

스크립트 언어를 사용한 DSSAT 모델 기반 격자형 작물 생육 모의 시스템 개발

유병현¹ · 김광수^{1,2*} · 반호영^{1,2}

¹서울대학교 식물생산과학부, ²서울대학교 농업생명과학연구원
(2018년 5월 14일 접수; 2018년 8월 29일 수정; 2018년 9월 4일 수락)

Development of a gridded crop growth simulation system for the DSSAT model using script languages

Byoung Hyun Yoo¹, Kwang Soo Kim^{1,2*} and Ho-Young Ban^{1,2}

¹Department of Plant Science, Seoul National University, Seoul, Korea

²Research Institute of Agriculture and Life science, Seoul National University, Seoul, Korea

(Received May 14, 2018; Revised August 29, 2018; Accepted September 4, 2018)

ABSTRACT

The gridded simulation of crop growth, which would be useful for shareholders and policy makers, often requires specialized computation tasks for preparation of weather input data and operation of a given crop model. Here we developed an automated system to allow for crop growth simulation over a region using the DSSAT (Decision Support System for Agrotechnology Transfer) model. The system consists of modules implemented using R and shell script languages. One of the modules has a functionality to create weather input files in a plain text format for each cell. Another module written in R script was developed for GIS data processing and parallel computing. The other module that launches the crop model automatically was implemented using the shell script language. As a case study, the automated system was used to determine the maximum soybean yield for a given set of management options in Illinois state in the US. The AgMERRA dataset, which is reanalysis data for agricultural models, was used to prepare weather input files during 1981 - 2005. It took 7.38 hours to create 1,859 weather input files for one year of soybean growth simulation in Illinois using a single CPU core. In contrast, the processing time decreased considerably, e.g., 35 minutes, when 16 CPU cores were used. The automated system created a map of the maturity group and the planting date that resulted in the maximum yield in a raster data format. Our results indicated that the automated system for the DSSAT model would help spatial assessments of crop yield at a regional scale.

Key words: Parallel processing, GIS, Crop model, NetCDF, Mask



* Corresponding Author : Kwang Soo Kim
(luxkwang@snu.ac.kr)

I. 서 론

특정 지역내에서 이상기상 및 기후변화에 의한 작물 생산성의 변동을 추정하기 위해 작물 생육 모델들이 사용되어 왔다. 특히, 지역 특이적인 기후변화 적응 대책을 수립하기 위해 작물모델을 이용하여 격자별로 작물의 공간적인 생산성 변동 양상들을 분석한다. 예를 들어, Rosenzweig *et al.*(2014)은 5 개의 전지구 모델로부터 수집된 격자형 기상자료를 7 개의 작물모델 입력자료로 사용하여 2100년 까지의 작물 생산성 변화를 모의하였다. 국내에서도 작물에 대한 기후변화 영향을 평가 하기 위해 지점 기반의 모델인 ORYZA 2000모델이나 DSSAT (Decision Support System for Agrotechnology Transfer) 모델을 사용하여 공간적인 작물생산성 변동 예측 연구를 수행하였다(Kim *et al.*, 2017; Lee *et al.*, 2011; Kim *et al.*, 2013). 예를 들어, Lee *et al.*(2011) 는 기상청 격자형 기후변화 시나리오 자료의 지점별 자료를 바탕으로 ORYZA 2000 모델을 구동하여 적정 출수기 등의 변화를 전망하였다.

일정 지역내에 위치하는 농경지를 대상으로 작물모형을 구동하기 위해 상당한 전산자원이 필요하다. 예를 들어, 전구 수준에서 0.5°의 해상도의 기후변화에 따른 작물 생산성 영향평가를 수행할 경우, 2.5×10^7 번의 작물모델 구동이 필요하다. 그러나, 실질적인 생육모의를 위해 재배관리 시나리오가 사용될 경우, 이 보다 상당히 많은 생육모의가 수행되어야 한다. 따라서, Elliott *et al.*(2014)은 격자형 작물 생육 모의를 수행하기 위해 96 개의 노드로 구성된 슈퍼컴퓨터를 사용하였다.

격자형 작물모형 구동을 위해 요구되는 전산자원을 만족시키기 위해, 슈퍼컴퓨터와 같이 여러 대의 컴퓨터들을 단일 시스템처럼 활용할 수 있는 클러스터 컴퓨터를 활용할 수 있다. 그러나, 일반 연구자들의 경우 이러한 전산자원의 구축이나 확보가 어렵다. 예를 들어, 클러스터 시스템에 사용되는 노드들에는 보조기억장치가 탑재되지 않아 네트워크를 통해 운영체제를 전달받는 방식이 주로 사용된다. 또한, 클러스터 컴퓨터에 적합한 작물모델 구동을 위해, 작물모델의 소스코드를 수정해야 할 수도 있다. 따라서, 일반적으로 사용할 수 있는 워크스테이션에서 격자형 작물 생육 모의를 가능하게 할 수 있는 시스템이 개발될 경우, 보다 많은 연구자들이 지역규모에서의 작물 생산성 변동 양상을 파악하는 연구에 참여할 수 있다.

본 연구에서는 단일 워크스테이션에서도 활용 가능한 격자형 작물모형 구동 시스템을 개발하고자 하였다. 특히, Linux 운영체제에서 기본적으로 제공되는 스크립트 언어와 R을 기반으로 전문적인 컴퓨터 프로그래밍 개발자가 아니더라도 활용할 수 있는 대용량의 작물 생육 모의를 수행하기 위한 시스템을 개발하고자 하였다. 또한, 작물 모델들의 소스코드를 변경하지 않고도 격자단위에서 작물 생육 모의를 수행할 수 있는 시스템을 개발하는 것을 목표로 하였다.

II. 재료 및 방법

2.1. DSSAT 모델

DSSAT은 과정 중심 모형으로 물리화학적 수식을 바탕으로 작물의 일별 생육을 모의하는 모형이다. DSSAT은 주요 식량 작물인 벼, 옥수수 등을 포함한 40여개의 작물에 대한 생육 모의를 지원하며, 기상, 토양, 재배관리 등의 입력 자료를 필요로 한다. 기상 및 토양 등의 환경 요인 뿐 아니라 시비, 관개 등 재배 관리에 대한 고려가 가능하므로 특정 지점에서 재배관리 방법에 따른 생산성을 모의할 수 있다.

DSSAT 모델(version 4.6)은 일최고기온(°C), 일최저기온(°C), 강우량(mm d⁻¹), 일사량(MJ m⁻² d⁻¹)의 4 가지 필수 기상 변수만으로 구동이 가능하며 6개의 보조적인 기상 요소의 고려가 가능하다(Hoogenboom *et al.*, 2015). 추가적인 기상 요소로는 증발산량 산정과 관련된 이슬점 온도(°C), 상대습도(%), 수증기압(kPa), 풍속(km d⁻¹)이 포함되며, 광합성량 계산과 관련된 유효광합성복사량(MJ m⁻² d⁻¹) 및 CO₂(ppm) 농도의 입력을 지원한다.

DSSAT 모델에 사용할 수 있는 기상입력자료를 생성하기 위해 특정한 형식을 가진 텍스트 파일로 기상자료들을 정리해야 한다. 예를 들어, WTH 확장자를 가지는 기상 입력 파일을 구성하기 위해 테이블 형태로 일별 기상자료가 저장된다. 또한, 기상입력 파일의 헤더 부분에는 해당 기상자료가 얻어진 지역의 정보와 제공된 기상 자료의 변수명을 지정하기 위해 네자리의 지역 코드, 위도(°), 경도(°), 고도(m)와 연평균 기온(°C), 연평균 일교차(°C), 기준 높이(m) 등의 정보가 저장된다.

2.2. R script 기반의 기상입력자료 생성 도구 구현

작물 모델에서 사용하는 기상 입력자료들은 ASCII (American Standard Code for Information Interchange)

형식의 텍스트 자료들이다. 반면, 공간적인 기후 및 기상자료인 미래 기후변화 시나리오자료와 수치기상 모델 출력 자료들은 주로 바이너리 형식의 격자형 자료들이다. 예를 들어, CORDEX(Coordinated Regional Downscaling Experiment)에서는 수치기상 모델을 통해 각 지역별로 상세화된 자료를 network Common Data Form (NetCDF) 형식으로 제공한다. Agricultural Model Intercomparison and Improvement Project (AgMIP) 에서 작물 모형에 대한 검증을 지원하기 위해 전세계의 과거 기상 자료를 재분석하여 격자형식으로 생산된 AgMERRA (AgMIP climate forcing dataset based on the NASA Modern-Era Retrospective Analysis for Research and Applications) 자료 역시 NetCDF 형식이 사용된다(Ruane *et al.*, 2015). 이러한 격자형 자료를 처리하기 위해 다양한 도구들이 개발되어 왔다. Yoo *et al.*(2017)은 격자형 지역 기후 자료인 CORDEX 자료를 처리할 수 있는 CDSL을 개발하였으며, Zender(2008)는 NetCDF 파일을 전문으로 다루는 NetCDF Operators 를 개발하였다. 그러나, 이러한 도구들은 NetCDF, HDF, ESRI Grid 등 다양한 격자 파일 형식 중 일부만 처리가 가능하거나, 통합 처리가 가능한 도구들도 작물모델의 입력자료를 생성하기 위한 기능들이 제공되고 있지 않다.

격자자료로부터 DSSAT 모델의 기상 입력자료를 생성하기 위해 R script 기반의 함수들을 구현하였다. R은 공개 통계 패키지로, 통계 처리에 필요한 다양한 함수를 제공하고 있으며 R 패키지의 형태로 다양한 추가 기능들이 제공된다. 특히, 격자형 자료가 가지는 대용량 공간적 정보를 처리하여 지점별 기상입력자료를 생성하기 위해 필요한 GIS 자료나 병렬처리를 지원하는 R 패키지들이 제공되고 있다. 예를 들어, raster 패키지의 경우 래스터 형식의 격자 자료의 입력 및 공간 분석 기능을 가지고 있다(Hijmans *et al.*, 2014).

격자자료를 사용하여 원하는 지점에서의 기상입력 자료를 생성하기 위해 R 스크립트를 기반으로 특정 기능을 수행하는 모듈들로 구성된 도구를 개발하였다. 이러한 모듈 중심의 개발 방식은 코드의 재사용과 유지 보수의 측면에서 유리하다. 이들 도구는 관심지역을 선정하기 위한 모듈, 기상자료 지정 모듈, 기상자료 입력 모듈, 기상자료 처리 모듈 및 기상입력자료 출력 모듈 등으로 구성되도록 설계되었다. 특히, 워크스테이션뿐 아니라 일반 데스크탑 컴퓨터에서도 소켓 클러스터를 활용한 병렬처리를 지원하는 doSNOW 패키지

를 활용하여 병렬처리 기능을 구현하였다(Analytics and Weston, 2014).

처리 대상지역을 선정하기 위해 공간정보를 가진 Mask 격자자료를 사용하거나 사용자가 지정한 위치를 관심지역으로 선정하기 위한 함수들을 정의하였다(Fig. 1). raster 패키지에서는 특정 위경도 좌표에 위치하는 격자를 찾는 기능을 제공하고 있어 이를 활용하면 관심 지역의 자료만을 추출할 수 있다. 따라서, 관심 지점의 위경도 좌표값 입력을 지원하기 위해 Mask 격자자료로부터 각 격자별 위경도 좌표값을 계산하여 저장하는 readMaskGrid 함수와 사용자가 텍스트 파일로 저장한 좌표를 읽어오는 readPoints 함수를 구현하였다.

다양한 종류의 격자형 기상자료들을 처리하기 위해, 처리 대상 자료의 고유한 파일명 형식에 맞추어 파일 이름들을 지정할 수 있는 함수가 구현되었다. 예를 들어, AgMERRA 자료의 경우, "AgMERRA", 연도, 및 변수이름이 "_" 로 연결되어 있는 형식을 가진다. 처리 대상 격자파일들을 지정하는 createFilelist 함수는 이러한 형식에 맞추어 입력된 연도와 변수별로 AgMERRA 자료의 파일명을 지정하고 자료를 동적으로 관리할 수 있는 list 형 변수에 이들을 할당한다. 격자형 자료의 파일이름 형식에 따라 createFilelist 함수를 수정할 경우, 격자형 파일을 읽는 함수들을 그대로 사용할 수 있어 비교적 쉽게 AgMERRA 이외의 격자자료들을 처리할 수 있다. 예를 들어, CORDEX 자료의 파일 이름은 도메인명, 변수명, CMIP5 관련 정보, GCM과 RCM 정보, 시간해상도 및 시작, 종료 연도 등으로 구성되어 있다. 따라서, 이러한 파일명 규칙을 반영하도록 createFilelist 함수를 수정한다면 CORDEX 자료를 활용하여 작물 모델의 기상입력자료를 생성할 수 있다.

작물모델의 기상 입력자료를 생성하기 위해 사용되는 격자형 자료를 불러들이기 위해 readGrid 함수를 작성하였다. readGrid 함수는 기상입력자료 생성을 지원하기 위해 지정된 격자형 기상자료를 위경도 좌표 및 시간 축을 가지는 3차원 배열 형태로 저장한다. 불러들인 격자형 기상자료는 기상 변수명과 연도 등의 메타 데이터와 함께 여러 변수를 통합하여 관리 가능한 list 형식의 변수에 저장된다. 만들어진 각 list 들은 다시 또다른 list 변수에 저장되어 통합적으로 관리할 수 있도록 하였다. 병렬처리를 위해 한 개 년도에 해당하는 여러 변수를 동시에 읽어 처리하는 방식으로 구

현하였다.

메모리상에 저장된 격자자료를 사용하여 작물모델 기상입력자료를 생성하기 위해 `organizeWTH` 함수를 구현하였다. 이 함수에서는 변수별로 저장된 격자 자료로부터 관심 지점의 값을 `WTH` 파일의 형식에 맞는 테이블 형태로 재배열하는 기능을 수행한다. 이러한 변수값의 재배치 과정에 상당한 처리시간이 소요되므로, 관심지점들을 분할하여 처리할 수 있도록 병렬처리 기능을 적용하였다. 또한 `organizeWTH`는 지정된 작기 동안의 기상자료를 메모리에 저장하며, 이들 자료는 개별 지점별로 `list` 형식의 변수에 저장된다. 기상 자료와 함께 위도, 경도, 및 격자의 행과 열의 위치값을 해당 변수에 추가되어 자료의 활용이 용이하도록 구성하였다. 각 지점별로 정리된 `list` 변수들은 다시 전체 지역 단위로 `list` 변수에 저장되어 하나의 변수로 출력된다.

`writeWTH` 함수는 개별 지점별로 메모리에 저장된 기상자료들을 텍스트 형식의 파일에 출력하도록 구현하였다. 우선, `organizeWTH` 함수에서 생성된 복수의 지점들에 위치한 기상자료들이 매개변수의 형태로 `writeWTH` 함수에 전달된다. 또한, `writeWTH` 함수에서 `DSSAT` 모델과 호환되는 기상입력자료를 생성하기 위한 헤더가 작성되고 전달받은 지점별 기상자료들을 시간별로 파일에 출력한다. 텍스트 형식의 파일은 병렬처리를 통해 출력 하더라도 하드디스크의 처리속도에 영향을 받으므로 처리 시간의 큰 단축을 기대할 수 없어 `writeWTH` 함수에는 병렬처리 기능을 적용하지 않았다.



Fig. 1. The Nassi-schneiderman chart for `extract_point_AgMERRA.R` to prepare weather input files for the AgMERRA data files that have a gridded data format.

2.3. R script 기반의 DSSAT 모델 구동 도구 구현

여러 재배관리 시나리오 조건에서 병렬적으로 `DSSAT` 모델을 구동하기 위해 shell script 언어를 사용한 도구(`prepare_dssat.sh`)를 개발하였다(Fig. 2). `DSSAT` 모델은 파일 기반의 입력 자료를 바탕으로 구동되기 때문에, 같은 경로에서 복수의 `DSSAT` 모델이 구동될 경우 입력과 결과 자료들이 덮어쓰워질 수 있다. 동시에 `DSSAT` 모델을 구동하여 워크스테이션의 전산 자원을 최대한 활용하기 위해 재배관리 시나리오에 따라 폴더를 생성하고, `DSSAT` 모델구동을 위한 파일들을 작업 경로에 복사하는 기능을 구현하였다.

격자별, 연도별, 재배관리별로 `DSSAT` 모델 구동을 위한 R script (`shell_wrapper.R`)를 작성하였다(Fig. 2). 재배관리 설정 입력파일인 `xfile`의 경우, 하나의 파일에 여러 재배관리 조건을 저장할 수 있으나, 복수의 연도에 대한 처리에는 한계가 있다. 따라서, `xfile`에 포함되어 있는 모델 구동 연도를 기상자료의 시점과 동일하게 설정할 수 있도록 구현하였다. 또한, 구동 결과 중 주요 인자인 개화일, 성숙일, 수량 및 립수를 개별 격자별로 언어진 출력파일로부터 수집하여 연도별, 재배관리 시나리오별로 하나의 파일에 저장하는 기능을 추가하였다.

추가적으로 각 디렉토리별로 작성된 `DSSAT` 구동 스크립트들을 동시에 실행시킬 수 있도록 백그라운드에서 각 경로별 R script를 구동하기 위해 shell script (`dssat_bg.sh`)를 작성하였다. shell script로 명령어를

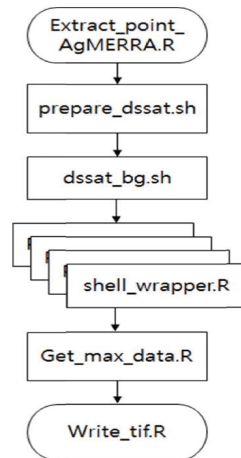


Fig. 2. A flow of the scripts to perform a gridded simulation of crop growth using the `DSSAT` model.

수행할 경우 기본적으로 하나의 명령어가 완료될 때까지 다음 명령이 수행되지 않는다. 이를 방지하기 위해, linux 운영체제에서는 명령어의 완료 여부와 관계없이 다음 명령어를 처리할 수 있는 백그라운드 처리 기능을 활용할 수 있다. 디렉토리별로 DSSAT 모델을 동시에 구동하기 위한 명령어들이 자동으로 생성되어 백그라운드에서 처리될 수 있도록 ‘&’ operator를 활용하였다.

2.4. 격자형 출력자료의 후처리 모듈의 구현

재배관리 시나리오에 따라 나뉘어 텍스트 형식으로 저장된 출력자료들을 모아 분석할 수 있는 후처리 모듈을 구현하였다. 후처리 모듈은 사용자의 분석 요구도에 따라 유연하게 활용될 수 있어야 하므로, 공통적으로 사용되는 파일의 출력값을 정리하는 모듈과, 연구 주제에 따라 특이적으로 사용될 분석 모듈로 구분하여 작성하였다. 자료 정리 모듈의 경우, 각 요소에 해당하는 값을 연도별로 지점과 재배관리를 행과 열로 정리하도록 구현하였다. 본 연구에서는 자료 분석 모듈을 사용하여 각 요소별로 정리된 자료로부터 격자별 최대 수량과 이에 해당하는 성숙군과 파종일을 파악할 수 있도록 구현하였다.

각 지점별로 저장된 자료를 GIS 형식의 지도 파일로 출력하는 후처리 R 스크립트를 작성하였다. 지도 형식의 출력 파일은 각 격자별로 분석된 결과의 공간 분포 파악에 용이하다. 각 지점별 결과값을 격자자료로 변환할 때, 각 격자가 가장 가까운 지점의 자료를 나타낼 수 있도록 nearest neighbor 방식의 보간법을 활용하였다. 이를 구현하기 위해 R의 gstat 라이브러리의 gstat 과 interpolate 함수를 사용하였다. 이를 통해 생성된 메모리상의 격자자료는 Tagged Image Format (TIF) 파일 형식으로 저장되도록 작성하였다.

III. 사례 연구

본 연구에서 최대 수량을 얻을 수 있는 품종들이 재배된다는 가정하에 개발된 스크립트들을 사용하여 공간적인 수량 예측을 하고 이에 따르는 품종과 파종일의 분포를 추정하는 과정을 자동화하였다. 공간적인 작물 생육 모의 수행을 위해 해당 지역에서 주로 사용되고 있는 재배방식을 반영할 수 있는 재배관리 설정이 사용되어야 한다. 그러나, 격자별로 정확히 어떠한 방식으로 재배관리가 수행된다는 정보는 알려지지 않

는다. 예를 들어, 미국 생산량의 14 %에 해당하는 콩이 생산되는 Illinois 주에서는 대개 위도에 따라 여러 종류의 성숙군으로 분류되는 품종들이 재배된다 (NASS, 2017). 그러나, 성숙군 3-5로 분류되는 이 지역 재배 품종들에 대한 공간적 분포를 확인할 수 있는 관측정보는 생산되지 않는다(Mourtzinis and Conley, 2017).

미국의 Illinois 주를 대상으로 DSSAT 모델의 기상 자료의 생성부터 최대 수량과 재배관리 자료 지도의 생산까지의 과정이 스크립트 언어들로 구현된 모듈들을 사용하여 자동으로 수행되었다. Illinois 주는 위도와 경도가 각각 36.9° - 42.4°와 -91.3° - -87.2° 범위 내에 위치한다. 이와 같은 공간적 범위에서 1981년부터 2005년까지의 기간 동안 연도별로 최고온도, 최저온도, 강우량, 일사량, 풍속 및 습도 자료를 가진 기상입력자료를 생산하고 이를 사용하여 DSSAT 모델을 구동하였다.

기상입력자료를 생성하기 위해 작물모형의 입력변수로 요구되는 기상자료들이 제공되는 AgMERRA 자료를 사용하였다. AgMERRA 자료는 일별 최저, 최고 온도, 강우량, 일사량, 상대습도, 풍속 등 개별 기상변수별로 1년간의 기상자료가 하나의 NetCDF 파일로 제공된다(<https://data.giss.nasa.gov/impacts/agmipcf/agmerra/>). 0.25°의 공간해상도를 가지는 AgMERRA 자료는 1980년부터 2010년까지의 30년 기간 동안의 자료가 사용될 수 있다. AgMERRA 격자자료를 사용하여 Illinois 주의 기상자료를 생성하기 위해 약 10km 단위로 나눈 모든 지점에 대한 위경도 자료를 준비하였다. Illinois 주에 해당하는 지역을 약 10 km단위로 나누어 모의자료를 준비할 경우 총 2109 개의 지점이 포함되는 것으로 나타났다. 생성된 지점별 위경도 값들은 LAT과 LON 열을 가지는 테이블 형태의 텍스트 파일에 저장되었다(Table 1).

Table 1. Table for specify coordination of Illinois

LAT	LON
268.7	36.9
268.7	37
268.7	37.1
...	...
272.8	42

병렬처리에 따른 처리 속도 증가효과를 비교하기 위해 병렬처리에 사용할 CPU 코어 수를 달리하여 기상입력자료 생성을 반복 수행하였다. 전산처리 용량에 따른 시간단축 효과를 분석하기 위해 최대 24개의 CPU 및 256GB의 메모리가 지원되는 워크스테이션 (KFT46-S8230G50-230, Tyan, Korea)을 사용하였다. 하나의 CPU 코어를 사용하였을 때 소요되는 시간을 기준으로 하여 2, 4, 8, 16 개의 CPU 코어가 사용되었을 때의 처리시간을 측정하였다. 구동 시간의 보다 정확한 측정을 위해 R 에서 제공되는 Sys.time 함수를 사용하여 R script의 시작과 끝 부분의 시간을 저장하고, 이를 바탕으로 사용된 CPU 별 스크립트의 구동 시간을 측정하였다.

연구대상 지역에서 사용될 수 있는 재배관리 방식에 대한 시나리오를 설정하고 모든 조건에 대해 작물 생육 모의를 수행하였다. 최대 수량을 가지는 성숙군 품종이 재배된다는 가정하에, 개별 격자에서 성숙군 1-5에 해당하는 5가지 품종 모수를 적용하여 각각의 품종별로 모의를 수행하였다. 해당 품종을 생물학적 및 물리학적 특성을 나타내기 위한 품종 모수들은 DSSAT 모델에서 제공되고 있는 표준 품종 모수를 사용하였다. 파종일의 경우, 콩의 일반적인 파종시기를 고려하기 위해 DOY 150일부터 10일 간격으로 5개의 파종일을 선정하였다. 따라서, 각 격자에서 년도별로 25번의 생육모의가 수행되었다. 품종과 파종일을 제외한 다른 재배관리 설정을 위해 미국에서 통상적으로 이루어지는 재배관리 설정값을 사용하였다. 예를 들어, 파종 2일 전 N-P-K를 30-30-34 의 비율로 시비하는 것으로 설정하였다. 관개의 경우, 강우에 의존하는 것으로 설정하였다. 토양 모수는 1km 단위의 잠재 수량을 모의하기 위하여 일반적으로 농업에 사용되는 사양토의 DSSAT 기본 제공 자료(IB00000008) 를 사용하였다.

작성된 스크립트들을 활용하여 maturity group 1-5에 해당하는 품종이 다섯 번에 걸쳐 파종되는 시나리오에 대한 수량이 모의되었다. 각 격자점에서 재배관리 시나리오별 작물 생육 모의를 25년간 구동하여 얻어진 625 개의 모의결과 출력 파일이 생성되었다. 후처리 모듈을 구동하여 각 연도별로 모의된 지점별 최대 수량 자료를 추출하였으며, 이에 해당하는 콩 성숙군과 파종일을 선정하고 raster 형식의 공간자료를 생성하였다.

IV. 결과 및 고찰

Illinois 주에 대한 기상 자료를 처리하여 1800 여개 지점에 대한 기상 입력자료를 생성하였다. 연구 대상 지역에 포함된 2109개의 격자들에는 미시간 호에 해당하는 격자들이 포함되어있어 AgMERRA의 자료가 제공되지 않는 250 지점이 존재하였다. 따라서, 이를 제외한 1,859 개의 지점에서 25년간의 자료를 생산하도록 R script를 구동하여 총 46,475개의 기상입력 자료를 생성하였다.

병렬처리에 CPU 수 증가에 따른 기상자료 생성을 위한 처리 시간 감소 효과는 이론적인 효과와 유사하게 나타났다. 하나의 CPU 코어를 사용할 경우 1,859 지점들에 대한 1년간의 기상자료를 생산하기 위해 7.38 시간이 소요되었다(Fig. 3). 그러나, CPU 코어 수의 증가함에 따라 구동시간은 상당히 단축되었다. 예를 들어, 16개의 CPU 코어를 사용하여 병렬처리한 경우 35분만에 모든 자료의 처리가 가능하였다. 특히, 이론적으로 얻을 수 있는 처리시간이 28 분이었던 점을 고려하면 실용적인 측면에서 처리시간의 큰 차이를 보이지 않았다. 또한, 작성된 스크립트들을 사용하여 모든 연도와 지점, 재배관리 시나리오에 대해 DSSAT을 구동하기 위해 linux 운영체제에서 제공하는 백그라운드 실행 기능을 사용하여 워크스테이션에 탑재된 CPU 코어들을 활용하여 DSSAT 모델을 동시에 구동할 수 있었다. 예를 들어, 5개의 코어를 활용하여 동시에 구동한 결과 25개의 재배관리에 대해 모든

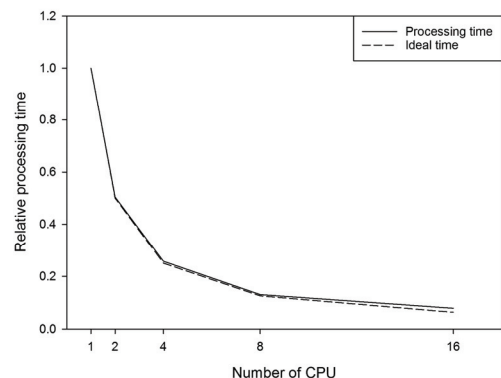


Fig. 3. The relative wall clock time that was taken to process gridded data using different numbers of CPU cores. The idea time is a theoretical wall clock time for a given number of CPU core.

격자에 대한 1년간 작물 생육 모델 구동을 위해 4.4 시간이 소요되었다.

본 연구에서 개발된 기상자료 생성 도구의 경우 기상입력자료 생성에 있어 병렬처리에 따른 처리 속도 증가 효과가 큰 것으로 나타났다. AgMERRA 자료로부터 전 세계 자료를 생산한다면 265,420개의 격자에 대하여 처리해야 한다. 만일 1개의 코어만을 사용한다면, 약 44일이 소요될 수 있으나, 16개의 코어가 사용된다면 3일만에 처리될 것이다. 이러한 병렬처리기능을 가지는 도구의 개발로 대용량으로 제공되는 고해상도의 격자형 기상자료들의 신속한 처리가 가능할 것으로 보인다. 특히, 최근 8개 이상의 CPU 코어가 탑재된 고성능 PC가 널리 보급되고 있다. 이러한 PC들을 사용할 경우, 고가의 워크스테이션을 사용하지 않더라도 공간적인 작물 생산성 예측이 가능할 것이다.

Illinois 주에서는 기존에 알려진 것과 같이 위도에 따라 성숙군 3-5가 최대 수량을 가지는 것으로 나타났다(Fig. 4A). 파종일의 경우 지역에 따라 140 일부터 180일까지 다양하게 나타났으며 남부 지역과 북부 지역의 파종일은 각각 5월 중순 및 6월 말 경인 것으로 모의되었다(Fig. 4B). 반면, 수량의 경우 972 - 3024 kg ha⁻¹의 범위에서 모의되었으며, 이는 최대 수량을 선택하였음에도 불구하고 낮은 수량이 모의된 지역이 분포하였다는 것을 의미하였다(Fig. 4C). 이러한 결과는 관개가 시행되지 않는 조건에 모의가 수행되었기

때문인 것으로 사료되었다.

본 연구를 통해 격자 자료인 AgMERRA를 사용하여 연구대상 지역에서 격자별 DSSAT 모델 구동을 지원할 수 있게 되었다. 특히, script 언어들로 구성된 시스템을 개발하여 다량의 자료가 처리되어야 하는 기상 입력파일 생성 및 시나리오 별 DSSAT 모델 구동 과정을 자동화할 수 있었으며 비교적 적은 노력으로 이를 수행할 수 있었다. 또한, 여러 조건에서의 작물 생육 모의를 병렬적으로 수행할 수 있어, 단일 프로세서를 사용하는 것 보다 빠른 처리가 가능하였다. 따라서, AgMERRA 자료와 같은 격자자료를 사용하여 일정 지역내 품종 모수 및 재배관리 시나리오에 대한 평가하는 것에 도움이 될 것으로 보인다. 예를 들어, 미국 지역에는 행정단위인 county 별로 작황 통계자료가 제공되고 있어, 작물 모형의 결과와 비교 및 검증을 통해, 작황 예측을 위한 재배관리 설정을 수행할 수 있다.

기후변화 영향평가를 위해 AgMERRA 자료 외에 지역 기후자료인 CORDEX 자료나 기상청 국가 표준 시나리오 자료 등의 격자형 자료를 사용하여 작물 생육 모의가 수행될 수 있다. 해당 자료들 역시 작물 모형의 구동에 필요한 변수들이 제공되고 있으며, 파일 명에 규칙이 있어, 각 자료에 대응되도록 모듈을 구성하기 용이하다. 특히, 이들 자료들은 지역에 대한 상세화된 기후 변화 시나리오가 제공되고 있어 본 연구에서 개발된 도구들을 사용하여 작물 모델을 사용한 기

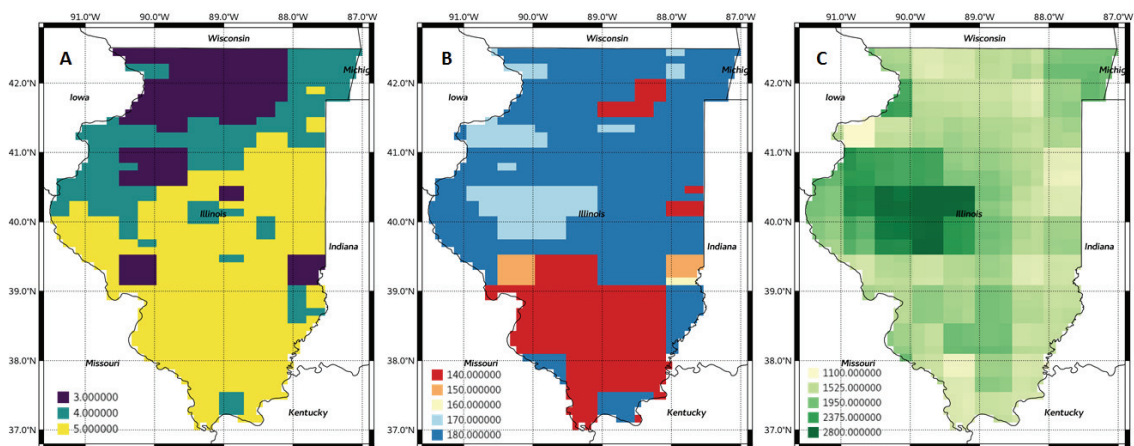


Fig. 4. Maps of maturity group for a variety (A) and planting date (B) that resulted in the maximum yield (C) in Illinois in 1991. The maximum yields were obtained using the DSSAT model with crop management scenarios that include five varieties and five planting dates.

후변화 영향 평가 연구를 지원할 수 있을 것이다.

본 연구에서 개발된 모듈들을 바탕으로 명령어 사용자 인터페이스 상으로 구동될 수 있는 다른 모델들을 지원하는 모듈이 개발될 수 있다. 예를 들어, 격자형 자료의 입력이 지원되지 않는 ORYZA 모델이나 격자형 기상자료가 지원되더라도 병렬처리가 지원되지 않는 DNDC 모델을 사용하여 대용량 자료 처리를 수행할 때 본 연구에서 개발된 시스템을 활용할 경우 처리시간을 상당히 단축시킬 수 있을 것으로 보인다 (Li *et al.*, 2017; Giltrap *et al.*, 2010). 개별 모델별로 모델 특이적인 기상 입력자료와 명령어 형식을 요구하므로 이와 호환되는 모듈 개발이 필요하다. 반면, 국내에서 작물이나 토양의 반응을 모의하기 위해 사용되는 모델들은 비교적 적기 때문에 이들 모델들을 대상으로 모듈들을 추가하는 것이 비교적 용이할 것이다.

여러 작물모델을 지원하는 시스템 개발을 통해, 기후변화 영향 평가의 불확도를 줄일 수 있는 앙상블 연구를 지원할 수 있을 것으로 보인다. 단일 작물 모형이나 단일 기후변화 자료를 사용한 기후변화 영향 평가에는 불확실성이 따른다(Olsen *et al.*, 2007). 이에 따라 불확실성을 줄이기 위해 여러 모형으로부터 생산된 기후 자료들을 활용하거나 다양한 작물 모형의 결과를 종합하는 앙상블 방식이 사용되고 있다(Palmer *et al.* 2004; Rosenzweig *et al.*, 2014). 본 시스템은 비교적 쉽게 여러 작물 모형과 기상자료의 지원이 가능하도록 개발되어 있기 때문에, 일정 지역에 대한 다양한 기상자료 및 작물 모형을 이용한 앙상블 연구를 수행하는데 도움이 될 수 있을 것이다.

적 요

작물 생산량의 시·공간적 분석은 정책입안자와 이해관계자들에게 중요한 정보를 제공할 수 있으나, 이를 위해서는 공간적 기상자료를 처리하고 이에 맞추어 작물 모형을 구동할 수 있는 작업들이 필요하다. 이에 따라 DSSAT (Decision Support System for Agrotechnology Transfer)을 기반으로 지역내 작물 생산성 자료를 생산할 수 있는 자동화 시스템을 개발하고자 하였다. 이 시스템은 전문적인 컴퓨터 프로그래머가 아니더라도 사용가능한 R과 shell script를 기반으로 개발되었다. 먼저, 격자형 기상자료의 각 격자에 해당하는 정보를 텍스트 형식의 기상 입력자료 형식으로 변환하는 기능을 가지는 모듈을 작성하였다. 다음

으로 R 패키지를 사용하여 GIS자료 처리와 병렬 처리 기능이 구현된 R script을 작성하였다. 또한, 작물 모델을 자동으로 구동하는 기능을 shell script를 사용하여 구현하였다. 사례 연구로, 미국 Illinois 주에서 콩의 최대 수량을 얻을 수 있는 재배관리 조건의 공간적인 분포를 파악하고자 하였다. 개발된 도구를 통해 AgMERRA 자료로부터 Illinois 주의 1981 - 2005년까지의 기상입력자료를 생산하였다. 해당 지역에서 1개의 CPU 코어를 사용하여 1년간의 자료를 처리하기 위해 7.38 시간이 걸렸으나, 병렬처리를 통해 16개의 CPU 코어를 사용하였을 때 처리 시간이 크게 줄어, 35분만에 처리가 가능하였다. 이렇게 생산된 기상 입력자료들을 작물 모형 자동 구동 시스템에 활용하여 해당 지역에서의 최대 수량과, 최대 수량을 가지는 성숙군 및 파종일 지도를 작성할 수 있었다. 특히, 본 연구에서 개발된 도구는 DSSAT 모델뿐만 아니라 국내에서 사용되는 다른 작물모델들에게 적용될 수 있어 공간적 작물 생산성 평가에 도움을 줄 수 있을 것으로 보인다.

감사의 글

본 논문은 농촌진흥청 공동연구사업 (과제번호: PJ010115022017)의 지원에 의해 이루어진 것임.

REFERENCES

- Elliott, J., D. Deryng, C. Müller, K. Frieler, M. Konzmann, D. Gerten, M. Glotter, M. Flörke, Y. Wada, N. Best, S. Eisner, B. M. Fekete, C. Folberth, I. Foster, S. N. Gosling, I. Haddeland, N. Khabarov, F. Ludwig, Y. Masaki, S. Olin, C. Rosenzweig, A. C. Ruane, Y. Satoh, E. Schmid, T. Stacke, Q. Tang, and D. Wisser, 2014: Constraints and potentials of future irrigation water availability on agricultural production under climate change. *Proceedings of the National Academy of Sciences* **111**(9), 3239-3244.
- Giltrap, D. L., C. Li, and S. Saggat, 2010: DNDC: A process-based model of greenhouse gas fluxes from agricultural soils. *Agriculture, Ecosystems & Environment* **136**, 292-300.
- Hijmans, R. J., and J. van Etten, 2014: raster: Geographic data analysis and modeling. *R package version 2*, 15.
- Hoogenboom, G., J. Jones, P. Wilkens, C. Porter, K.

- Boote, L. Hunt, U. Singh, J. Lizaso, J. White, and O. Uryasev, 2015: Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.6. 1.0 (www.DSSAT.net) DSSAT Foundation, Prosser, Washington, USA.
- Kim, H.-Y., J. Ko, S. Jeong, J.-H. Kim, and B. Lee, 2017: Geospatial delineation of South Korea for adjusted barley cultivation under changing climate. *Journal of Crop Science and Biotechnology* **20**(5), 417-427.
- Kim, K., M. kang, H. Jeong, and J. Kim, 2013: Comparison of crop growth and evapotranspiration simulations between Noah multi physics model and CERES-Rice model. *Korean Journal of Agricultural and Forest Meteorology* **15**(4), 282-290.
- Lee, C.-K., K.-S. Kwak, J.-H. Kim, J.-Y. Son, and W.-H. Yang, 2011: Impacts of climate change and follow-up cropping season shift on growing period and temperature in different rice maturity types. *Korean Journal of Crop Science* **56**(3), 233-243.
- Mourtzinis, S., and S. P. Conley, 2017: Nass (National Agriculture Statistic Service). 2017: Quick stats (<https://quickstats.nass.usda.gov>) Delineating soybean maturity groups across the United States. *Agronomy Journal* **109**, 1397pp.
- NASS (National Agriculture Statistic Service), 2017: Quick stats (<https://quickstats.nass.usda.gov>)
- Olesen, J. E., T. R. Carter, C. H. Díaz-Ambrona, S. Fronzek, T. Heidmann, T. Hickler, T. Holt, M. I. Minguéz, P. Morales, J. P. Palutikof, M. Quemada, M. Ruiz-Ramos, G. H. Rubæk, F. Sau, B. Smith, and M. T. Sykes, 2007: Uncertainties in projected impacts of climate change on European agriculture and terrestrial ecosystems based on scenarios from regional climate models. *Climatic Change* **81**, 123-143.
- Palmer, T. N., A. Alessandri, U. Andersen, P. Cantelaube, M. Davey, P. Délécluse, M. Déqué, E. Diez, F. J. Doblas-Reyes, H. Feddersen, R. Graham, S. Gualdi, J. F. Guérémy, R. Hagedorn, M. Hoshen, N. Keenlyside, M. Latif, A. Lazar, E. Maisonnave, V. Marletto, A. P. Morse, B. Orfila, P. Rogel, J. M. Terres, and M. C. Thomson, 2004: Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (Demeter). *Bulletin of the American Meteorological Society* **85**, 853-872.
- Analytics, R., and S. Weston, 2014: doSNOW: Foreach parallel adaptor for the snow package. *R Package version 1*, 12.
- Rosenzweig, C., J. Elliott, D. Deryng, A. C. Ruane, C. Müller, A. Arneth, K. J. Boote, C. Folberth, M. Glotter, N. Khabarov, K. Neumann, F. Piontek, T. A. M. Pugh, E. Schmid, E. Stehfest, H. Yang, and J. W. Jones, 2014: Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the National Academy of Sciences* **111**(9), 3268-3273.
- Ruane, A. C., R. Goldberg, and J. Chrystanthacopoulos, 2015: Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. *Agricultural and Forest Meteorology* **200**, 233-248.
- Li, T., O. Angeles, M. Marcaida III, E. Manalo, M. P. Manalili, A. Radanielson, and S. Mohanty, 2017: From ORYZA2000 to ORYZA (v3): An improved simulation model for rice in drought and nitrogen-deficient environments. *Agricultural and Forest Meteorology* **237-238**, 246-256
- Yoo, B. H., and K. S. Kim, 2017: Development of a gridded climate data tool for the COordinated Regional climate Downscaling EXperiment data. *Computers and Electronics in Agriculture* **133**, 128-140.
- Zender, C. S., 2008: Analysis of self-describing gridded geoscience data with netCDF Operators (NCO). *Environmental Modelling & Software* **23**, 1338-1342.