

작물모형 평가를 위한 통계적 방법들에 대한 비교

김준환* · 이충근 · 손지영 · 최경진 · 윤영환

농촌진흥청 국립식량과학원 답작과

(2012년 11월 8일 접수; 2012년 12월 7일 수정; 2012년 12월 10일 수락)

Comparison of Statistic Methods for Evaluating Crop Model Performance

Junhwan Kim*, Chung-Kuen Lee, Jiyoung Shon, Kyung-Jin Choi and Younghwan Yoon

Rice Research Division, National Institute of Crop Science, RDA

(Received November 8, 2012; Revised December 7, 2012; Accepted December 10, 2012)

ABSTRACT

The objective of this short communication is to introduce several evaluation methods to crop model users because the evaluation of crop model performance is an important step to develop or select crop model. In this paper, mean error, mean absolute error, index of agreement, root mean square error, efficiency of model, accuracy factor and bias factor were explained and compared in terms of dimension and observed number. Efficiency of model and index of agreement are dimensionless and independent of number of observation. Relative root mean square, accuracy factor and bias factor are dimensionless and not independent of number of observation. Mean error and mean absolute error are affected by dimension and number of observation.

Key words: Crop model evaluation, Index of agreement, Root mean square error, Model efficiency, Accuracy, Bias

I. 서 론

토양 및 기상환경과 품종에 따른 작물 생육 양상을 예측하기 위해 작물모형이 사용되어 왔다(Lee, 2008). 작물모형은 그 개발 역사가 오래되었음에도 불구하고 국내에서는 큰 관심을 끌지는 못하였다. 그러나, 최근 기후변화에 따른 미래의 작물 생산성 변화 평가와 작목의 적지 및 적기 변동을 예측하기 위해 작물모형을 적극 활용하고 있으며(Shin and Lee, 1995; Chung *et al.*, 2006), 이와 더불어 작목별 작황예측을 위한 도구로서 그 쓰임이 확대되고 있다. 이러한 추세에 따라 작물모형 개발과 활용을 위한 연구도 예전과 달리 활발한 상황이다.

작물 모형 개발과 활용을 위해서는 작물 모형에 대한 정확한 평가가 필수적이다. 그러나 작물학을 기반으로 한 연구자들은 소수의 수식과 지표만을 사용하여 작물모형의 정확도를 평가해 왔다. 예를 들어, Lee (2008)는 root mean square error (RMSE)를 사용하여 작물모형으로부터 얻어진 건물중과 수량의 예측값과 실측값을 비교분석 하였다.

그러나, 작물모형을 대상으로 한 지표 뿐만 아니라 다양한 분야에서 사용되어온 지표들이 작물모형의 정확도를 평가하기 위해 사용될 수 있다. 예를 들어, Nash and Sutcliffe(1970)이 제안한 efficiency of model은 수문학 분야에서 사용되어 왔으나 작물모형의 정확도 분석을 위해서는 많이 사용되어오지 않았다. 그런데, 사용되는 지



* Corresponding Author : Junhwan Kim
(sfumato@korea.kr)

표에 따라 작물모형의 정확도가 다른 양상으로 분석될 수 있어 각 지표간의 장단점을 파악하는 것이 중요하다.

이번 연구에서는 작물모형 평가에 보편적으로 사용되는 평가지표와 더불어 다른 분야에서 사용되는 정확도에 관련된 지표를 사용하여 작물모형의 예측값에 대한 정확도 분석을 소개하고자 한다. 특히, 각 지표들을 dimension 유무에 따라 분류하고 각 지표군들의 특성 분석을 통해 향후 작물모형을 이용한 예측값들 분석에 어떻게 활용될 수 있는지를 제시하였다. 또한, 실제 사례를 통한 지표들의 장단점 분석을 위해 Lee (2008)이 보고한 건물중과 수량의 작물모형 예측값들을 중심으로 각 지표간의 차이를 분석하였다.

II. 적용모형 및 평가 지표

2.1. 적용모형

Lee(2008)는 Oryza2000을 사용하여 특정 벼 품종

의 건물중과 수량을 예측하였다. Oryza2000은 국제미 작연구소와 Wageningen 대학이 공동개발한 벼 생육모형이다(Bouman and Van Laar, 2006). 이 모형의 구동을 위해서는 발육단계, 비염면적 및 건물분배계수 등의 입력자료가 필요한데 이 자료는 품종에 따라 변화되기 때문에 모형에 내재된 default값을 바로 사용할 수 없는 경우가 많다. 포장실험을 통해 새로 얻어진 품종모수를 입력하여 계산된 건물중 및 수량 예측값들은 기본값(default)으로 주어진 품종모수를 이용한 경우보다 더 정확할 것이다.

이러한 정확도의 차이를 수치적으로 분석하기 위해 본 보에서는 Lee(2008)가 수행한 이앙기, 시비량, 재식밀도 처리가 조합되어 다양한 조건이 고려된 포장실험 결과를 모형검증을 위해 사용하였으며 품종모수가 고려되어 예측된 결과와 기본값(default)을 이용하여 예측된 결과를 각각 비교하였다. 품종모수 도출을 위한 포장실험은 수원에 소재한 국립식량과학원 답작포

Table 1. Abbreviation and formula of evaluation indices

Abbreviation	Index	Formula
ME	Mean Error	$ME = \frac{\sum_1^n (O_i - S_i)}{n}$
MAE	Mean Absolute Error	$MAE = \frac{\sum_1^n O_i - S_i }{n}$
RMSE	Root Mean Square Error	$RMSE = \sqrt{\frac{\sum_1^n (O_i - S_i)^2}{n}}$
RRMSE	Relative root mean square error	$RRMSE = \frac{100}{\bar{O}} \times \sqrt{\frac{\sum_1^n (O_i - S_i)^2}{n}}$
EF	Model Efficiency	$EF = 1 - \frac{\sum_1^n (O_i - S_i)^2}{\sum_1^n (O_i - \bar{O})^2}$
d	Index of Agreement	$d = 1 - \frac{\sum_1^n (O_i - S_i)^2}{\sum_1^n (S_i - \bar{O} + O_i - \bar{O})^2}$
Bf	Bias factor	$Bf = 10^{\left\{ \frac{1}{n} \sum_1^n \log(S_i/O_i) \right\}}$
Af	Accuracy factor	$Af = 10^{\left\{ \frac{1}{n} \sum_1^n \log(S_i/O_i) \right\}}$

*O : observed data, S: predicted data, \bar{O} : average of observed data, n : number of observation

장에서 1999년(미출판)수행된 자료, Lee *et al.*(2004b)에 의해 수행된 자료, 서울대학교 부속 실험농장에서 수행된 Cui and Lee(2002) 자료 등을 사용하였다. 또한 생육단계 추정을 위해 phytotron에서 수행된 Choi *et al.*(2002)과 Lee *et al.*(2004a)의 자료가 사용되었다.

2.2. 정확도 평가 지표

본 보에서 소개할 평가 지표와 각 지표의 약어는 Table 1과 같으며 이후 모든 약어는 Table 1의 기준

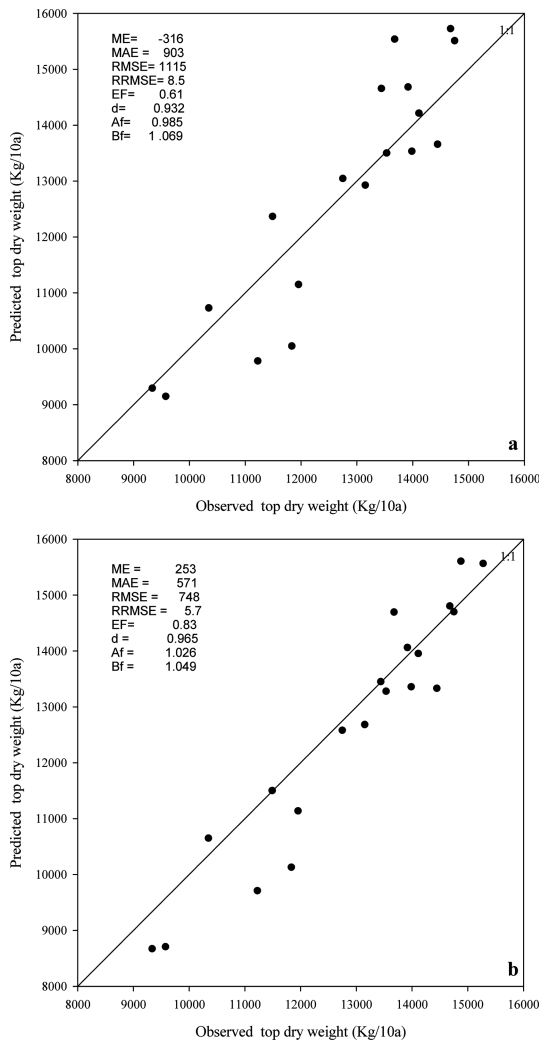


Fig. 1. Scatter plot of observed vs. predicted top dry weight. Predicted dry weight (a) was calculated by Oryza 2000 with default input data, Predicted dry weight (b) was calculated by Oryza 2000 with modified input data. Evaluation indices including ME, MAE, RMSE, RRMSE, EF, d, Bf and Af are described in Table 1.

에 따름을 밝혀둔다.

2.2.1. Mean Error (ME)

일반적으로 모형을 평가할 때는 실측값에 대한 예측값의 대응 그래프에 앞으로 설명하고자 하는 지표들을 병기하여 이루어진다. 예를 들어 Fig. 1과 같은 형태이다.

실측값에 대한 1:1대응선을 중심으로 예측값이 어떻게 분포하였는지를 평가하는데 모형에 대한 평가는 결국 모형에 의해 얻어진 결과가 실측치에 얼마나 근접하였는지를 판단하는 것으로 결국 정확도에 대한 판단이다. 따라서, 직관적으로 식 (1)을 생각할 수 있다.

$$ME = \frac{\sum_{i=1}^n (O_i - S_i)}{n} \quad (1)$$

식 (1)은 실측값인 O와 모형예측값인 S와의 차이를 의미하며 mean error 또는 mean bias라고 한다. 각 대응 값에 대한 차이를 평균한 것 이기 때문에 0에 가까울수록 실측값을 중심으로 편이가 없음을 의미하고 반대로 - 또는 + 방향으로 값이 커질수록 실측값에 대해 과대 또는 과소평가가 되었음을 파악할 수 있다.

2.2.2. Mean Absolute Error (MAE)

ME의 경우 각 대응하는 실측값과 예측값의 차의 총합이기 때문에 양수 또는 음수로 나온 값들이 상쇄되어 실제로 예측값이 실측값과 어느정도 차이가 발생했는지를 파악할 수는 없다. 이를 보완하기 위해 식 (2)를 생각할 수 있다.

$$MAE = \frac{\sum_{i=1}^n |O_i - S_i|}{n} \quad (2)$$

식 (2)는 mean absolute error라고 하며 문자 그대로 오차 절대값의 합이다. 따라서 실측값과 예측값간에 발생한 절대값 차이의 총합으로서 모형이 얼마나 실측값에 근접하였는지를 설명해줄 수 있다.

2.2.3. Root Mean Square Error (RMSE)

식 (1)과 (2)를 활용하면 모형을 평가하는 데는 이론적으로 문제가 없어 보인다. 그러나 이들 수식을 활용할 때 관측수 n에 의해 ME와 MAE 두 수치가 크게 변할 수 있다. 만일 ME계산에서 실측값과 예측값간의 차의 합이 동일한데 n의 수가 크다면 당연히

ME의 값도 줄어든다. 즉 관측수에 따라 모형에 대한 잘못된 판단을 내릴 수 있다. 만일 두 개의 모형을 동일한 관측수로 비교한다면 이러한 문제가 없지만 한 개의 모형에 대한 평가를 할 시에는 주의가 필요하다. 식 (1)과 (2)는 작물모형관련 논문에서 널리 쓰이기는 하나 더 보편적으로 쓰이는 것은 다음의 식 (3)이다.

$$RMSE = \sqrt{\frac{\sum_1^n (O_i - S_i)^2}{n}} \quad (3)$$

식 (3)은 root mean square error로 일반적으로 알고 있는 표준편차와 유사하여 직관적으로도 이해하기 쉽다. RMSE는 실측값과 예측값간의 차이에 의해 RMSE값이 크게 변동되고 관측수가 커지면서 발생하는 영향력은 상대적으로 축소된다. 따라서 식 (2)에서 발생한 문제가 일정부분 상쇄된다. 그러나 RMSE는 관측수의 영향력이 여전히 존재한다.

2.2.4. Relative Root Mean Square Error (RRMSE)/ Normalized Root Mean Square Error (NRMSE)

지금까지 열거한 방법들이 공통적으로 안고 있는 또 다른 문제는 dimension의 문제이다. 이것은 모형이나 이 수식들에서 기인한 문제가 아닌 단순한 문제이지만 그 결과는 크다. Table 2을 보면 단순히 단위만을 바꿨을 뿐인데 마치 t ha⁻¹ 단위에서는 더 정확한 결과가 나온 것으로 착각할 수 있다. 이를 보완하기 위해 RMSE에서 dimension을 제거할 필요가 있는데 이와 관련해서는 연구자별로 동일한 수식에 대해 다른 이름을 붙인다던가 이름은 같은데 수식이 다르다던가 하는 혼돈이 있어 정리가 필요하다. 단위를 제거하기 위해서는 식 (4)와 식 (5)을 생각할 수 있다.

Table 2. Comparison of evaluation indices in different units. This indices were calculated using data of Fig 2b

	Yield (Kg 10a ⁻¹)	Yield (t ha ⁻¹)
ME	535	0.535
MAE	698	0.698
RMSE	825	0.825
EF	0.27	0.27
d	0.866	0.866
Bf	0.91	0.91
Af	1.12	1.12
RRMSE	12.1	12.1

$$RRMSE \text{ or } NRMSE = \frac{100}{O} \times \sqrt{\frac{\sum_1^n (O_i - S_i)^2}{n}} \quad (4)$$

$$RRMSE = \sqrt{\frac{\sum_1^n \left(\frac{O_i - S_i}{\max(O) - \min(O)} \right)^2}{n}} \quad (5)$$

식 (4)는 연구자에 따라 relative root mean square error (Rinaldi *et al.*, 2003; Loague and Green, 1991) 또는 normalized root mean square error (Bouman and van Laar, 2006)로 표현하였다. 식 (4)는 실측값 평균(\bar{O})으로 RMSE를 나눈 것으로 변이계수와 사실상 동일한 개념이다. Dimension이 없어지고 실측과 예측간의 차이가 표준화되는 특징이 있어 단위와 숫자의 크기에서 기인한 판단오류를 방지할 수 있다(Table 2). 이와는 별도로 식 (5)역시 relative root mean square error로 쓰이는 경우가 있다(Nicolardo *et al.*, 2001). 식 (5)는 식 (4)와는 달리 평균이 아닌 실측의 최고값과 최저값을 이용하여 표준화한 것이라고 생각할 수 있다. 보통 작물모형에서 흔히 볼 수 있는 형태는 식(4)로 relative RMSE (RRMSE)로 쓸 수 있다.

2.2.5. Model Efficiency (EF)

RRMSE가 어느 정도 수치가 나와야 정확하거나 잘 된 모형이라고 할 수 있는지를 파악할 필요가 있다. 우선 실측값의 변이계수보다는 작아야 할 것이다. 만일 예측값의 RRMSE가 실측값의 변이계수보다 크다면 이 모형을 활용한 예측은 사실상 평균값을 활용하는 것보다 예측력이 떨어지며 모형은 무의미해진다. 이를 수치화한 것이 식 (6)이다.

$$EF = 1 - \frac{\sum_1^n (O_i - S_i)^2}{\sum_1^n (O_i - \bar{O})^2} \quad (6)$$

식 (6)은 efficiency of model이라고 하며 실측값의 지능함으로 실측과 예측간의 차이의 지능함을 나눈 것을 기반으로 하며 Nash and Sutcliffe(1970)에 의해 처음 제안하였다. 이 식은 결국 RRMSE를 실측값의 변이계수로 나눈 후 제곱하여 1에서 빼준 것이다. RRMSE가 실측값의 변이계수보다 같아지거나 커질수록 EF는 0또는 음의 값을 보이며 반대로 RRMSE가 관측값 변이계수보다 작다면 1에 가까운 값을 나타내게 된다. 따라서 0보다 큰 값을 보이면 일단 받아들일 만한 모형으로 판단할 수 있다. 이 지표 역시

dimension에 관계없이 일정하다(Table 2).

2.2.6. Index of Agreement

위의 지수들은 결국 실측과 예측의 편차가 어느 정도 발생하는지에 대한 것들로 모형 예측값의 방향성을 표시하지는 못한다. 즉 관측값이 크기가 증가함에 따라 예측값도 그에 대응하여 증가되는지 또는 감소하는지를 판단하지는 못한다. 이를 파악하기 위해 Willmott (1981, 1982)에 의해 다음의 식 (7)이 제안되었다.

$$d = 1 - \frac{\sum_1^n (O_i - S_i)^2}{\sum_1^n (|S_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad (7)$$

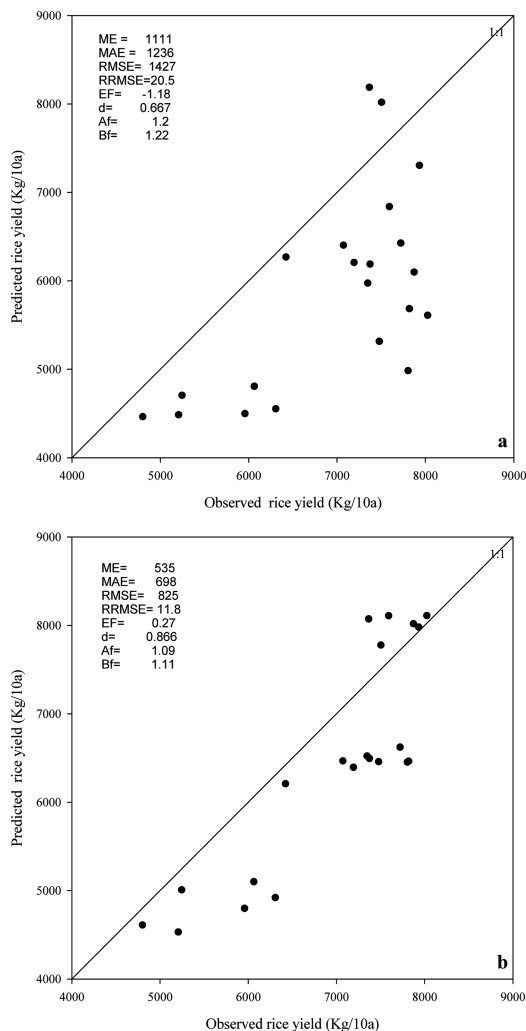


Fig. 2. Scatter plot of observed vs. predicted yield. Predicted yield (a) was calculated by Oryza 2000 with default input data, Predicted yield (b) was calculated by Oryza 2000 with modified input data. Evaluation indices including ME, MAE, RMSE, RRMSE, EF, d, Bf and Af are described in Table 1.

식 (7)은 Index of agreement라고 하며 1에 가까울수록 실측과 예측간의 일치성이 큰것이고 반대로 0에 가까우면 일치도가 떨어지는 것이다. 예를 들어 Fig. 2a와 Fig. 2b를 비교하면 Fig. 2b는 실측값의 크기가 커짐에 따라 예측값의 크기도 비례해서 증가하는 반면 Fig. 2a에서는 일정 수치이상에서 예측값이 일정해진다. 따라서 Fig. 2b의 d값이 더 높음을 알 수 있다. 즉 d값은 단순한 편차가 아닌 편차의 추세를 반영할 수 있다. 이러한 추세편차는 곧 모형에 의한 계통오차라고 할 수 있다(Willmott, 1982). 지금까지 소개한 지표들은 그 개발 기원이 수리학 또는 기상학 분야의 모형에서 기인한 것들로서 대부분 실측과 예측간의 차이에 대한 평균 자승합의 개념을 기반으로 한다.

2.2.7. Bias factor and Accuracy factor

이와는 별도로 최근 미생물의 성장예측분야에서 사용하는 지표들이 있는데 작물모형분야에서는 알려지지 않아 여기서 소개하고자 한다. 미생물의 성장예측 모형의 예측 지표는 기본적으로 식 (1)과 (2)에서 기인한다고 할 수 있는데 이는 Ross(1996)가 제안한 것으로서 식은 다음과 같다.

$$Bf = 10^{\left\{ \frac{1}{n} \sum_1^n \log(S_i/O_i) \right\}} \quad (8)$$

식 (8)은 결국 식 (1)의 변형인데 실측과 예측에 log 변환하여 ME를 구한 후 이 값을 10의 거듭제곱 값에 넣어 환원하는 방식이다. ME가 의미하는 비와 마찬가지로 미생물 분야에서는 이 값을 Bias factor라고 지칭한다. 1이면 편차가 없는 것이고 1보다 작거나 크면 과소 또는 과대 평가되었다고 판단한다. 이 식에서는 dimension이 제거된다.

$$Af = 10^{\left\{ \frac{1}{n} \sum_1^n |\log(S_i/O_i)| \right\}} \quad (9)$$

식 (9) 역시 식 (8)과 같은 절차로 구해지는 변형된 식 (2)라고 할 수 있으며 accuracy factor라고 한다. MAE와 뜻은 같으며 1에 가까운 값이 나올수록 정확한 것이고 멀어질수록 부정확한 모델임을 의미한다. Af 역시 dimension이 제거된다. Bf와 Af역시 관측수에 의한 영향에서 역시 자유롭지 못하는데 다른 지표들과는 달리 일정수준의 관측수 이상에서는 값이 일정해지는 성질이 있어 관측수의 영향력이 감소된다(Fig. 3).

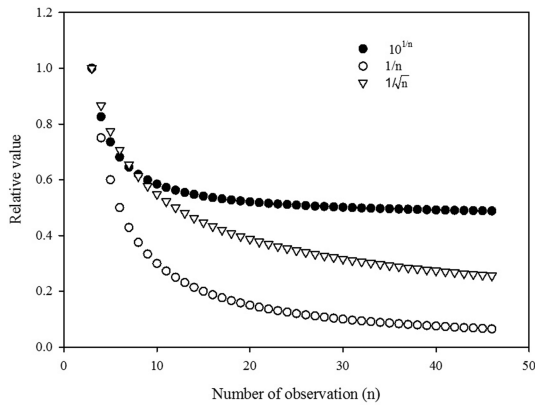


Fig. 3. Relative value of number of observation for $10^{1/n}$, $1/n$ and $1/\sqrt{n}$. Relative value is $F(n)/F(3)$. $F(n)$ is function of $10^{1/n}$ or $1/n$ or $1/\sqrt{n}$. n is number of observation.

이 때문에 관측수가 큰 경우에는 오히려 평가에 문제가 발생 할 수 있어 주의가 필요하다. 하지만 이 방법은 ME와 MAE에서 dimension을 제거하여 표준화된 수치를 제공해줄 수 있어 dimension에 따른 판단요류를 줄여줄 수 있다.

III. 평가지표 비교와 활용

지금까지 설명한 평가 지표들의 특성을 요약하면 Table 3과 같다. dimension과 관측수(n) 모두에 대해 비교적 자유로운 지표는 efficiency of model (EF)와 index of agreement (d)이다. dimension에 대해서만 자유로운 것은 relative root mean square error (RRMSE), bias factor (Bf)와 accuracy factor (Af)이다. 그 외의 지표들은 dimension과 n에 의해 값의 크기가 결정되기 때문에 주의할 필요가 있다. 그러나ME와

MAE, RMSE는 평균적인 오차범위를 확인할 수 있기 때문에 의미가 없는 것은 아니다.

모형을 검정할 때 어떤 지표를 사용해야만 한다고 지정할 수는 없다. 각각의 목적에 따라 지표의 특성을 이해하여 취사선택을 하여 종합적으로 판단하면 된다. 또한 이 지표들이 어떤 값 이상이어야지 잘 맞는 모형 인지는 명확하게 말 할 수는 없다. 다만 목적과 모형에 따라 그 허용범위를 사전에 결정할 수 있는데 Bellocchi *et al.*(2002)은 일사모형에서 Fuzzy 시스템에서 활용하기 위해 Favorable (RMMSE \leq 20)과 Unfavorable (RMSE \geq 40) 이 두 기준을 전문가 판단에 따라 설정하였고, Ross *et al.*(2000)은 성장속도 예측시 Bf가 0.9~1.05의 범위에 있는 것이 적절하다고 설정하였다. 이렇듯 각 해당지표는 그 모형의 목표하는 오차범위에 따라 판단하여 결정할 수 있다.

이러한 지표를 활용하는 표준적인 순서는 존재하지 않지만 본 논문에서 제안하는 것은 우선 EF값으로 실측치의 변이보다 모형의 변이가 높은 지 확인하되 만약 음수가 나올 경우 이 모형에 대해서는 재검토가 필요하다. 예를 들어 Fig. 2a 수량 예측의 EF값이 음수로 판정되었다. 이는 default 입력자료를 이용한 수량예측은 무의미하며 상당한 부분에서 개선이 요구된다고 할 수 있다. 예시의 경우에는 입력자료를 변경하여 예측능력을 개선할 수 있었다. 그러나 경우에 따라서 작물모형에 사용되는 일부 함수 또는 모듈 전체를 검토하고 수정해야 할 상황도 발생할 수 있다.

EF확인 후 d값을 구하여 모형이 실측값 변동을 충실히 반영하는지 확인한다. 대부분 모형에 큰 문제가 없다면 이 수치는 0.9내외 또는 1에 가까운 값을 보인다. Fig. 1과 2를 비교하여 보면 Fig. 1의 경우 default

Table 3. Summary for characteristics of evaluation indices

Index*	Limit of value		Best case	Dimension	Dependency on number of observation
	Upper	Lower			
ME	$+\infty$	$-\infty$	0	yes	yes
MAE	$+\infty$	0	0	yes	yes
RMSE	$+\infty$	0	0	yes	yes
RRMSE	100	0	0	no	yes
EF	1	$-\infty$	1	no	no
d	1	0	1	no	no
Bf	$+\infty$	>0	1	no	yes
Af	$+\infty$	1	1	no	yes

*Description of evaluation index is listed in Table 1.

와 modified모두 0.9이상의 값을 보이는 반면 Fig. 2에서는 0.9미만인 값을 보여준다. 즉 이 모형은 건물중에 대한 예측에는 뛰어난 성능을 보이지만 수량에 대해서는 상대적으로 문제가 있다고 판단할 수 있다. 이 후에는 분포되어 있는 형태를 표현해주는 지표들인 RMSE 등(당연히 EF값을 포함하여)을 사전에 결정된 목표 정확도를 결정하여 검토한다. 만일 목표 정확도가 RRMSE 10%미만이었던 건물중의 경우 default 값을 사용해도 큰 문제가 없다고 판단할 수 있다. 하지만 관측수량의 변이보다 모형의 오차가 적은 것을 원할 경우 EF값이 1에 가까운 것을 선택해야 한다. 이 경우에는 입력변수를 변경한 Fig. 2가 해당되며 건물중 수량은 modified된 입력변수를 활용하면 상당부분 예측할 수 있음을 의미한다.

이와 같이 지금까지 설명한 지표를 활용하면 모형의 정확도를 손쉽게 검토할 수 있으며 모형간 비교 또는 모형 내 출력값 간의 정확도 역시 비교 가능하다. 그러나 각 지표를 활용할 때에는 그 것의 성격과 의미를 잘 파악하여 사용해야 한다.

적 요

작물모형 평가에 사용되거나 사용할 수 있는 9가지 지표를 소개하였으며 이들의 특징은 다음과 같다. efficiency of model (EF)와 index of agreement (d)은 dimension이 없고 관측수(n)에 의존적이지 않았으며, dimension에 대해서만 자유로운 것은 relative root mean square error (RRMSE), bias factor (Bf)와 accuracy factor (Af)이다. Root mean square error, mean absolute error들은 관측수와 dimension에 영향을 받기 때문에 판단 시 주의가 필요하다. 따라서 이들의 특징을 파악하여 목적에 맞게 모형의 성능을 파악하여야 한다.

감사의 글

본 논문은 농촌진흥청 연구사업(ATIS 주관과제번호: PJ008679)의 지원에 의해 이루어진 것임.

REFERENCES

Bellocchi, G., M. Acutis, G. Fila, and Donatelli, Marcello,

2002: An Indicator of Solar Radiation Model Performance based on a Fuzzy Expert System. *Agronomy Journal* **94**, 1222-1233.

Bowman, B. and H. van Laar, 2006: Description and evaluation of the rice growth model ORYZA2000 under nitrogen-limited conditions. *Agricultural Systems* **87**, 249-273.

Choi, K., J. Lee, N. Chung, and W. Yang, 2002: The effect of temperature and day-length on the heading of rice cultivars. *Treatises of Crop Researches* **3**, 163-170.

Chung, U., K. Cho, and B. Lee, 2006: Evaluation of site-specific potential for rice production in Korea under the changing climate. *Korean Journal of Agricultural and Forest Meteorology* **8**, 229-241. (in Korean with English abstract)

Cui, R. X. and B. W. Lee, 2002: Spikelets number estimation model using nitrogen nutrition status and biomass at panicle initiation and heading stage of rice. *Korean Journal of Crop Science* **47**, 390-394.

Shin, J. and M. Lee, 1995: Rice production in south Korea under current and future climates. *Modeling the impact of climate change on rice production in Asia*. Matthews R. M., Kroff M. J., Bachelet D. and van Larr H.H. (eds), IRRI, CAB International 199-213.

Lee, C., J. Shin, D. Kim, K. Choi, T. Park, and J. Kim, 2004a: Flowering response of rice varieties on photoperiod at different temperature regimens. *Treatises of Crop Researches* **5**, 258-264. (in Korean with English abstract)

Lee, C., J. Shin, D. Kim, K. Choi, T. Park, and J. Kim, 2004b: Growth simulation of Ilpumbyeo under Korean environment using rice growth simulation model "Oryza 2000". *Treatises of Crop Researches* **5**, 250-257. (in Korean with English abstract)

Lee, C. 2008: Development and application of model for estimating grain weight and grain N content in rice. Seoul National University. Ph.D thesis 102-134. (in Korean with English abstract)

Loague, K. and R. E. Green, 1991: Statistical and graphical methods for evaluating solute transport models: overview and application. *Journal of Contaminant Hydrology* **7**, 51-73.

Rinaldi, M., N. Losavio, and Z. Flagella, 2003: Evaluation and application of the OILCROP-SUN model for sunower in southern Italy. *Agricultural Systems* **78**, 17-30.

Nash, J. and I. Sutcliffe, 1970: River flow forecasting through conceptual models Part1-A discussion of Principles. *Journal of Hydrology* **10**, 282-290.

Nicolardot, B., S. Recous, and B. Mary, 2001: Simulation of C and N mineralisation during crop residue decomposition: A simple dynamic model based on the C:N ratio of the residues. *Plant and Soil* **228**, 83-103.

Ross, T., 1996: Indices for performance evaluation of predictive models in food microbiology. *Journal of Applied Bacteriology* **81**, 501-508.

Ross, T., P. Dalgaard, and S. Tienungoon, 2000: Predictive modelling of the growth and survival of *Listeria* in shery products. *International Journal of Food Microbiology* **62**, 231-245.

Willmott, C., 1981: On the validation of models. *Physical*

Geography **2**(2), 184-194.

Willmott, C., 1982: Some comments on the evaluation of model performance. *Bulletin American Meteorological Society* **63**(11), 1309-1313.