)	17.10 (0.30)	1.54 (0.03)
14	-25.86 (0.03)	5.16 (0.06)	2.84 (0.10)	17.49 (0.25)	1.55 (0.02)
15	-26.20 (0.03)	3.18 (0.06)	1.93 (0.07)	15.59 (0.41)	0.76 (0.01)

Values are the means with standard deviation in parentheses.

Details of scenarios 1 and 2 are described in the text.

시나리오 1에서는 오염원을 5 개로 고정하고 각 오염원의 기여도가 상이한 3 개의 유역 (Sink 1–3) (Table 2b)을 대상으로 추적자 end-member 의 종류와 수를 2–5 개로 달리하여 모두 8 개의 end-member 경우의 수 조합으로 구성하였는데, 조합에는 2 가지 동위원소비 ($\delta^{13}\text{C}$ 와 $\delta^{15}\text{N}$)를 기본으로 하고 나머지 C/N, N/P, Na/Ca 를 계승적 (factorial)으로 포함시켰다. 사용한 end-member 조합은 end-member 2 개 ($\delta^{13}\text{C}$ - $\delta^{15}\text{N}$), 3 개 ($\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -C/N, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -N/P, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -Na/Ca), 4 개 ($\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -C/N-N/P, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -C/N-Na/Ca, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -N/P-Na/Ca), 5 개 ($\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -C/N-N/P-Na/Ca)였다. 시나리오 2에서는 추적자 end-member 를 5 개로 고정하고, 오염원의 종류가 각각 3, 4, 5 개인 3 개 유역에서 각 오염원의 오염 기여도가 5 가지로 변하는 15 개의 세부시나리오 조건 (Table 2c)으로 모형의 정확도를 검증하였다. 일반적으로 SIAR 에서 요구하는 추적자 end-member 개수는 (n-1) 개이지만 (Parnell, 2008), 시나리오 2에서는 추적자 end-member 의 종류와 개수의 영향이 배제된 조건에서 개별 오염원의 기여도가 달라질 때 SIAR 모형의 정확도를 검증하였다.

오염원 (Source)과 각 수질 시료 (Sink)의 end-member 를 SIAR 모형에 입력하여 각 오염원의 기여도를 산출하였으며, end-member 에 대한 분할 지수는 0 으로 하여 분할 작용은 고려하지 않았다. SIAR 모형의 정확도 검증을 위해 가상 시나리오에서 설정한 각 오염원의 기여도 값 (설정치)과 SIAR 모델을 통해 산출된 오염원의 기여도 값 (예측치)을 회귀 분석을 통해 비교하였으며, 결정 계수 (r^2)가 0.90 이상일 때 모형의 신뢰도가 높다고 판단하였다.

실제 유역 분석 자료를 이용한 SIAR 모형의 적용성 검토 이상과 같은 시나리오에서는 각 오염원 별로 수질 시료의 특정 오염물질 (예를 들면, 총탄소, 총질소, 총인)에 대한 기여도 보다는 잠재적 전체 오염원의 오염 기여도를 평가할 수 있다. 하지만, 농촌지역 수질 관리를 위해서는 잠재적 오염원 기여도 보다는 특정 오염원에 대한 각 오염원 기여도 평가가 필요하다. 이를 위해, 영농-축산-주거-산림이 복합적으로 구성된 전남 장성군 풍영정천 유역에서 수행된 Jeon (2015)의 연구 결과를 활용하여 복합토지이용 유역 수질의 총탄소, 총질소, 총인에 대한 오염원 별 기여도 평가에서 SIAR 모형의 적용성을 검토하였다. 이 유역에는 5 가지의 오염원 (논, 밭, 과수원, 생활하수, 산림)이 있는 것으로 조사되었다. Jeon (2015)의 결과 중 SIAR 모형 분석을 위해 필요한 유역의 수질 시료 및 오염원의 end-member 데

Table 3. Data extracted from Jeon (2015) for application of SIAR model: (a) end-members of tracers for sources and (b) end-members of tracers in sinks.

Source-Sink	C/N	N/P	C/P	$\delta^{13}\text{C}$ (‰)	$\delta^{15}\text{N}$ (‰)
(a) Source					
Paddy	10.2 (1.4)	6.9 (2.7)	69.9 (1.1)	-26.9 (0.8)	1.8 (0.7)
Upland	2.4 (1.0)	7.6 (16.7)	18.1 (1.7)	-24.3 (1.5)	1.9 (0.9)
Orchard	10.1 (7.5)	1.5 (18.2)	14.8 (0.8)	-23.4 (1.0)	3.0 (1.6)
Residence	0.8 (0.1)	6.8 (0.4)	5.6 (1.6)	-26.3 (0.1)	6.5 (0.1)
Forest	1.6 (0.4)	40.3 (9.7)	62.5 (0.3)	-28.5 (0.5)	-2.5 (1.0)
(b) Sink [†]					
1	5.8 (2.7)	22.1 (4.3)	118.0 (42.9)	-22.4 (3.5)	11.0 (1.9)
2	2.1 (0.5)	34.2 (5.7)	65.1 (14.9)	-26.5 (0.8)	14.5 (2.0)
3	2.1 (0.6)	21.5 (8.0)	31.9 (7.3)	-26.9 (1.8)	10.3 (1.2)

Values are the means with standard deviation in parentheses.

[†]The concentrations of total C of sinks 1, 2, and 3 were 11.9 (2.5), 6.7 (1.3), and 5.3 (1.6) mg C L⁻¹, respectively; those of total N were 2.5 (0.6), 3.5 (0.5), and 3.1 (0.8) mg N L⁻¹, respectively; those of total P were 0.12 (0.03), 0.11 (0.02), and 0.16 (0.02) mg P L⁻¹, respectively. Data set of sinks 1, 2, and 3 were used for SIAR model for the estimation of sources contribution for C, N, and P, respectively.

이터는 Jeon (2015)의 결과를 활용하였다 (Table 3). Jeon (2015)은 이 유역으로 유입되는 오염물질의 탄소와 질소, 인 (P) 함량 및 탄소와 질소 동위원소비를 조사하였으며, 약 5 개월간 유역의 수질 시료를 정기적으로 채취하여 동일한 항목을 조사하였다.

본 연구에서는 Jeon (2015)의 연구결과 중 대표적인 수질오염물질 측정 항목인 탄소, 질소, 인 함량이 상대적으로 높은 하천 수질 시료에 대해 SIAR 모델을 적용하여 총탄소, 총질소, 총인에 대한 5 개 오염원의 기여도를 평가하였다 (Table 3). 총탄소, 총질소, 총인 오염원 기여도 평가를 위해 원소 함량비 (C/N, N/P, C/P)와 동위원소비 ($\delta^{13}\text{C}$, $\delta^{15}\text{N}$)를 추적자로 추출하여 각각의 end-member 를 사용하였으며, end-member 에 대한 분할 지수는 0 으로 가정하였다. 총탄소, 총질소, 총인 오염 기여도 평가에는 각 오염물질과 관련도가 높은 C/N-C/P- $\delta^{13}\text{C}$ (총탄소), C/N-N/P- $\delta^{15}\text{N}$ (총질소)와 N/P-C/P (총인)를 추적자로 선정하여 각 추적자의 평균값을 end-member 로 사용하였다. Sink 1, 2, 3 수질 시료의 총탄소 농도는 각각 11.9 ± 2.5 , 6.7 ± 1.3 , $5.3 \pm 1.6 \text{ mg C L}^{-1}$, 총질소 농도는 각각 2.5 ± 0.6 , 3.5 ± 0.5 , $3.1 \pm 0.8 \text{ mg N L}^{-1}$, 총인 농도는 각각 0.12 ± 0.03 , 0.11 ± 0.02 , $0.16 \pm 0.02 \text{ mg P L}^{-1}$ 였다.

Results and Discussion

시나리오 1: 추적자 종류와 개수에 따른 SIAR 모형의 정확도 시나리오 1 에서 오염원을 5 개로 고정하고 각 오염원의 기여도가 상이한 3 개의 유역 (Sink 1–3) (Table 2b)을 대상으로 추적자의 종류와 수를 2–5 개로 달리하여 SIAR 모형의 정확도를 검토한 결과, end-member 개수가 2 개일 경우 (Combination A), 설정치와 예측치의 차이가

Table 4. Contribution of sources to sinks estimated by SIAR model under the Scenario 1 with fixed contribution of sources and varying combinations of tracers.

Source	Given contribution (%)	Estimated contribution (%) estimated by using different combinations (A–H) of tracers [†]							
		A	B	C	D	E	F	G	H
(a) Sink 1									
Paddy	17	20 (+3)	13 (-4)	23 (+6)	14 (-3)	16 (-1)	12 (-5)	17 (0)	14 (-3)
Upland	12	20 (+8)	17 (+5)	18 (+6)	16 (+4)	15 (+3)	17 (+5)	15 (+3)	17 (+5)
Orchard	15	20 (+5)	17 (+2)	20 (+5)	15 (0)	14 (0)	15 (0)	14 (0)	13 (-2)
Residence	44	22 (-22)	37 (-7)	24 (-20)	40 (-4)	41 (-3)	41 (-3)	41 (-4)	43 (-1)
Forest	12	17 (+5)	16 (+4)	15 (+3)	15 (+3)	13 (+1)	15 (+3)	13 (+1)	13 (+1)
(b) Sink 2									
Paddy	33	24 (-9)	35 (+2)	24 (-9)	34 (+1)	33 (0)	33 (0)	32 (-1)	34 (+1)
Upland	25	21 (-4)	24 (-1)	22 (-3)	22 (-3)	24 (-2)	27 (+2)	26 (0)	23 (-2)
Orchard	23	18 (-5)	21 (-3)	17 (-6)	24 (+1)	23 (0)	21 (-2)	23 (0)	25 (+2)
Residence	2	24 (+22)	7 (+5)	23 (+21)	4 (+2)	5 (+3)	3 (+1)	3 (+1)	3 (+1)
Forest	16	14 (-2)	14 (-3)	14 (-2)	15 (-1)	16 (-1)	15 (-1)	16 (0)	16 (0)
(c) Sink 3									
Paddy	13	16 (+3)	20 (+7)	11 (-2)	21 (+8)	15 (+2)	23 (+10)	13 (0)	16 (+3)
Upland	28	15 (-13)	20 (-8)	17 (-11)	12 (-16)	23 (-5)	8 (-20)	24 (-4)	21 (-7)
Orchard	13	13 (0)	14 (+1)	15 (+2)	22 (+9)	16 (+3)	24 (+11)	16 (+3)	16 (+3)
Residence	4	21 (+17)	10 (+6)	18 (+14)	7 (+3)	5 (+1)	8 (+4)	5 (+1)	6 (+2)
Forest	42	35 (-7)	36 (-6)	39 (-3)	38 (-4)	41 (-1)	37 (-5)	42 (0)	40 (-2)

Values in parentheses are the numerical difference between estimated and given contribution.

[†] A, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$; B, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -C/N; C, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -N/P; D, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -Na/Ca; E, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -C/N-N/P;

F, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -C/N-Na/Ca; G, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -N/P-Na/Ca; H, $\delta^{13}\text{C}$ - $\delta^{15}\text{N}$ -C/N-N/P-Na/Ca.

-22 ~ +22%로 오염원 기여도 평가의 정확도가 낮았다 (Table 4). 이는 오염원 기여도 설정치와 예측치의 회귀 분석 결과에서 r^2 이 0.00 - 0.56 으로 낮은 것으로도 확인되었다 (Fig. 2a). 반면, 3 개의 end-member (Combination B-D)를 활용하면 오염원의 기여도의 정확도가 증가하였다. 비록 Combination C 의 경우, 설정치-예측치 회귀식의 r^2 이 0.65 (Sink 3)이하로 낮았지만, Combination B 와 D 에서는 대부분의 r^2 이 0.90 이상으로 높아 예측치의 정확도가 향상되었다 (Fig. 2b-d). 또한, 4 개의 end-member (Combination E-G)를 사용하면 설정치-예측치 회귀식의 r^2 이 일부를 제외하고 0.92-0.99 로 높았다 (Fig. 2e-g). 하지만, 3 개의 end-member 를 활용한 Combination D 와 4 개의

Fig. 2. Correlation between the given and SIAR-estimated contributions of sources to a sink under the Scenario 1 with fixed contribution of sources and varying combinations of tracers (Table 2): (a) end-members $\delta^{13}\text{C}-\delta^{15}\text{N}$, (b) end-members $\delta^{13}\text{C}-\delta^{15}\text{N}-\text{C}/\text{N}$, (c) end-members $\delta^{13}\text{C}-\delta^{15}\text{N}-\text{N}/\text{P}$, (d) end-members $\delta^{13}\text{C}-\delta^{15}\text{N}-\text{Na}/\text{Ca}$, (e) end-members $\delta^{13}\text{C}-\delta^{15}\text{N}-\text{C}/\text{N}-\text{N}/\text{P}$, (f) end-members $\delta^{13}\text{C}-\delta^{15}\text{N}-\text{C}/\text{N}-\text{Na}/\text{Ca}$, (g) end-members $\delta^{13}\text{C}-\delta^{15}\text{N}-\text{N}/\text{P}-\text{Na}/\text{Ca}$, and (h) end-members $\delta^{13}\text{C}-\delta^{15}\text{N}-\text{C}/\text{N}-\text{N}/\text{P}-\text{Na}/\text{Ca}$. Data depicted in the figure are provided in Table 4.

end-member 를 활용한 Combination F 에서 Sink 3 에 대한 설정치-예측치 회귀식의 r^2 이 각각 0.55 와 0.34 로 낮은 점은 SIAR 모형의 정확도가 end-member 개수에 결정되지 않음을 보여준다. 일반적으로, 오염원 개수가 n 개일 때 SIAR 모형 구동을 위해서 $n-1$ 개의 end-member 가 필요하지만 (Parnell, 2008), 본 연구 결과는 end-member 의 개수는 물론 해당 end-member 의 오염원 대표성 역시 중요함을 제시한다. 예를 들면, 시나리오 1 에서 Sink 3 의 경우 상대적으로 산림의 기여도가 큰 데에 반해 (Table 4), Combination D 와 F 에는 산림유출수의 특성을 보여 줄 수 있는 N/P 가 end-member 로 포함되지 않았다 (Table 4). 일반적으로 인산질 비료가 사용 되지 않으면서, 대기로부터 질소 강하물이나 질소 고정과 같은 경로로 질소가 유입되는 산림은 농경지 등의 타 오염원에 비해서 N/P 가 높은 것으로 알려져 있다 (Jeon et al., 2015; Shen et al., 2019; Choi et al., 2020).

한편, 5 개의 오염원에 대해 5 개의 end-member 를 활용한 Combination H 에서는 설정치-예측치 회귀식의 r^2 이 0.94-0.98 로 매우 높았다 (Fig. 2h). 따라서, 이상의 결과를 종합하면, 추적자의 오염원 대표성이 높을 경우에는 n 개의 오염원에 대해 ($n-1$)개의 end-member 로 신뢰도 높은 오염원 기여도 평가가 가능하지만, 추적자에 대한 신뢰도가 낮을 경우에는 가급적 다수의 end-member 를 사용하면 오염원 평가의 정확도가 향상되는 것을 알 수 있다 (Parnell, 2008).

시나리오 2: 오염원 개수가 상이한 조건에서 각 오염원의 기여도 변화에 따른 SIAR 모형의 정확도 시나리오 2 에서 오염원의 종류가 3-5 개인 조건에서 추적자를 5 개로 고정하고, 오염원의 오염 기여도가 5 가지로 변하는 조건에서 SIAR 의 정확도를 평가한 결과, SIAR 의 오염원 평가 정확도는 오염원 개수와 end-member 개수의 조합은 물

Table 5. Contribution of sources to sinks estimated by SIAR model under the Scenario 2 with varying number of sources (3-5) and their contribution and fixed 5 tracers.

Source	Source contribution type									
	Type 1		Type 2		Type 3		Type 4		Type 5	
	Given	Estimated	Given	Estimated	Given	Estimated	Given	Estimated	Given	Estimated
(a) Sink with 3 sources (Sinks 1-5)										
Paddy	4	3 (-1)	17	18 (+1)	31	31 (0)	2	4 (+2)	2	2 (+1)
Upland	12	27 (+15)	30	27 (-3)	31	31 (0)	49	39 (-10)	1	9 (+7)
Orchard	84	70 (-14)	53	55 (+2)	38	38 (0)	50	58 (+8)	97	89 (-8)
(b) Sink with 4 sources (Sinks 6-10)										
Paddy	3	3 (0)	4	3 (-1)	9	7 (-2)	7	6 (-1)	24	24 (0)
Upland	9	9 (0)	31	37 (+5)	9	18 (+9)	16	22 (+6)	24	24 (0)
Orchard	11	11 (0)	33	28 (-5)	41	33 (-8)	33	27 (-5)	26	26 (0)
Residence	77	77 (0)	31	32 (+1)	41	42 (+1)	44	44 (0)	26	26 (0)
(c) Sink with 5 sources (Sinks 11-15)										
Paddy	6	5 (-1)	6	4 (-2)	6	4 (-2)	3	6 (+3)	20	18 (-2)
Upland	5	6 (+1)	4	5 (+1)	5	12 (+7)	23	10 (-13)	21	21 (0)
Orchard	5	5 (0)	6	6 (0)	30	24 (-5)	26	26 (0)	20	27 (+7)
Residence	6	6 (0)	41	41 (0)	30	30 (0)	24	30 (+6)	20	7 (-13)
Forest	77	77 (0)	44	44 (0)	29	29 (0)	24	29 (+5)	20	27 (+8)

Values in parentheses are the numerical difference between estimated and given contribution.

Fig. 3. Correlation between the given and SIAR-estimated contributions of sources to a sink under the Scenario 2 with varying number of sources (3–5) and their contribution and fixed 5 tracers (Table 2): (a) sink with 3 sources, (b) sink with 4 sources, and (c) sink with 5 sources. Data depicted in the figure are provided in Table 4.

론, 오염원별 기여도 변화에 의해서도 달라지는 것으로 나타났다 (Table 5). 예를 들면, 오염원이 각각 3 개와 4 개인 조건에서 오염원별 기여도 설정치가 각각 30%내외와 20% 내외로 유사한 경우 (각각 Type 3 과 5), SIAR 의 예측치는 설정치와 정확하게 일치하였다. 하지만, 오염원이 5 개인 조건에서 오염원별 기여도 설정치가 유사한 Type 5 의 예측치는 설정치와 -13 ~ +8 의 차이가 있었고 (Table 5), 설정치와 예측치 회귀분석의 r^2 이 0.09 로 매우 낮았다 (Fig. 3). 또한, 오염원이 5 개인 조건에서 Type 4 의 r^2 도 0.54 로 매우 낮았다. 한편, 오염원 보다 더 많은 end-members 를 활용한 조건에서 r^2 이 0.90 이상으로 높았다. 이러한 결과는 오염원 개수 보다 많은 수의 end-member 를 활용하면 오염원의 기여도가 다양하게 변하더라도 SIAR 모형 예측 결과가 상대적으로 정확함을 보여준다 (Parnell, 2008). 따라서, SIAR 을 활용한 오염원 기여도 평가의 정확도 확보를 위해서는 가급적 많은 end-member 를 활용할 필요가 있다.

실제 유역을 대상으로 한 SIAR 적용 Jeon (2015)의 실측 자료를 활용하여 SIAR 모형을 적용한 결과에 의하면, 오염원 별 기여도는 총탄소의 경우 과수원 (22%) > 산림 (20%) > 논 (21%) > 밭 (19%) > 주거 (18%)의 순서였고, 총 질소는 산림 (37%) > 주거 (28%) > 밭 (21%) > 논=과수원 (7%), 총인은 산림 (23%) > 과수원=밭=주거 (20%) > 논 (17%)의 순서였다 (Fig. 4). 실제 오염원 별 기여도에 대한 신뢰도가 확보된 결과가 없기 때문에 본 연구의 오염원 평가 결과의 정확도를 평가하기는 어렵다. 하지만, 본 연구에서는 전체 오염원이 5 개였고, end-member 는 총탄소

Fig. 4. Contribution of source to sink (water samples) with different water chemistry: (a) total C of sink 1, (b) total N of sink 2, and (c) total P of sink 3. Details of water chemistry are provided in Table 3.

(C/N, C/P, $\delta^{13}\text{C}$)와 총질소 (C/N, N/P, $\delta^{15}\text{N}$)는 3 개, 총인 (N/P, C/P)은 2 개를 사용하여, SIAR 모형의 기본 end-member (n-1)인 4 개보다 적었기 때문에 오염원 기여도 평가의 정확도는 낮을 것으로 판단된다 (Parnell, 2008). 하지만, SIAR 의 사례 연구 적용을 통해 SIAR 의 활용 가능성을 확인하였다는 측면에서 연구 결과의 의미를 찾을 수 있다. 앞의 시나리오별 SIAR 모형 결과에 의하면 총탄소, 총질소, 총인의 오염원 별 특징적 추적자와 해당 end-member 를 추가하면 오염원 별 기여도 평가 결과의 정확도를 높일 수 있을 것으로 기대된다.

Conclusions

SIAR 모형은 지금까지 전세계에서 수질 오염원을 추적하기 위한 연구에 활용되어 왔지만, 복합토지이용 구역의 수질 오염원을 추적하기 위한 SIAR 모형의 정확성이 검증된 바 없어 산출된 오염원의 기여도에 대한 신뢰도를 향상시킬 필요가 있다. 본 연구에서 시나리오별로 SIAR 모형의 오염원 기여도 평가 정확도를 검증한 결과, n 개의 오염원에 대한 정확한 오염 기여도 평가를 위해서는 최소 n-1 개의 end-member 가 필요하지만, end-member 로 사용된 추적자가 오염원의 특성을 제대로 반영하지 못할 때는 오염 기여도의 정확도가 낮아지는 것으로 나타났다. 따라서, 추적자의 오염원 별 대표성이 불확실할 경우에는 오염원 개수와 동일하거나 많은 end-member 를 사용하면 이와 같은 문제를 극복할 수 있는 것으로 나타났다. SIAR 의 사례연구 적용 결과에 의하면, 총탄소, 총질소, 총인 오염에 대한 오염원 별 기여도 결과의 정확도를 평가할 수는 없었지만, SIAR 을 이용한 오염원 기여도 평가 가능성은 확인할 수 있었다. 향후 주요 오염물질 별 오염원의 특징을 대표할 수 있는 추가적인 추적자와 해당 end-member 를 보완하면 보다 정확한 오염원 평가가 가능할 것으로 기대된다.

Acknowledgement

This work was carried out with the support of “Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ015010022020)”, Rural Development Administration, Republic of Korea.

References

- Choi, W.J., G.H. Han, S.M. Lee, G.T. Lee, K.S. Yoon, S.M. Choi, and H.M. Ro. 2007. Impact of land-use types on nitrate concentration and $\delta^{15}\text{N}$ in unconfined groundwater in rural areas of Korea. *Agric. Ecosyst. Environ.* 120:259-268.
- Choi, W.J., J.H. Kwak, H.J. Park, H.I. Yang, S.I. Park, Z.H. Xu, S.M. Lee, S.S. Lim, and S.X. Chang. 2020. Land-use type, and land management and disturbance affect soil ^{15}N : a review. *J. Soils Sediments* 20:3283-3299.
- Choi, W.J., J.H. Kwak, S.S. Lim, H.J. Park, S.X. Chang, S.M. Lee, M.A. Arshad, S.I. Yun, and H.Y. Kim. 2017. Synthetic fertilizer and livestock manure differently affect $\delta^{15}\text{N}$ in the agricultural landscape: A review. *Agric. Ecosyst. Environ.* 237:1-15.
- Choi, W.J., J.W. Jung, W.J. Choi, K.S. Yoon, D.H. Choi, S.S. Lim, J.H. Jeong, B.J. Lim, and N.I. Chang. 2011. Estimation of pollution sources of Oenam watershed in Juam lake using nitrogen concentration and isotope analysis. *J. Korean Soc. Water Qual.* 27:467-474.
- Garzon-Garcia, A., J.P. Laceby, J.M. Olley, and S.E. Bun. 2017. Differentiating the sources of fine sediment, organic matter and nitrogen in a subtropical Australian catchment. *Sci. Tot. Environ.* 575:1384-1394.
- Hong, Y.J. 2001. Identification of the sources of nitrate using stable isotope mass ratio in rural watersheds. Master's thesis. Seoul National University, Seoul, Korea.
- Jeon, B.J. 2015. Source identification of rural non-point source pollution using isotopic and elemental characteristics of pollutants. Master's thesis. Chonnam National University, Gwangju, Korea.
- Kang, M.W. and S.S. Lee. 2019. Assessment and its control of non-point source pollution in Korea: Review. *J. Korean Soc. Water Wastewater* 33:457-467.
- Lee, Y.D., S.G. Hyun, and H.K. Song. 2003. Estimation of sources for nitrate in groundwater at Ongpo stream watershed in Jeju island. *J. Korean Soc. Water Qual.* 19:455-461.
- Lu, L., H. Cheng, X. Pu, J. Wang, Q. Cheng, and X. Liu. 2016. Identifying organic matter sources using isotope ratios in a watershed impacted by intensive agricultural activities in Northeast China. *Agric. Ecosyst. Environ.* 222:48-59.
- Matiatos, J. 2016. Nitrate source identification in groundwater of multiple land-use areas by combining isotopes and multivariate statistical analysis: A case study of Asopos basin (Central Greece). *Sci. Tot. Environ.* 541:802-814.
- Murphy, B.L. and R.D. Morrison. 2014. Introduction to environmental forensics. Academic Press.
- Nam, T.H., H.S. Ryu, T.W. Kang, Y.U. Han, J.H. Kim, K.H. Lee, S.H. Hwang, and K.H. Kim. 2019. Quantifying nitrogen source contribution ratios using stable isotope method: Application of Bayesian mixing model. *J. Korean Soc. Water Environ.* 35:510-519.
- Park, S.I., H.I. Yang, H.J. Park, B.S. Seo, Y.J. Jeong, and W.J. Choi. 2020. Changed stoichiometric ratios of carbon, nitrogen, and phosphorus in runoff water from maize upland fields by rice straw cover. *Korean J. Soil Sci. Fertil.* 53:258-267.
- Park, S.I., H.I. Yang, H.J. Park, B.S. Seo, Y.J. Jeong, S.S. Lim, J.H. Kwak, H.Y. Kim, K.S. Yoon, S.M. Lee, and W.J. Choi. 2021. Rice straw cover decreases soil erosion and sediment-bound C, N, and P losses but increases dissolved organic C export from upland maize fields as evidenced by $\delta^{13}\text{C}$. *Sci. Tot. Environ.* 753:142053.
- Parnell, A.C. 2008. SIAR: stable isotope analysis in R. R package version 3.
- Parnell, A.C., D.L. Philips, S. Bearhop, B.X. Semmens, E.J. Ward, J.W. Moore, A.L. Jackson, J. Grey, D.J. Kelly, and R. Inger. 2013. Bayesian stable isotope mixing models. *Environmetrics* 24:387-399.
- Parnell, A.C., R. Inger, S. Bearshop, and A.L. Jackson. 2010. Source partitioning using stable isotopes: coping with too much variation. *PLoS one* 5:e9672.

- Shen, F., J. Wu, H. Fan, W. Liu, X. Guo, H. Duan, L. Hu, X. Lei, and X. Wei. 2019. Soil C/P and C/P ratio regulate the responses of soil microbial community composition and enzyme activities in a long-term nitrogen loaded Chinese fir forest. *Plant Soil* 436:91-107.
- Spruill, T.B., W.J. Showers, and S.S. Howe. 2002. Application of classification-tree methods to identify nitrate sources in ground water. *J. Environ. Qual.* 31:1538-1549.
- Sun, B., L. Zhang, L. Yang, F. Zhang, D. Norse, and Z. Zhu. 2012. Agricultural non-point source pollution in China: causes and mitigation measures. *AMBIO* 41:370-379.
- Wang, J., N. Lu, and B. Fu. 2019. Inter-comparison of stable isotope mixing models for determining plant water source partitioning. *Sci. Total Environ.* 666:685-693.
- Xue, D., B. De Baets, O. Van Cleemput, C. Hennessy, M. Berglund, and P. Boeckx. 2012. Use of a Bayesian isotope mixing model to estimate proportional contributions of multiple nitrate sources in surface water. *Environ. Pollut.* 161:43-49.
- Yoo, S.H., W.J. Choi, and G.H. Han. 1999. An investigation of the sources of nitrate contamination in the Kyonggi province groundwater by isotope ratios analysis of nitrogen. *Korean J. Soil Sci. Fertil.* 32:47-56.
- Yu, L., T. Zheng, X. Zheng, Y. Hao, and R. Yuan. 2020. Nitrate source apportionment in groundwater using Bayesian isotope mixing model based on nitrogen isotope fractionation. *Sci. Tot. Environ.* 718:137242.
- Yuan, R., T. Zheng, X. Zheng, D. Liu, J. Xin, L. Yu, and G. Liu. 2020. Identification of groundwater nitrate pollution sources in agricultural area using PCA and SIAR methods. *Episodes* 43:739-749.