

식품용수 수질자료를 이용한 지하수 오염 예측 모델 개발 및 소규모 유역에서의 검증

남성우¹ · 박은규^{2*} · 이명재¹ · 전선금¹ · 정혜민¹ · 김정우¹

¹(주)지오그린21

²경북대학교 지구시스템과학부

Development of Prediction Model of Groundwater Pollution based on Food Available Water and Validation in Small Watersheds

Sungwoo Nam¹ · Eungyu Park^{2*} · Myeong-jae Yi¹ · Seonkeum Jeon¹ · Hyemin Jung¹ · Jeongwoo Kim¹

¹Geogreen21 Co, Ltd, Seoul, Korea

²Department of Geology, Kyungpook National University, Daegu, Korea

ABSTRACT

Groundwater is used in many areas in food industry such as food manufacturing, food processing, cooking, and liquor industry etc. in Korea. As groundwater occupies a large portion of food industry, it is necessary to predict deterioration of water quality to ensure the safety of food water since using undrinkable groundwater has a ripple effect that can cause great harm or anxiety to food users. In this study, spatiotemporal data aggregation method was used in order to obtain spatially representative data, which enable prediction of groundwater quality change in a small watershed. In addition, a highly reliable predictive model was developed to estimate long-term changes in groundwater quality by applying a non-parametric segmented regression technique. Two pilot watersheds were selected where a large number of companies use groundwater for food water, and the appropriateness of the model was assessed by comparing the model-produced values with those obtained by actual measurements. The result of this study can contribute to establishing a customized food water management system utilizing big data that respond quickly, accurately, and preemptively to changes in groundwater quality and pollution. It is also expected to contribute to the improvement of food safety management.

Key words : food processing water, groundwater, prediction model, segmented regression, data aggregation

1. 서 론

식품용수란 식품의 제조, 가공, 조리, 소분 등에 사용하는 물을 말하며, 우리나라에서는 지하수가 식품용수로 식품의 제조, 가공, 조리 등 많은 부분에 활용되고 있다. 이러한 경향은 지하수를 식품용수로 사용하는 업소의 증가로 나타나고 있으며, 식품 산업의 발달과 깨끗한 물에 관

한 관심 증가로 업소와 식품용수로 사용하는 지하수의 이용량은 지속해서 증가할 것으로 예상된다.

식품용수로 사용되는 지하수는 오염될 경우 해당 식품의 불특정 이용자에게 동시다발적으로 위해를 미칠 수 있는 과급성을 지니고 있으며, 식품업소의 영업 형태 등에 따라 지역에만 국한되지 않는다는 점에서 심각성을 내포하고 있다. 이에 현행 「식품위생법」에서는 수도물이 아닌 지하수 등을 먹는 물 또는 식품의 가공은 물론 조리·세척 등에 사용하는 경우 먹는물 수질검시기관에서 검사를 받아 마시기에 적합하다고 인정된 물을 사용하도록 규정하고 있다. 수질검사의 주기는 같은 법 시행규칙에 따라 1년마다 실시하고, 음료류 등 마시는 용도의 식품이면 6개월마다 수질검사를 수행하도록 의무화한 안전장치를 마련해두고 있다.

주저자: 남성우, (주)지오그린21, 전무

공저자: 이명재, (주)지오그린21, 대표이사; 전선금, (주)지오그린21, 이사; 정혜민, (주)지오그린21, 사원; 김정우, (주)지오그린21, 부사장

*교신저자: 박은규, 경북대학교 지구시스템과학부, 교수
Email: egpark@knu.ac.kr

Received : 2021.12.07 Reviewed : 2021.12.09

Accepted : 2021.12.20 Discussion until : 2022.02.28

그러나 영업 형태에 따라서는 먹는물 기준의 검사가 2년 주기로 이루어지기도 하며, 검사 결과에서 용수로서의 적합성 여부만을 확인하는 경우가 대부분을 차지하고 있었다. 일부 영업소의 경우 「식품위생법」 제23조(식품 등의 재검사) 조항을 잘못 해석하여 부적합 항목에 대한 재검사를 시행하는 등 식품용수 지하수질 관리의 사각지대가 일부 확인되기도 하였다.

한 개의 영업소에서 식품용수로 사용되는 지하수에 대한 수질검사를 수행하면 검사 용도별로 회 당 12~46개 항목의 수질검사 결과가 생성되는데 이를 전국으로 확대해보면 생성되는 수질 자료의 개수는 연간 약 1,500개(2019년 식품업소 기준)로 추산된다. 또한, 시료 당 약 27만 원에 달하는 먹는물 수질검사 수수료를 고려하면 데이터 생성에 드는 비용이 막대하다는 것을 알 수 있다. 생성되는 데이터의 양과 그 비용을 고려할 때, 지하수질의 적합성 여부만 확인하는 등의 단편적인 관리에서 벗어나 사장되는 수질 자료의 활용성을 증대시킬 방안이 필요하다. 식품용수로 사용되는 지하수의 오염으로 인한 위해성과 과급력을 고려할 때, 지하수질 관리의 사각지대를 보완하여 부적합이 우려되는 지하수가 식품의 제조, 가공, 조리 등에 사용되는 것을 차단하고, 그동안 많은 비용을 들여 생성된 식품용수 지하수질 데이터를 효과적으로 활용하기 위하여, 데이터 기반의 지하수질 오염예측 모델의 개발이 필수적이며, 예측된 결과를 식품업체에 제공하여 선제적으로 식품용수 안전관리를 할 수 있는 방법론의 개발이 요구된다.

유럽연합에서는 식품으로 이용되는 지하수를 별도로 분리하여 관리하지는 않지만, 이미 지하수 수질 추세 평가를 통해 지하수 수질을 주기적으로 감시하고 있으며, 수질 항목의 농도가 지속적으로 증가하는 추세를 보일 경우 문턱값(수질기준 농도의 75%)에 도달하기 전 추세를 역전시키기 위한 조치를 시행하여 수질악화를 방지하는 노력을 기울이고 있다(EC, 2009). 한편, 미국에서는 농업지역 관정의 약 20%에서 질산성질소 농도가 미국 환경보호국(Environmental Protection Agency, EPA)이 정한 음용수 최대오염수준인 10mg/L를 초과함에 따라 질산염 우선지역(Nitrate Priority Areas, NPA)을 선정하고 각 주에서 질산성질소 오염관리 계획을 수립 및 이행하기 위해 일반 전략과 세부 목표를 세워 관리하고 있다(The Groundwater Association, 2020). 2014년 NPA 선정 시에는 Nonpaired data analyses, Paired data analyses, Ratio의 3가지 방법을 사용하였다. 선정된 Bruneau 지역의 세부 목표 및 전략으로는 질산성질소 오염에 대한 개인 관정의 수질검

사 활성화, 질산성질소 오염원에 대한 교육 기반의 지하수 수질 개선 및 관리, 인센티브 제도를 통한 가축사육시설로 인한 오염 방지 등이 있다(Korea Environment Institute, Jeju Province, 2020). 또한, 영국에서도 4개 그룹의 모니터링을 통해 질산염의 농도 현황과 추세를 분석하여 지하수 질산염 취약지구를 선정, 지도화한 사례가 있다. 해당 연구에서는 모니터링 결과 예외적으로 높거나 낮은 농도의 이상치를 제거하기 위해 Multiple Outlier Test를 실시하고, 그룹별로 데이터 수가 충분한 모니터링 지점은 실측자료를 기반으로 총무기질소 농도의 95% 분위값을 산정하여 미래의 수질 모형기반으로 총무기질소 농도를 추정하는 방법을 사용하였다. 예측된 결과는 통계적 보간 기법인 크리깅을 이용하여 영국 전역을 1km²의 해상도로 지도화 하여 모니터링되지 않는 지역의 질산염 농도를 예측할 수 있는 성과를 도출하였다(An, 2020).

현재 우리나라는 「지하수법」 제18조와 「수질보전등에 관한 규칙」 제9조에 의거하여 국가지하수 수질측정망을 통한 지하수 수질현황과 수질변화 추세를 정기적으로 파악하고 있으나, 「지하수법」에 의한 지하수 관리는 관리범위가 광범위하고 오염우려지역에 국한되어 있어 개별 시설 위주인 식품용수로 사용되는 지하수의 관리에까지 그 실효성이 미치지 못하고 있는 실정이다. 다만, 수질의 예측이 필요하다는 추세에 맞춰 국내에서도 지하수를 관리해야 할 필요성이 대두되고 있고, 여러 가지 방안도 제시되고 있다.

한국수자원공사·제주특별자치도(2016)는 제주도에서 문제 되고 있는 질산성질소 항목에 대해 지하수 관정의 질산성질소 농도 자료 22,169건 중 지점별 평균값을 이용하여 ordinary Kriging을 통해 제주도 전 역의 질산성질소 농도 분포를 예측하고, 배경수질농도(1.0 mg/L), 보전수질농도(배경농도 상한치 2.0 mg/L)와 먹는물 수질기준치인 10.0 mg/L를 기본으로 AAA, AA, A, B 4개 등급으로 분류도를 나타내었다. 나아가 Mann-Kendall 시험 및 Sen 기울기 시험을 시행하여 질산성질소 농도 추세의 크기를 설명하였다. 한국환경정책평가연구원·제주특별자치도(2020)는 한국수자원공사·제주특별자치도(2016)의 연구를 바탕으로 지하수 수질 등급을 먹는물수질기준(10 mg/L)과 생활용수 수질기준(20 mg/L)에 따라 1, 2, 3등급인 A, B, C 등급으로 나누어 관리구역을 지정하여 관리하는 방안과 제도를 추가로 제안하였다. A 등급은 먹는물 수질기준 미만의 청정상태로 수처리 없이 음용수로 사용 가능할 수 있는 상태를 말하며, B 등급은 먹는물 수질기준을 초과하지만, 지하수가 생활용수로 활용되고 농도의 증가 추세 역

전만 필요한 경우이고 C등급은 생활용수수질기준 미만으로 지하수 수질 저하가 더 이상 지속되지 않는 상태로 관리하는 지역을 의미한다.

본 연구에서는 자료의 한계성으로 인하여 질산성질소 항목에 국한된 분석을 시행하였다. 질산성질소는 대표적인 인위적 오염물질로, 지표 또는 지표 인근의 지하 오염 물질이 지하로 침투하여 지하수를 오염시키는 양상을 잘 설명한다. 본 연구는 수질검사(질산성질소) 자료 기반 지역 지하수 오염추세 평가기법을 개발하고, 이를 식품용수 안전성 분야에 적용하였다. 이들 수질검사 자료는 식품용수의 적합 여부 판정 후 예기되던 자료로, 자료의 신뢰성이 크지는 않지만, 전국을 포괄한다는데 의미가 크다.

본 연구에서는, 10년간의 식품용수(지하수) 수질검사 자료를 확보하여 GIS 기반의 DB로 구축하였으며, 향후 빅데이터로 발돋움하기 위한 표준 절차를 마련하였다. 또한, 식품용수로 사용되는 지하수의 장기적인 변화를 예측할 수 있도록 예측모델을 개발하고 적용하여 사전적 개념의 식품용수 관리를 위한 기반을 마련하였다.

2. 연구방법

2.1. 수질 검사 자료 확보

식품용수로 사용되는 지하수는 「식품위생법」 제44조(영업자 등의 준수사항)에 따라 수도물이 아닌 지하수 등을 먹는 물 또는 식품의 제조·가공 등에 사용하는 경우이므로 「먹는물관리법」 제43조에 의거하여 먹는 물 수질 검사기관에서 「식품위생법 시행규칙」 제57조(식품접객영업자 등의 준수사항)로 정하는 바에 따라 검사를 받아 마시기에 적합하다고 인정된 물을 사용하여야 한다. 해당 규정에 따라 식품업소가 수질검사기관에 수질검사를 의뢰해 수질검사 자료를 지자체에 제출하면 지자체에서는 새

올행정시스템을 이용하여 수질검사 결과를 관리한다. 「식품위생법」에 근거한 지하수 수질검사 자료의 경우, 지자체에서 담당하고 있는 모든 식품업소의 현황이 기재되어 있으며, 영업소별로 인허가번호, 급수시설(상수도 또는 지하수), 업소 상태, 업종, 업소명, 주소 등의 기본 정보와 함께, 수질검사 일자, 해당 검사 연도의 검사 항목 수(전 항목 또는 일부항목), 수질검사에 관련된 검사 결과 등 연도별로 최근 5년간의 수질검사 이력을 보유하고 있다. 다만, 검사 결과는 적합 또는 부적합 여부로만 관리되기 때문에 개별 수질 항목별 농도는 수질검사 기관에서 발행한 수질검사 성적서를 확인해야 하는 번거로움이 있고, 이 과정에서 수질검사 성적서에는 영업소 명이나 주소 등의 정보가 누락되는 경우가 일반적이라 매칭시키는 과정이 발생한다.

따라서 본 연구에서는 새올행정시스템 상의 먹는물 수질 검사자료(식품업소)와 개별 지하수 이용 영업소 정보를 확보하여 정제 및 표준화 과정을 거쳐 사용하였다.

2.2. 수질 예측 모델 개발

본 연구에서는 정확도 및 신뢰도가 높은 예측 모델을 개발하기 위해 3단계로 구성되는 최적의 프로세스를 설정하였다. 1단계는 전처리과정(pre-processing)으로 데이터 전처리 과정이며, 2단계는 주·처리(main-process)과정인 모델추정으로 수질 예측을 위한 추세분석 과정, 3단계는 후처리(post-process) 과정으로 추세 해석과 등급 분류 과정이다.

2.2.1. 전처리과정

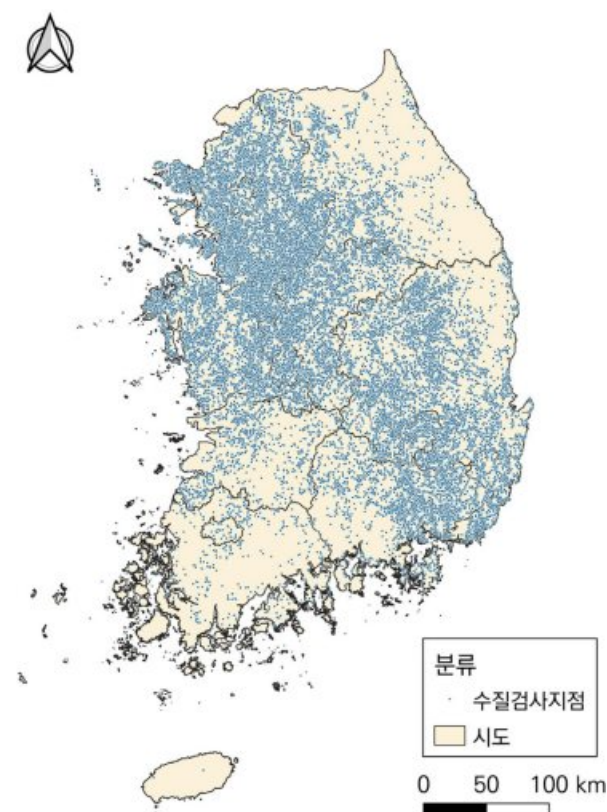
전처리과정은 수집한 지하수 사용 식품업소의 수질검사 자료를 추세분석에 적합하도록 정제·통합·결합 등 가공하는 모든 과정을 의미한다. 전처리과정을 거친 데이터를 토



Fig. 1. The process of developing a model for predicting the water quality trend of groundwater.

Table 1. The number of water quality test data used by underground water use businesses

구분	전체 자료 수	미활용 자료			최종 활용 자료 수
		주소 미기재	주소 부정확	유역 외	
자료 수(개)	67,975	7,428	709	84	59,754
비율(%)	100	10.9	1.0	0.1	87.9

**Fig. 2.** Distribution of water quality (groundwater) test points in food businesses.

대로 주처리과정에서 수질 예측 결과를 얻기 때문에 전처리과정은 예측에 타당한 면적의 범위와 자료의 수 등을 결정하는 중요한 절차이다.

전처리과정에서는 먼저, 전산화되어 있지 않은 수질검사 자료의 DB화를 위해 구주소 및 도로명주소를 WGS84 좌표로 구성된 SHP 파일로 변환해주는 지오코딩 툴(Geocoding Tool(v20.10.15))을 사용하여 위치 좌표를 생성하였다. 수집한 67,975개의 수질검사 자료 중 87.9%인 59,754개 자료에서 위치 좌표가 생성되었고, 위치 좌표가 생성되지 않은 12.1%는 수질검사 성적서에 주소가 기재되어 있지 않거나, 정확하지 않은 주소가 기재된 경우, 섬 지역 등 표준유역에서 벗어난 경우 등이다. 이와 같은 자료들은 정형 및 비정형 자료를 포함하고 있으므로, 빅데이터적 성격을 지닌다.

앞서 언급하였듯이, 수질검사 기준을 충족하기 위해 반복적으로 재검사가 진행된 사례가 있어, 이로 인한 중복 데이터를 최대한 정제하여 최초 검사자료를 활용하고자 재분석된 결과를 선별하는 과정을 거쳤다. 수질검사 항목 중 수소이온농도 값의 부적합으로 재검사하는 사례는 없을 것으로 가정하여 중복 자료 선별을 위해 수소이온농도 항목을 이용하였으며, 수소이온농도 값에 미검사 등으로 표기된 해당 검사일의 검사자료는 자료 전체를 분석에서 제외하였다.

위치 좌표가 생성된 59,754개의 자료 중 수소이온농도 값을 이용하여 재분석으로 추정되는 11,876개를 제외한 결과, 추세분석에 활용할 수 있는 자료 수는 총 47,878개이다.

다음으로 예측 및 관리 대상 면적을 결정하였다. 본 연구는 개별 관정이 아니라 일정 면적 단위에서의 추세분석을 통해 예측 및 관리가 필요하고, 지하수의 흐름 특성을 고려하면 광역적 규모의 유역보다는 준 유역 규모가 효과적이므로(Kim, 2020), 이를 만족하는 적합한 면적을 결정하기 위해 대권역, 중권역, 표준유역이 제공되는 환경부의 수자원단위지도와 집수구역이 제공되는 KRF 하천망분석도를 모두 검토하였다.

추세분석에 적합한 면적 검토 시, (1) 유역의 면적 및 개수가 관리에 적합하여야 하고, (2) 추세분석이 가능한 자료의 수가 많아야 하며, (3) 추세분석 결과의 정확도와 신뢰도가 높아야 한다는 기준을 두었다.

대권역, 중권역, 표준유역, 집수구역으로 구분된 유역 수는 각각 21개, 118개, 851개, 8,118개이며, 추세분석이 가능한 유역은 각각 15개(71.4%), 23개(19.5%), 76개(8.9%), 61개(0.8%)로 나타났다.

대권역과 중권역의 경우, 추세분석 가능 비율이 높지만, 두 권역 모두 관리에 적합한 면적 범위를 벗어나며, 정확도와 신뢰도가 높은 추세를 나타내지는 못한다. 집수구역의 경우 평균 면적이 10 km²로 적어 관리가 용이하나, 면적 분할에 대한 정보가 정확하지 않아 신뢰도가 낮고, 해당 면적에 포함되는 기본적인 자료의 수를 충족하지 못해 추세분석이 가능한 지역이 매우 적다.

표준유역의 경우, 지하수관리 및 상위 계획 등에서도 가

Table 2. Comparison of watersheds suitable for trend analysis

구분	대권역	중권역	표준유역	집수구역
코드	BBSNCD	MBSNCD	SBSNCD	CATID
유역 수	21개	118개	851개 (826개)	8,118개
평균면적	약 3,000 km ²	약 700 km ²	약 100 km ²	약 10 km ²
추세분석 가능 유역 수	15개	23개	76개	61개
추세분석 유역 비율	71.4%	19.5%	8.9%	0.8%

장 기본적으로 널리 사용하는 신뢰성 있는 유역단위이며, 평균 면적이 100 km²로 관리에도 용이하여 본 연구에서는 표준유역을 최소 단위로 사용하였다.

표준유역 각각에는 고유의 SBSNCD 코드가 부여되어 있어, 추세분석이 가능하도록 정제된 후의 위치 좌표를 가지는 47,878개의 자료를 각 유역과 결합하였다. 이때 검사 수 미달로 분석 결과를 신뢰할 수 없는 유역을 최소화하여 신뢰도와 정확도가 높은 추세분석 결과를 도출하기 위해 하나의 표준유역 당 최소 분석 자료 수를 50개 이상으로 설정하였다. 표준유역과 수질 자료를 결합한 결과, 추세분석이 가능한 표준유역은 308개이며, 표준유역 내 검사 결과는 모두 해당 표준유역을 대표한다고 가정하였다.

식품용수로 사용되는 지하수의 경우, 검사일 기준으로 1년마다 「먹는 물 수질검사 및 검사 등에 관한 규칙」 제 4조에 따른 마을상수도의 검사기준에 따른 일부항목 검사(12항목)를 하고, 「먹는 물 수질검사 및 검사 등에 관한 규칙」 제 2조에 따른 먹는 물의 수질기준에 따라 2년마다 모든 항목에 대한 검사(46항목)를 진행한다.

최대 46개의 수질검사 항목 중 추세분석 대상 항목을 선정하기 위해 추세를 나타낼 수 있는 오염 항목 중 질산성질소, 황산이온, 염소이온, 암모니아성질소 등을 검토하였다. 이 중 질산성질소와 암모니아성질소의 값이 가장 많은 것으로 확인되었으나, 암모니아성질소는 표준유역당 검사 수가 충분하지 않아 본 연구에서는 분석할 수 없어 제외하고, 검사 수가 충분히 축적된 질산성질소 항목을 대상 항목으로 선정하였다.

질산성질소는 분석값이 충분할 뿐 아니라, 세균류 다음으로 검사 횟수가 가장 많으며, 전국적인 분포를 보인다는 장점을 가지고 있다. 「먹는물관리법」 먹는물 수질기준 및 검사 등에 관한 규칙 제 4조에 따라 질산성질소는 주요 오염 관리 대상 항목으로 마을상수도 및 전용상수도에서 매 분기 1회 이상 검사해야 하는 등 대표적인 지표 오염 지시인자임에도 식품 먹는물 관련 관리 현황에 별도의 관리 매뉴얼이 없기 때문에 본 식품용수 수질오염 예측모델 개발에 질산성질소를 분석대상으로 선정하였다.

수질검사 자료는 검사 대상 관정별로 검사일이 다양하여 개별적으로 사용할 수 없기 때문에 개별 자료의 변동성을 최대한 억제하되 지하수의 계절별 수질 변화는 반영할 수 있도록 수질검사를 수행한 연도를 기준으로 4분기(1~3월, 4~6월, 7~9월, 10~12월)로 그룹화하였다. 그리고 식품용수로 이용되는 지하수의 수질검사는 공인된 분석기관에 의해서도 수행되지만, 대부분은 개인 사업자의 위탁하에 진행되는 현 상황을 고려하여 개별 수질 관측 자료의 신뢰성이 높지 않을 것으로 가정하여 매우 높거나 낮은 수질분석 값에 의한 영향을 최소화할 수 있도록 분기별로 그룹화된 검사 자료는 중앙값을 산정하여 적용하였다.

자료를 그룹화하여 사용하는 경우, 추세분석에 필요한 최소 자료 수를 충족하지 않을 수 있다. Frollini(2021)는 지하수 오염 추세분석을 진행하기 위해 최소 15개 이상의 데이터가 필요하다고 제안하였으며, 본 연구에서도 표준유역별로 그룹화한 질산성질소의 분기별 중앙값이 15개 이상으로 확인된 76개 표준유역을 추세분석 대상으로 선정하였다.

2.2.2. 주처리과정

주처리과정은 전처리과정을 거쳐 전처리 된 데이터를 활용하여 추세분석을 실시하는 과정이다. 주처리과정에서는 2010년부터 2019년까지 최근 10년간 식품업소 지하수 수질검사 자료를 바탕으로 2020년부터 2025년까지 향후 5년의 수질 추세를 예측하였으며, 지하수 수질 추세분석을 위한 통계적 방법으로 회귀분석, Least Trimmed Squares (LTS), Mann-Kendall test 등의 분석 기법을 사용하였다.

관측기간 동안 추세의 변화를 해석하기 위하여 본 연구를 통해 개발된 자료분할 회귀분석(segmented regression) 기법을 적용하였다. 이는 추세가 자료기간 중 급격하게 변화하는 자료들에 대하여 적요할 수 있는 모수적(parametric) 추세분석 기법이다. 또한, 아래의 모델에 대한 설명에서와 같이 수질자료가 이상치로 심각하게 오염되었을 때도 추세 및 추세의 역전을 예측 가능하도록 설계되어, 일반적으로 이상치에 큰 영향을 받는 선형회귀(linear regression)에 비하여 개선된 모델이다.

2000~2020년까지 10년간의 자료를 사용하는 만큼, 해당 기간 동안 기후변화, 도시개발, 토지이용 변화, 인구변동 등으로 인한 수질변화가 발생할 수 있으며, 이는 다른 추세에 보이는 변곡점으로 나타날 수 있다. 관측 기간 동안 최소 1회의 수질변화가 발생했을 것이라 가정하여 예측 모델에 변곡점을 설정하였으며, 변곡점 설정과 관련된 일련의 과정은 모두 자동화하여 적합한 변곡점으로 분석하는 코드를 추가하여 예측의 정확도를 향상시켰다.

본 연구에서는 하나의 변곡점을 가지고 두 가지의 다른 추세가 있는 경우를 모델링하기 위해 강력한 분할선형회귀 분석 방법을 개발하였다. 또한, 유역을 대표하는 지하수 수질의 변화가 일반적으로 급변하지 않는다는 가정하에 두 가지의 추세가 변곡점에서 만나도록 회귀 모델을 설계하였고, 따라서 분할선형회귀 방정식은 다음과 같다.

$$C(t; \omega_{10}, \omega_{11}, \omega_{20}, \omega_{21}, t_{BP}) = \alpha(\omega_{10} + \omega_{11}t) + (1 - \alpha)(\omega_{20} + \omega_{21}t) \quad (1)$$

여기서 t 는 시간, ω_{10} 과 ω_{11} 은 변곡점 이전의 추세를 설명하는 계수이며, t_{BP} 와 ω_{11} 과 ω_{21} 은 변곡점 이후의 계수이다. 두 추세가 변곡점에서 서로 교차하기 때문에 계수는 아래의 식 (2)와 같이 변곡점과 관련된다. 또한, 식품업소의 지하수 수질검사결과가 각기 다른 분석기관에 의해 생성되었기 때문에, 분석과정에서의 오차 등으로 발생하는 이상치의 영향을 배제하기 위하여 Least Trimmed Squared(LTS) 기법을 적용하였다. LTS는 50%의 Break-down point를 지녀 이상치의 영향을 가장 잘 배제할 수 있는 로버스트 회귀모델(robust regression)의 하나로, 일반적으로 많은 이상치를 포함하는 지하수 수질자료의 적용에 유리하고 선형회귀에 기반하여 계절적 지하수 수질 변동에 용이하게 적용된다는 장점이 있다.

수질의 느린 변화속도에 기초하여 추세의 연속성을 가정할 경우 분할지점은 식 (2)로 결정하였으며,

$$t_{BP} = \frac{w_{21} - w_{11}}{w_{12} - w_{22}} \quad (2)$$

추세식의 α 는 전반기 및 후반기의 지시인자로 식 (3)으로 계산하였다.

$$\alpha = \begin{cases} 1 & t < t_{BP} \\ 0 & t \geq t_{BP} \end{cases} \quad (3)$$

미래의 수질변화 예측 부분에는 식 (4)의 선형회귀 및 추세의 신뢰구간을 이용하였다. 신뢰구간은 평균 수질농도 추세가 위치할 가능성이 99% 이상인 범위로 선정하였으며, 향후 5년에 대한 수질변화 예측을 수행하였다.

$$CI(t) = \mu(t) \pm 2.58 \times \sigma_c \sqrt{\frac{1}{n} + \frac{(t - \bar{t})^2}{(N-1)\sigma_t}} \quad (4)$$

매 자료에 대하여 분할회귀 추정결과는 선형회귀 추정결과와 벤치마크 되었으며, 두 가지의 추정에 대하여 각각 Mann-Kendall test의 p-value 값을 얻었다. 이를 통해, 두 추정 중 더 낮은 p-value 값을 보이는 결과가 합리적인 추세라고 정의하였다.

도출 결과 파일은 Table 3과 같이 5가지 항목이 표기된다. 1)은 추세변화지점(breakpoint)로 이전의 자료 개수를 의미하고, 2)는 추세변화 이후 추세선의 기울기, 3)은 단순 선형회귀 분석결과 추세선의 기울기, 4)는 분할회귀 추정 결과 우측 선형추세의 p-value, 5)는 변곡점이 없는 단순 선형회귀분석추정의 p-value 값을 의미한다. 예를 들어, 유역코드 100701의 결과 값이 1) 32, 2) 0.09, 3) 0.03, 4) 0.08, 5) 0.23 이라면 해당 유역은 P-value 값인 4)와 5) 중 4)의 값이 0.08로 더 작으므로 변곡점이 있는 추세가 변곡점이 없는 추세보다 더 우세함을 뜻하며, 따라서 해당 유역의 기울기 값은 2) 0.09로 상승추세이다.

예측 결과는 그래프로 도출되도록 하였으며, 변곡점이 있거나 없을 가능성 중 더 높은 가능성의 추세가 실선으

Table 3. Trend analysis result file of prediction model

SBNCD	1)	2)	3)	4)	5)
100207	18	0.04	0.02	0.02	0.07
100610	11	0.01	0.01	0.79	0.99
100701	32	0.09	0.03	0.08	0.23
100702	12	0.03	-0.02	0.68	0.10
100703	6	-0.08	-4.80	2.76	0.00
100705	8	0.03	0.02	0.11	0.02
.
.

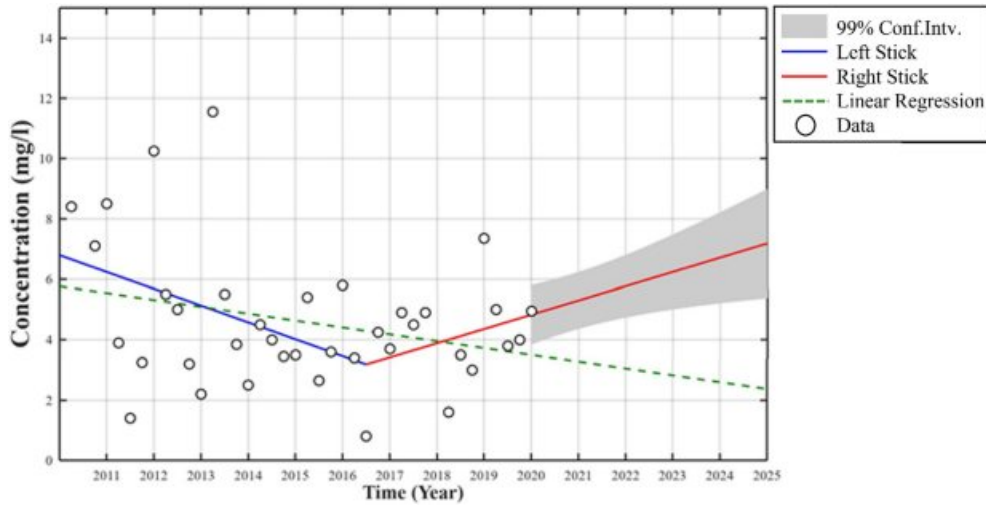
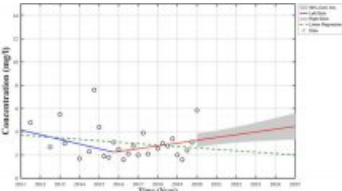
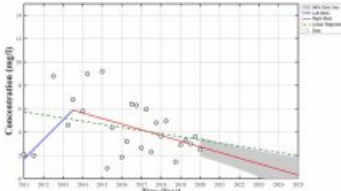
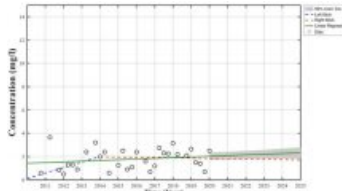
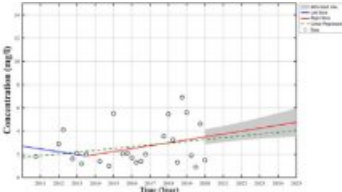
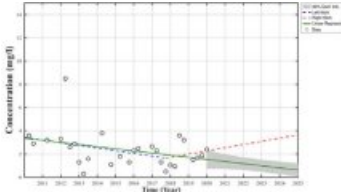
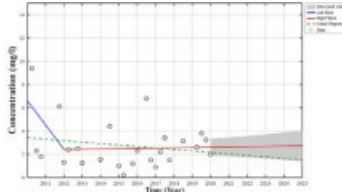
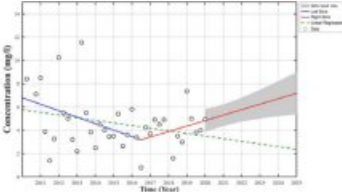
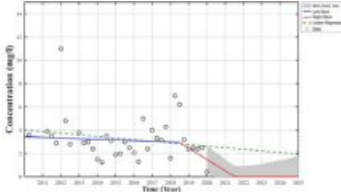
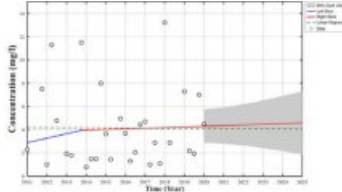


Fig. 3. Graph with breakpoint (Gray range indicates the interval with a probability of 99% or more where the predicted average water quality trend is located).

Table 4. The trend result of classifying as slope

상승추세	하강추세	유지추세
		
		
		

로 표현되고, 나머지가 점선으로 표현되도록 하였다. 또한, 수질농도 추세가 위치할 가능성이 99% 이상인 범위를 회색으로 표현하였다.

예측결과 그래프에서 점선이 변곡점을 설정하지 않은 결과이고 실선이 변곡점을 설정한 결과이다. 변곡점이 설정되지 않았을 때는 수질의 상승 추세를 예측할 수 없으

나 변곡점을 설정하면 2016년도 기점으로 수질이 상승하는 추세를 예측할 수 있다.

76개 표준유역에서 식품용수로 사용되는 지하수의 질산성질소항목에 대한 대표적인 추세분석 및 예측 결과는 Table 4와 같다. 개별 관정으로부터는 획득하기 어려운 표준유역을 대표하는 결과이며, 상승추세 유역 35개, 하강추

세 구역 35개, 유지 구역 6개이다.

2.2.3 후처리과정

후처리과정은 수질 추세 예측 모델을 식품용수 관리에 실 적용하기 위해 추세 결과를 해석하고 등급을 분류하는 과정이다.

추세 예측을 통한 수질 관리는 유럽, 영국, 미국, 제주도 등에서 적용(Korea Environment Institute, Jeju Province, 2020)한 사례와 같이 등급으로 분류하여 관리하는 것이 최선이며, 이를 위해서는 기준 설정이 중요하다. 본 연구에서는 질산성질소의 먹는물 수질기준(MCL, Maximum Contaminant Level)인 10 mg/L와 추세의 기울기 값에 따른 추세 방향을 고려하여 등급을 분류하였다.

먼저, 추세 해석을 위해 10년간의 자료에서 최근 5년에 해당하는 분기별 중앙값이 MCL을 초과하는 경우는 나쁨(poor)으로, 그렇지 않은 경우 좋음(good)으로 구분하고, 이를 다시 문턱값(MCL의 75%)의 초과 또는 미만으로 구분하였다. MCL 초과, MCL 75% 초과, MCL 75% 미만에서 기울기(a) 값에 따라 상승추세($a > 0.01$), 유지추세($-0.01 \leq a \leq 0.01$), 하강추세($a < -0.01$)의 9가지로 구분하였다. 다음으로, 수질 등급은 행정 실무자나 식품업소 등이 해당 구역의 수질을 빠르게 판단하고 조치하기 위해 관리의 중요도와 위험도에 따라 9가지 예측 추세를 긴급, 조치, 관찰, 안심의 4단계 등급으로 분류하였다.

추세 해석 결과 최근 5년간 분기별 중앙값이 MCL을 초과하면서 증가추세인 경우 긴급등급(RED) 지역, MCL을 초과하지만 유지추세이거나 하강추세인 경우 조치등급(ORANGE) 지역, 최근 5년간 분기별 중앙값이 MCL 기준 75%(문턱값)를 초과하면서 증가추세인 경우 관찰등급(GREEN) 지역, 초과하지만 하강, 유지추세이거나 문턱값 미만인 모든 지역은 안심등급(BLUE) 지역으로 분류된다(Table 5).

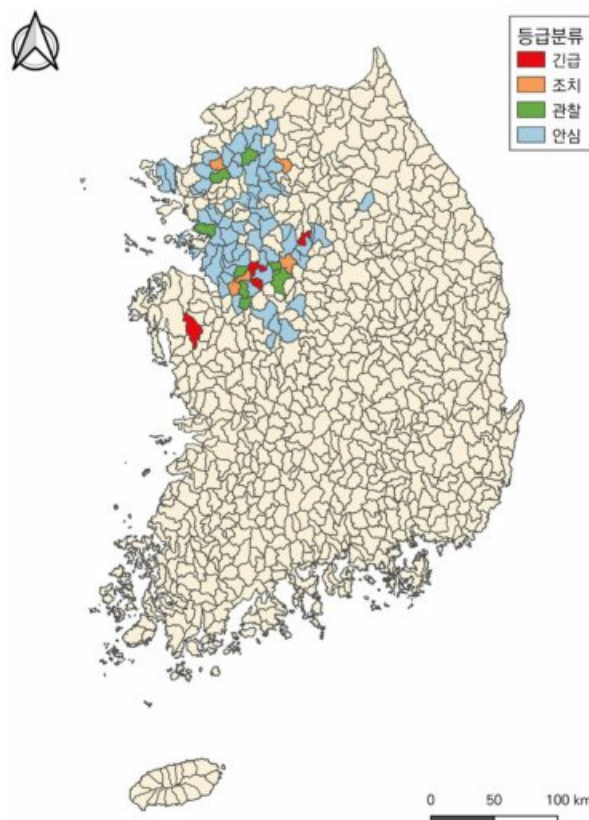


Fig. 4. Distribution of food water quality (groundwater) trend grade.

2.3. 시범지역 검증

개발된 예측 모델을 실무에 적극적으로 활용하기 위해서는 예측 모델의 정확도와 신뢰도 등 모델의 적절성이 확인되어야 하며, 이에 따라 모델을 검증하는 것은 매우 중요한 과정이다.

본 연구에서는 예측 모델을 검증하기 위해 시범지역을 선정하여 수질시료를 채취·분석하였으며, 현장 수질이 예측 추세와 일치하는지를 확인하였다.

Table 5. Standard for trend and grade classification of prediction results

기준		추세	등급	색상
나쁨	poor	MCL 초과 / 증가추세	긴급	RED
		MCL 초과 / 유지추세	조치	ORANGE
		MCL 초과 / 하강추세	관찰	GREEN
좋음	good	MCL 75% 초과 / 증가추세	안심	BLUE
		MCL 75% 초과 / 유지추세		
		MCL 75% 초과 / 하강추세		
		MCL 75% 미만 / 증가추세		
		MCL 75% 미만 / 유지추세		
		MCL 75% 미만 / 하강추세		

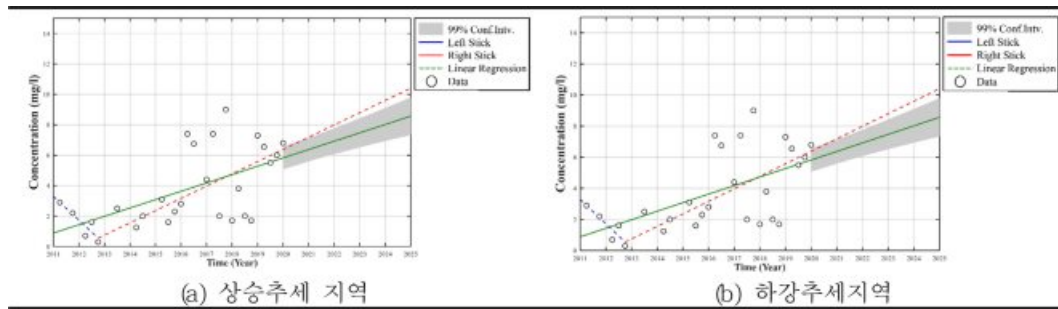


Fig. 5. Trend graph of watershed subject to field test.

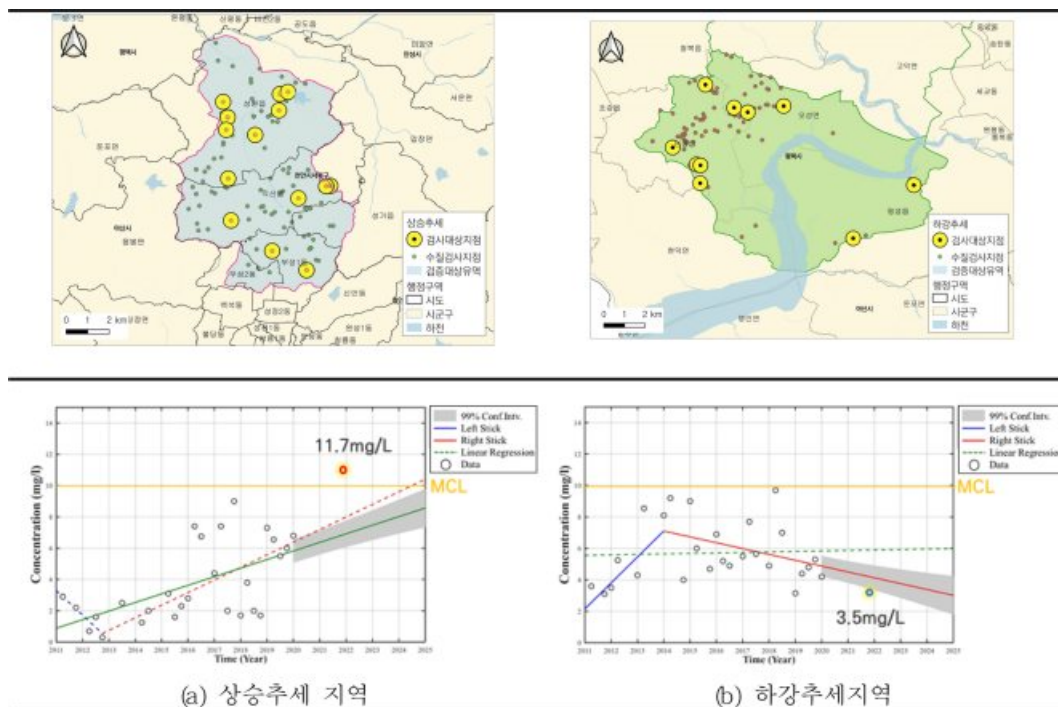


Fig. 6. Distribution of points in the field test basin.

시범지역은 오염 가능성이 있거나 오염된 지역, 그렇지 않은 경우를 모두 검증하기 위해 추세 결과 중 상승추세로 확인된 1개 표준유역과 하강추세로 확인된 1개 표준유역을 각각 선정하였다.

상승추세로 확인된 표준유역의 유역코드는 110106이며, 하강추세로 확인된 표준유역의 유역코드는 110116이다. 앞서 그래프로 도출되는 예측 결과는 변곡점이 있거나 없을 가능성 중 더 높은 가능성의 추세가 실선으로 표현되도록 하였고, 이에 따라 110106은 변곡점이 없는 실선이 상승하는 추세로, 110116은 변곡점이 있는 실선이 하강하는 추세로 더 우세하게 나타난 지역이다(Fig. 5).

두 지역에서의 추세가 상승과 하강으로 예측된 원인을

확인하기 위해 해당 지역의 2015~2018년의 잠재오염원, 인구변동, 토지피복도 등을 검토하였다.

상승추세인 110106 유역은 인구가 353,520명에서 398,877명으로 12.8% 증가하였으며, 가축두수도 45,909마리에서 49,641마리로 8.1% 증가하였고, 하수처리량, 농공단지 가동률, 수질(폐수)배출 사업장 모두 증가한 것으로 확인되었다(Statistical Yearbook of Cheonan, 2020). 토지피복도에서는 과수재배지 면적의 급격한 상승이 확인되어, 질산성질소 농도를 증가시킬 수 있는 비료나 거름 등의 사용량 증가가 예상되었고, 환경부의 토양지하수정보시스템(Statistical Geographic Information Service, SGIS)을 통해 해당 지역은 지하수 오염우려지역으로 지정된 지

역임이 확인되었다.

하강추세인 110116 유역은 인구 증가는 있었으나, 가축 두수가 152,230마리에서 129,099마리로 15.2% 감소하였고, 하수처리량, 수질(폐수)배출 사업장 모두 감소한 것으로 확인되었다(Statistical Yearbook of Pyeongtaek, 2020). 또한 토지피복도에서는 농경지 면적이 급격히 감소한 것으로 확인되었다.

상승추세와 하강추세로 확인된 두 개의 유역에서 각각 유역의 상류와 하류에 조사 지점이 고루 분포하도록 식품용수로 사용되는 지하수 관정 10~15개소를 선정하여 9월, 10월, 11월에 각각 1회씩 총 3회 수질시료를 채취하여 분석하였다.

분석 결과, 상승추세 지역인 110106 유역의 질산성질소 중앙값은 각각 13.70, 10.05, 12.0 mg/L이며, 전 회차의 중앙값 평균은 11.7 mg/L로, 예측 모델 결과 그래프의 가파른 상승추세와 같은 결과를 보였다.

하강추세 지역인 110116 유역의 질산성질소 중앙값은 각각 2.70, 3.70, 3.55 mg/L이며, 전 회차의 중앙값 평균은 3.50 mg/L로, 예측 모델 결과 그래프의 하강추세와 같은 결과를 보여, 본 연구를 통해 개발된 수질 예측 모델의 추세 예측이 정확하며 신뢰도가 매우 높은 것으로 검증되었다.

3. 결 론

본 연구는 식품용수로 사용되는 지하수의 사전예방적 수질관리의 필요성이 대두됨에 따라 기존의 수질검사 자료를 활용하여 장기적인 변화를 예측할 수 있는 모델을 개발하고자 수행되었다.

모델 개발을 위해 지난 10년간 먹는물 검사기관에서 진행된 식품용수 지하수의 수질검사 결과를 수집하여 자료의 형태를 통합·정제 및 DB화 하였으며, 3단계의 PROCESS를 설정하여 수질 오염 예측모델 개발을 진행하였다. 모델은 분할선형회귀, LTS, Mann-Kendall 분석법을 이용하였으며, 수질의 변화를 반영하기 위하여 변곡점을 적용하였고 추세의 신뢰구간은 평균 수질농도의 추세가 위치할 가능성이 99% 이상인 범위가 되도록 개발하였다. 개발된 모델의 결과값은 시범지역 2곳의 시료채취 및 분석을 통해 검증하였고 예측의 신뢰도가 높음을 확인하였다.

수질검사 전체 자료 중 10% 이상이 전처리 과정 중 활용 불가로 사장되었다. 이는 먹는물 수질검사 기관의 성격서 양식의 통일 작업, 결과 자료의 전산화·DB화 등이

필요함을 보여주는 사례이다. 또한, 본 연구를 통해 개발된 예측모델을 이용한 선제적 수질 관리가 지속적으로 이루어지고 향후 빅데이터 기반의 실시간 예측시스템이 운영되기 위해서는 먹는물 수질검사기관에서 생성되는 식품업소의 지하수 수질검사 결과가 관련 부처 예측시스템으로 실시간 자동 전송이 가능하도록 먹는물 검사기관-지자체-환경부-식품의약품안전처 간 통합시스템 구축이 필요하다고 판단된다.

개발된 모델을 통해 산출된 유역별 그래프와 예측 수치를 기반으로 한눈에 전국적인 예측 수질 현황을 파악하기 용이하도록 먹는물 수질기준 MCL과 그래프의 기울기 값(추세의 방향)을 이용하여 긴급, 조치, 관찰, 안심 4단계로 등급을 분류하고 각각 색상을 부여하였다. 제한한 수질 등급분류와 등급별 색상을 토대로 GIS와 연동한 시스템을 통해 시각화하여 식품업체와 소비자에게 제공한다면 유역별 예측 수질 현황을 빠르게 파악할 수 있다. 특히, 식품업자의 경우 업체가 속한 지역의 지하수 미래 수질을 사전에 파악할 수 있어, 사용 중인 지하수질이 먹는물 수질기준에 도달하기 이전에 자체적으로 시설개수, 주변 오염원 관리 등의 조치를 취할 수 있다. 지자체 담당자의 경우 관할 구역이 속한 유역의 수질 등급을 통보 받고 필요한 사전 조치를 취할 수 있어 부적합 수질의 지하수 사용으로 인한 사고를 사전에 차단하여 국민건강에 기여할 수 있을 것으로 기대된다.

식품용수로 사용되는 지하수의 경우 관련법이 상이하고 관계 부처 간 시스템이 통합되지 않는 등 수질관리를 위한 제도적 개선사항이 일부 존재하는 것으로 판단된다. 식품용수의 지하수 관리 매뉴얼 가운데 음료류인 주류 관리 매뉴얼에서는 지하수 수질검사 결과 부적합 시 조치사항을 마련하고 있지만 이미 수질기준을 초과한 이후에 대한 조치이므로 오염을 사전에 관리할 수 있는 방안이 필요하며, 또한, 수질기준 초과로 인한 재검사가 시행되기 전에 원인파악을 통한 시설개수, 오염물질 유입 차단 등 적절한 조치가 선행되어야 한다. 식품관리제도 중 가장 강력하다고 알려진 식품안전관리인증기준(HACCP) 역시 노로바이러스, 식중독균 등에 대한 관리에 초점을 맞추고 있어 전반적인 지하수질 관리에 대한 절차보완이 필요해 보인다. 본 연구에서 제안한 수질등급과 연계하여 식품업체에서 일상적으로 지하수 시설과 오염항목별 지하수질을 관리할 수 있도록 “(가칭)안전한 식품용수 확보를 위한 지하수 관리매뉴얼” 마련이 필요한 실정이다.

본 연구는 오랜기간과 비용이 투입되었지만 미활용되고 있는 수질검사 자료를 활용하여 수질 추세분석 모델 개발

을 통해 단순한 기준 초과여부에 따른 규제가 아니라 식품업소의 지하수 수질을 예측하고 사전적으로 관리할 수 있는 시스템을 제안하였다는 것에 의의가 있으며, 향후 자동화된 예측 시스템 구축에 기초자료로서 과학적 판단 기준을 통한 선제적이고 체계적인 관리를 통해 국민의 식품 안전에 대한 불안감을 해소하고 식품용수에 대한 신뢰도를 높여 식품 산업 발전에 원동력이 될 수 있는 기반을 제공한다는 데 의의가 있다.

사 사

본 연구는 식품의약품안전처의 식품 등 안전관리 R&D 연구개발사업(과제번호 : 20162MFDS034)의 지원을 받아 수행되었습니다. 이에 감사드립니다.

References

- An, H.S., Jin, S.W., Lee, S.J., Hyun, Y.J., Yoon, H.S., and Kim, R.H., 2015, Suggestion of a groundwater quality management framework using threshold values and trend analysis, *Journal of Soil and Groundwater Environment*, **20**(7), 112-120.
- EC, 2000, Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy.
- EC, 2006, Directive 2006/118/EC of the European Parliament and of the Council of 12 December 2006 on the protection of groundwater against pollution and deterioration.
- EC, 2009, Common implementation strategy for the water-framework directive (2000/60/EC) Guidance document No. Guidance on groundwater status and trend assessment.
- Frollini, E.N., Preziosi, E.B., Calace, N.L., Guerra, M.Z., Guyennon, N.C., Marcaccio, M.C., Menichetti, S.F., Romano, F.N., and Ghergo, S.F., 2021 Groundwater quality trend and trend reversal assessment in the European Water Framework Directive context: an example with nitrates in Italy, *Environmental Science and Pollution Research*, **28**, 22092-22104.
- Korea Water Resources Corporation, Jeju Province, 2016, Investigate potential sources of groundwater pollution and prepare management plans for each water quality level.
- Korea Environment Institute, Jeju Province, 2020, Establishment of Jeju-do groundwater quality management standards and preparation of application plan.
- Kim, G.B. and Hwang, C.I., 2020, A proposal of unit watershed for water management based on the interaction of surface water and groundwater, *Journal of Korea Water Resources Association*, **53**(1), 755-764.
- Ministry of Food and Drug Safety, Park, E.G., 2019, Research on the development of a water quality change prediction model to secure safe food processing water.
- The Groundwater Association, 2020, Groundwater Use in the United States of A merica.
- 광주지방식품의약품안전청, 2021, 안전한 주류제조를 위한 주류 위생관리 매뉴얼의약품안전청.
- 식품의약품안전처, 2020, 2020년 안전관리인증기준(HACCP) 평가기준 매뉴얼.
- 천안시, 2020, 2020년 천안통계연보.
- 평택시, 2020, 제25회 2020 평택통계연보.
- 환경부, 2021, 2020 지하수조사연보.