

# GIS기반의 기후-환경 분석 지원을 위한 통합DB 구축 방안†

## A Study on GIS-based Construction of the Climate-Environment Database

남기범\* · 김계현 · 이철용 · 송현오

Gibeom Nam\* · Kyehyun Kim · Cholyoung Lee · Hyunoh Song

인하대학교 공과대학 지리정보공학과

{gib2om\*, khsakura82}@inhaian.net, kye Hyun@inha.ac.kr, Roiz@inha.edu

### 요 약

최근 기후변화에 따른 생태계의 파괴, 해빙, 해수면 상승 등의 현상이 발생함에 따라 코펜하겐에서는 유엔기후변화협약 당사국 간의 온실가스 감축 논의를 위한 총회가 개최되는 등 전 세계적으로 기후변화에 대한 관심이 고조되고 있다. 국내에서는 기후변화가 지구환경시스템과 서로 밀접한 관계가 있음에 따라, 기후변화에 따른 환경의 변화를 함께 고려하는 연구가 진행되고 있다. 그러나 현재 기후-환경 통합 연구에 필요한 자료는 산재되어 있고, 필요한 자료의 확보에 많은 시간과 노력이 필요로 하기 때문에 원활한 연구 진행이 어려운 실정이다. 따라서 본 연구에서는 기후-환경 분야의 분석 지원을 위한 통합DB의 구축방안에 대해 연구하였다. 통합DB의 구축 시 발생하는 문제점에 대해 분석하고, GIS의 적용에 필요한 위치체계의 정의, 데이터 포맷에 따른 관리 방안과 통합DB 구축을 위한 데이터 표준화 방법을 이용하여 GIS기반 기후-환경 통합DB 구축 방안에 대해 제시하였다. 본 논문을 통해 구축되는 통합DB는 기후-환경 데이터의 원활한 공유를 통해 기후-환경 상호작용의 분석 기반을 마련하고, GIS기반의 공간분석에 지원이 가능하리라 기대된다. 나아가 기후-환경 변화에 대한 분석, 모의를 통해 기후-환경 변화에 따른 대응 전략 수립이 가능할 것으로 사료된다.

### 1. 서론

문명의 발달에 따른 산업화는 인류의 발전에 많은 이익을 가져다주었으나, 도시의 인구급증, 부의 불균형, 환경파괴 등의 문제점을 낳았다. 특히 화석 연료 사용 급증은 이산화탄소 및 메탄 등의 온실가스 배출량을 증가시켰고, 이는 지구 온난화 현상을 초래하여 해수면의 상승 및 생태계의 변화, 각종 기후 변화를 야기하였다. 이에 따라, 2009년 12월 덴마크 코펜하겐에서는 유엔기후변화협약 당사국 간 지구 온난화 문제의 해결 방안을 모색

하고, 새로운 기후협약 도출을 위해 총회가 개최되는 등, 전 세계적으로 기후변화에 대한 관심이 높아지고 있다.

국내에서도 기후변화에 대한 많은 연구가 이루어지고 있으며, 특히 기후변화와 지구환경시스템은 서로 밀접한 관계가 있음에 따라 상호작용 및 피드백에 대한 연구가 시도되고 있다. 이러한 기후-환경 상호작용의 분석과 모의는 기후변화뿐만 아니라 환경문제에 대한 사전 예측 및 대비 전략 수립이 가능하게 함으로써, 지속 가능한 사회를 실현할 수 있게 한다.

† 이 논문은 교육과학기술부의 재원으로 한국연구재단(No. 2010-0001906)과 국토해양부의 공간정보 전문인력 양성사업의 지원을 받아 수행된 연구임

그러나 현재 기후-환경의 분석과 모의에 필요한 데이터는 각각의 목적에 따라 독립적으로 구축되었고, 산재되어 있음에 따라 기후-환경 관련 연구자들이 연구를 위해 필요한 데이터를 원활하게 사용함에 있어서 제한적이다. 즉, 필요한 데이터의 소재 파악과 해당 데이터 사용 승인 등의 절차에 많은 시간을 소요하게 됨으로써 전체 연구의 진행이 지연되는 결과를 초래한다. 이와 함께 기후-환경 변화는 지구라는 공간적인 범위 내에서 일어나는 현상으로 이를 시각적으로 표현하고 공간 분석이 가능하게 하기 위해서는 GIS의 도입이 필수이나, GIS의 적용을 위해 필요한 위치체계가 기후-환경 분야마다 상이하고, 위치체계가 정의되지 않은 분야도 존재하는 실정이다. 이를 해결하기 위해서는 기후-환경 데이터의 원활한 공유를 위한 통합DB의 구축과 함께, GIS기반의 표출 및 공간분석을 위해 기후-환경 데이터가 현재 적용하고 있는 위치체계의 명확한 정의와 관리가 필요하다.

따라서 본 연구에서는 기후-환경의 상호작용 분석과 공간분석을 지원하기 위한 GIS기반의 기후-환경 통합DB 구축 방안에 대해 연구하였다.

## 2. 기후-환경 데이터의 통합 시 문제점

### 2.1 데이터 중복

기후와 환경은 서로 유기적으로 작용함에 따라, 분석과 모의에 공통적으로 사용되는 데이터가 존재한다. 이러한 기후-환경 데이터는 각 분야의 목적에 따라 독립적으로 구축되었고, 각각의 분야에서만 사용될 시에는 아무런 문제가 없었다. 그러나 기후-환경 통합 연구를 위해 통합DB가 필요하고, 공통되는 데이터에 대한 일련의 처리과정 없이 통합DB를 구축하게 되면 데이터가 중복으로 저장되는 상황이 발생한다. 중복된 데이터는 한 DB 내에서 불필요한 공간을 차지할 뿐만 아

니라, 각 분야의 분석에서 서로 다른 데이터를 사용하여 잘못된 결과를 도출하게 되므로 해결방안이 필요하다.

### 2.2 데이터 포맷의 상이함

기후-환경 데이터는 각 분야에서 사용하는 모델이나 분석 방법에 따라 데이터 포맷이 상이하다. 일반 텍스트 형태의 아스키 코드로 존재하는 데이터가 있는 반면, 엑셀이나 이미지 등 기타 특정 프로그램을 통해 읽을 수 있는 형태인, 바이너리 코드로 존재하는 데이터도 있다. 아스키 코드로 이루어진 데이터는 데이터를 통합DB에 입력하는데 큰 어려움이 없지만, 바이너리 코드의 데이터는 통합DB 내에서 사용자가 알아 볼 수 있는 문자 형태로 표현하기 위해서는 변환과정이 필요하다. 데이터의 변환 방법은 파일의 형태에 따라 상이하기 때문에 각 데이터에 따른 방법의 개발에 많은 시간이 소요된다. 또한 기후 데이터는 지정된 해상도를 기준으로 4차원에 걸쳐 데이터가 존재한다. 한 개의 격자 내에서도 100여 가지의 변수가 존재하기 때문에 많은 양의 데이터를 변환을 통해 통합DB에 직접적으로 입력하는 것은 시간적이나 데이터의 활용 차원에서 비효율적이다. 따라서 이러한 상이한 포맷의 데이터에 대해 적합한 저장 및 관리방안이 필요하다.

### 2.3 위치체계의 상이함

기후-환경 분야의 데이터는 해당 분야의 분석 범위와 연구의 필요에 맞게 서로 다른 위치체계를 적용하고 있다. 몇몇의 데이터는 위치체계의 사용 없이 해당 데이터의 관측 도시와 관측 값만을 가지고 있는 것도 존재한다. 비록 DB구축 단계에서는 분야 간의 위치체계를 일치시키지 않아도 되지만, 향후 GIS기반의 분석시스템을 개발하여 데이터를 표출, 분석하기 위해서는 현재 적용된 위치체계를 명확하게 정의하여 관리하는 것이 중요하다. 따

라서 데이터에 따라 사용하는 위치체계를 정확하게 정의하고, 위치체계가 정의되지 않은 데이터의 경우에는 하나의 체계를 선정하여 적용하는 것이 필요하다.

### 3. 연구내용

#### 3.1 연구범위 설정

기후-환경에는 수많은 분야가 존재한다. 본 연구에서는 몇 가지의 기후-환경 분야를 선정하고, 이를 기반으로 GIS기반의 기후-환경 통합DB 구축 방안을 마련하였다. 선정한 분야는 기후, 생물계절, 토양, 위성영상, GIS이고, 선정 기준은 기후-환경 분야에서 다양한 데이터 형태를 고려할 수 있는 범위 내에서 선정하였다.

기후분야는 기후변화에 대한 분석과 미래 기후변화 시나리오의 생성을 담당하고, 생물계절분야는 기후변화에 따른 식생, 계절의 변화 양상을 연구하는 파트이다. 토양분야는 기후변화에 따른 미세규모의 토양 변화를 예측하고, 위성영상은 지표면 및 대기 온도 자료 및 검증 자료로 활용된다. 그리고 GIS데이터는 지형을 표현하고, 데이터의 시각적 표현 및 공간분석에 활용되는 자료이다.

기후-환경 데이터의 목록은 각 분야에서 활용하는 모델이 존재할 경우에는 모델에 사용되는 데이터를 기준으로 하고, 없는 경우에는 해당 분야의 분석에 필요한 데이터를 기준으로 하였다. 이에 따라, 기후 분야에는 CCSM(Community Climate System Model), 토양 분야에는 LSPM(Land Surface Process Model)을 선정하였고, 생물계절과 위성영상, GIS데이터는 해당 분야에 사용되는 데이터를 기준으로 하였다.

#### 3.2 분야별 데이터 분석

앞서 선정된 기후-환경 분야의 데이터 범위를 기준으로 데이터를 분석하였다. 데이터의 분석을 통하여 공통적으로 사용되는 데이터를 찾아내고, 이를 통해 기후

-환경 데이터 간의 관계를 정의하였다(그림 1).

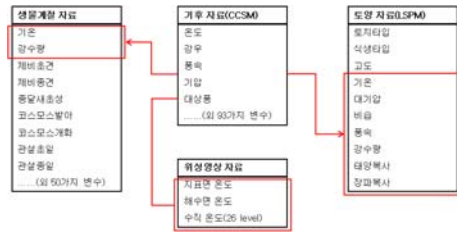


그림 1. 기후-환경 데이터 관계 정의

미래 기후의 모의를 위해 사용하는 CCSM은 4차원(X, Y평면과 높이, 시간) 범위에 대해 분석이 가능한 모델이다. 모의 단위는 격자로서 평면 해상도는 T85(140km)이고, 수직으로는 26레벨, 시간은 20분 단위로 98개의 변수에 대해 모의가 가능하도록 설계되었다. 대표적인 변수로는 경도, 위도, 표면온도, 기압, 풍속, 강수량 등이 있다.

생물종의 계절 변화 양상을 파악하기 위해 사용되는 생물계절 자료는 기온, 강수량과 함께 수십 가지의 연도별 생물 종의 관측 시기 자료가 있다. 분석에 쓰이는 기온, 강수량 자료는 기후 분야와 공통된 자료이고, 미래 기후 자료의 사용을 통해 미래의 생물종 계절 변화 양상을 파악할 수 있다.

토양 분야의 LSPM에 필요한 데이터는 토지타입과 식생타입 자료, 고도 자료와 함께 기온, 대기압, 비습, 풍속, 강수량, 태양복사, 장파복사의 기후 자료가 있다.

위성영상 자료는 기후 분야와 온도 자료의 중복이 있지만 타 분야의 검증자료로 활용함에 따라 통합DB에 입력하여 사용하도록 한다.

기후-환경 변화 연구의 기본 틀은 미래의 기후변화에 따른 환경변화의 분석이므로, 기후 자료를 기준으로 하여 환경 분야에서 공통되는 자료를 조사하고, 중복되는 자료는 데이터베이스에 입력 시 제외되어야 한다.

### 3.3 데이터 관리방안

기후 데이터는 일단위의 파일로 구성되어 있고, 파일은 바이너리 코드 형태로 되어 있으며, 한 개 파일의 용량은 약 1기가바이트 정도이다. 아스키 데이터와는 달리 바이너리 데이터는 통합DB 내에 테이블 단위로 넣고 관리하기 위해서는 코드 변환과정이 필요하다. 특히 기후 데이터의 경우에는 그 양이 방대하고, 데이터 변환을 위한 프로그램의 개발도 필요하기 때문에 테이블 단위로 저장하는 방법은 비효율적이다. 이 같은 경우에는 메타데이터를 구성하여 관리하는 방법이 저장면에서나 데이터 활용 면에서 더 효율적이다(표 1)([1][2]).

추가적으로 구축된 공간자료의 속성과 본 데이터를 연계하여 향후, 정보조회시스템 개발에 따른 시각적 표출 시 이용 가능하다. 생성된 생물계절 공간 데이터는 GIS 데이터와 함께 공간 자료 테이블을 생성하여 관리한다.

LSPM에 사용되는 토지타입, 식생타입, 고도 자료는 1회성 구축 자료로써, 그 업데이트 간격이 일정하지가 않다. 이들 자료는 기후 자료와 마찬가지로 바이너리 코드로 되어 있고, 작은 수의 파일 내에 모든 자료가 존재한다. 따라서 자료에 대한 정보를 테이블로 작성하여 관리하도록 한다(표 2).

표 1. 기후 자료 메타데이터 예시

순번	시나리오	날짜	해상도(km)	수직차원	시간간격	파일명
1	20C3M	1990-05-04	140×140	26levels	20min	T85_20C3M_19900504.nc
2	A2	2020-01-08	140×140	26levels	20min	T85_A2_20200108.nc
3	A2	2020-01-09	140×140	26levels	20min	T85_A2_20200109.nc
4	A1B	2020-01-08	140×140	26levels	20min	T85_A1B_20200108.nc

생물계절 자료는 HTML로 구성되어 있음에 따라 통합DB 내에 테이블을 구성하여 직접 저장하는 방법을 사용하면 된다. 그러나 공간상에 데이터를 표현하기 위한 수치적인 위치정보가 없으므로 도형 자료의 생성이 필요하다. 이 때 위치의 기준은 지역명이 되고, 구글 어스 등을 이용하여 지역 내 관측소의 좌표를 파악한 후, 포인트 형태의 공간자료를 구축하면 된다(그림 2).

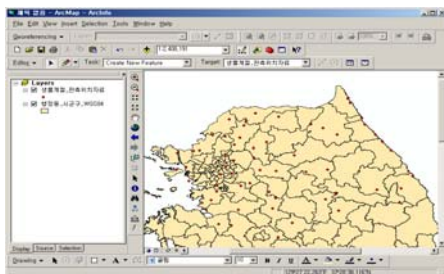


그림 2. 생물계절 관측소 공간 자료 예시

표 2. LSPM 자료 관리 테이블

종류	파일명	설명
ECOCLIMAP	ecoclimap_v2	토지타입
KLCV	KLCV_KOPE_1KM_A MODIS_ISOD_V1.bin	식생타입
	Standard_KOPE_La mc_Lon.bin	식생 좌표 정보(경도)
GTOPO30	Standard_KOPE_La mc_Lat.bin	식생 좌표 정보(위도)
	E020N40.DEM	고도 E020N40
	E100N40.DEM	고도 E100N40
	E100N90.DEM	고도 E100N90

위성영상 자료는 자료의 취득 시 xml 파일 형태의 메타데이터를 제공한다. 이 파일에는 해당 위성영상의 플랫폼 및 센서명, 관측시간, 동서남북 경계의 좌표,

파일명, 파일용량 등 48가지의 메타 정보를 제공한다. 따라서 해당 xml 파일을 이용하여 자료를 사용하는 방법을 이용한다.

GIS자료는 생물계절 분야에서 생성된 공간 데이터와 함께 토양자료와 마찬가지로 파일에 따른 테이블을 생성하여 저장, 관리하도록 한다. 이 때 칼럼은 분야와 데이터명, 파일명만 관리하고, 위치체계 관리 테이블과의 조인을 통해 파일의 위치체계 정보를 얻도록 한다.

### 3.4 위치체계의 정의

기후-환경 분야는 공간적인 범위 내에서 변화를 분석하고 모의하는 것이므로, 대부분이 위치체계를 가지고 있다. 그러나 각 분야에서는 분석 대상 공간에 적합한 위치체계를 사용함에 따라, 분야 별 위치체계가 상이하다. 심지어는 분야 내에서도 자료 마다 위치체계가 다른 경우도 존재한다. 그러므로 크게는 분야, 세부적으로는 분야 내 자료의 위치체계를 명확하게 정의하는 것이 필요하다(표 3).

표 3. 데이터 별 위치체계 관리 테이블

분야	데이터	해상도	좌표계	데이터 범위	파일 형태
기후	미래 기후 시나리오	140km	GRS80/경위도	전세계	nc
생물계절	년도별 시기자료	-	-	한국 74개 지점	xls, html
	관측소 위치자료	-	WGS84	한국 74개 지점	shp, dbf
토양	GTOPO30 (고도자료)	1km	WGS84	전세계	dem
	KLCV (식생자료)	1km	GRS80/위경도	한반도	bin
	ECOCLIMAP (토지자료)	1km	WGS84	전세계	-
위성영상	MODIS (MYD11C1)	약5.6km	WGS84	전세계	hdf, xml, jpg
GIS	국가경계	-	WGS84	전세계	shp, dbf
	행정경계	-	WGS84	한국	shp, dbf

### 3.5 데이터 표준화

데이터 표준화는 데이터가 통합DB에 입력되기 전, 데이터에 대한 명칭과 정의, 형식, 규칙 등을 표준화하는 것으로써, 필요한 데이터의 소재 파악에 소요되는 시간 및 노력을 줄일 수 있고, 일관된 데이터 형식 및 규칙의 적용으로 인한 데이터 품질 향상을 기대할 수 있다[3][4].

데이터 표준화에서는 일관된 데이터 형식을 정의하기 위해 표준단어, 표준용어, 표준도메인, 표준코드의 개념을 사용한다. 표준단어는 표준용어를 구성하는 최소한의 의미를 가지는 단어로써 용어에 대한 한글명과 영문명을 일관되게 정의하는 역할을 한다. 표준도메인은 칼럼에 대한 형태를 그룹핑한 개념으로써, 동일한 성질을 가진 칼럼의 데이터 타입 및 길이를 일관되게 관리할 수 있으며, 향후 칼럼 값에 대해 공통적인 데이터 검증 규칙의 적용이 가능하다. 표준코드는 도메인의 한 유형으로서 특정 도메인 값(코드값)이 이미 정의되어 있는 도메인이다. 표준용어는 통합DB 내 자료의 용어에 대한 표준을 정의하는 것으로써 용어의 사용 및 적용에 대한 혼란을 방지하고 자료 사용의 이해를 돕는다.

표준화를 위해 표준단어, 표준도메인, 표준코드, 표준용어에 대한 설계지침을 마련하고, 지침에 따라 작성된 표준은 테이블로 관리하여(그림 3), 향후 데이터 추가 시 일괄 적용이 가능하도록 한다.

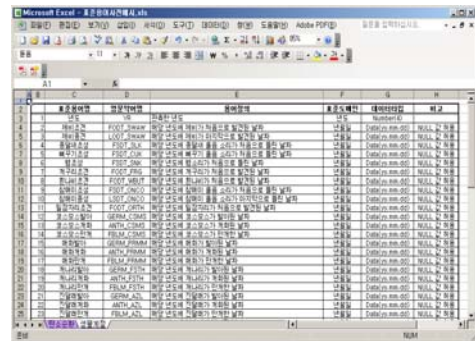


그림 3. 표준용어사전 예시

## 4. GIS기반 기후-환경 통합DB

### 구축방안

GIS기반의 기후-환경 통합DB를 구축하기 위해서는 산재되어 있는 기후-환경 데이터를 통합하는 과정에서 발생하는 문제점의 해결과 함께, 데이터의 위치체계의 정의, 데이터 포맷 별 관리 방안, 공간데이터의 관리 방안, 데이터 표준화 과정이 필요하다. 데이터 통합과정에서 발생하는 중복 데이터는 데이터의 분석을 통해 파악하고, 중복을 최소화 하는 방향에서 데이터베이스를 구축해야 한다. 데이터베이스 구축 과정에서 데이터 표준화 작업은 데이터의 명칭과 정의, 형식, 규칙을 명확히 함으로써, 필요한 데이터의 소재 파악에 소요되는 시간을 절약할 수 있고, 일관된 데이터 형식 및 규칙의 적용으로 데이터 품질의 향상을 기대할 수 있으므로 반드시 수행되어야 한다.

그리고 기후-환경 데이터는 데이터에 따라 위치체계가 상이함에 따라, 향후 GIS기반의 공간분석 시 정확한 위치체계의 적용을 위해 각 분야의 자료가 사용하는 위치체계가 명확하게 관리되어야 한다.

또한, 기후-환경 데이터는 분야마다 포맷이 상이함에 따라, 아스키 형태의 텍스트 데이터는 데이터베이스 내에 테이블의 구성을 통해 저장하고, 바이너리 코드 형태의 대용량 데이터는 메타데이터를 구성하여 저장, 관리하는 것이 효율적이다.

## 5. 요약 및 결론

본 연구에서는 기후-환경 변화 분석의 지원을 위한 GIS기반 기후-환경 통합DB 구축 방안에 대해 연구하였다. 연구의 진행을 위해 기후, 생물계절, 토양, 위성영상, GIS 분야를 선정하여 이들 데이터의 분석을 통해 연구를 진행하였다.

먼저 산재되어 있는 기후-환경 데이터의 통합 시 발생하는 문제점에 대해 분석하고, 이를 기반으로 해결 방안을 모색하였다. 각 분야에서 사용하는 모델을 기

으로 사용하는 데이터를 분석하여 분야 간 공통되는 데이터를 조사하고, 이를 기반으로 상관관계의 정의 및 중복 데이터를 제거하였다. 이후 다양한 형태의 데이터를 관리하기 위해 분야별 데이터 파일의 코드에 따라 저장 방법을 달리 적용하였다. 아스키 코드로 된 데이터의 경우 통합DB 내 테이블을 구성하여 저장하는 방법을 사용하였고, 바이너리 코드로 구성된 대용량 파일의 경우에는 메타데이터를 구성하여 저장, 관리하는 방안을 제시하였다. 그리고 분야, 파일에 따라 달리 적용된 위치체계를 명확하게 정의, 관리하여 공간분석의 기반을 마련하였다. 마지막으로 데이터 표준화를 통해 데이터에 대한 명칭과 도메인을 일괄적으로 적용하는 방안을 제안하였다.

본 연구를 통하여 구축되는 GIS기반의 기후-환경 통합DB는 기후-환경 연구자들의 원활한 연구를 지원하고, 나아가 GIS기반의 기후-환경 분석시스템에 적극 활용될 수 있을 것으로 기대된다. 또한, 제시된 데이터 표준화와 위치체계 관리, 대용량 데이터의 관리를 위한 메타데이터의 사용은 향후 통합DB의 확장 시 적용이 가능하여 확장에 필요한 시간 및 노력을 절감할 수 있을 것으로 판단된다.

그러나 통합DB의 구축만으로는 기후-환경 자료를 활용하는데 한계가 있으므로, 향후 데이터의 체계적인 관리와 함께 데이터의 조회, 입력, 수정, 삭제 등의 기능을 지원하는 DB관리시스템의 개발이 필요하다고 사료된다.

### 참고문헌

- [1] 김경배, 이용주, 박춘서, 신범주, GIS에서 대용량 파일을 위한 대용량 공유 디스크 파일시스템의 메타데이터 구조 “, 한국GIS학회지 제10권 제1호, pp.98-111, 2002.3
- [2] 김상완, 곽재혁, 함재균, 황영철, 관계형 데이터베이스를 이용한 계층적

파일 시스템의 메타데이터 관리 방법,  
한국정보과학회 2006 가을  
학술발표논문집 제33권 제2호(A),  
pp.547-551, 2006.10

- [3] 한국데이터베이스진흥센터, “2006  
데이터베이스백서”, 2006.6
- [4] 김진섭, 데이터표준화 사례를 통한  
데이터 품질 향상에 대한 연구,  
한국정보과학회 가을 학술발표논문집  
Vol.33, No.2(C), pp.210-213, 2006