植物营养与肥料学报 2012,18(5):1064-1072
Plant Nutrition and Fertilizer Science

DSSAT 作物模型的统计和图形校验方法

杨靖民[1], 杨靖一[2*], Hoogenboom Gerrit[3]
(1 吉林农业大学资源与环境学院,吉林长春 130118; 2 Greenhouse & Processing Crops Research Centre, Agriculture and Agri-Food Canada, 2585 County Road 20, Harrow, Ontario, Canada N0R1G0; 3 AgWeatherNet Program, Washington State University, Prosser, WA 99350-8694, USA)

摘要:本文以玉米,大豆,马铃薯和冬小麦试验为例,详细讨论了 DSSAT 模型的统计校验和图形校验方法及其支持软件 EasyGrapher 的应用。结果指出:R^2不是理想的模型校验统计,因为它是测量线性回归 $y = \alpha + \beta x + \varepsilon$ 的拟合度,其随机误差 ε 建立在正态性、相互独立和方差同质的假设之上。模型校验主要是检验残差 $d_i = y_i - x_i$(y 为测量值,x 为模拟值),不是估计回归系数 α、β。*RMSE*、*E*、*EF* 和 *d* 都是比较好的校验模型残差的统计量,它们不需要遵循回归模型的基本假设,统计量的物理意义明确,大样本容量下统计校验更为可靠。图形校验是模型校验必不可少的辅助工具,当有测量数据时,时间序列和残差校验图形是两种最基本的图形校验方法;没有测量数据,模拟图形仍然能够检验模型输出之间或与时间的动态关系,辨析模拟误差或者错误。应用 EasyGrapher 软件可以快速进行 DSSAT 模型的图形校验和统计校验。

关键词:DSSAT 模型;EasyGrapher 软件;统计校验;图形校验
中图分类号:S126　　**文献标识码**:A　　**文章编号**:1008-505X(2012)05-1064-09

Statistical and graphical methods for evaluation of DSSAT crop model

YANG Jing-min[1], YANG Jing-yi[2*], Hoogenboom Gerrit[3]
(*1 Faculty of Natural Resources and Environment, Jilin Agricultural University, Changchun 130118, China; 2 Greenhouse & Processing Crops Research Centre, Agriculture and Agri-Food Canada, 2585 County Road 20, Harrow, Ontario, Canada N0R1G0; 3 AgWeatherNet Program, Washington State University, Prosser, WA 99350-8694, USA*)

Abstract: This paper discussed statistical and graphical evaluation methods for DSSAT model using field experiments of maize, soybean, potato and winter wheat. The results indicated that R^2 is not a good statistic for model evaluation because it tests the goodness of fit of a linear regression $y = \alpha + \beta x + \varepsilon$ where random error, ε, was assumed to follow normality, independence and equal variance. Model evaluation aims at testing residual error $d = y - x$(y measured data, x simulated data), but not estimating regression coefficients, α, β, *RMSE*, *E*, *EF* and d are all good "difference measures", they do not need follow three assumptions, and they have clear physical meaning. Large sample size increase reliability of statistics. Graphical evaluation is a necessary method for model evaluation. Time series and residual error graphs are two basic graphical methods for model evaluation if there are measured data. Simulation graphs can also be used to display relationships among outputs or against time, to analyze residual errors or mistakes even if no measured data. EasyGrapher program is a useful tool for statistical and graphical evaluations of DSSAT model's output.

Key words: DSSAT model; EasyGrapher program; statistical evaluation; graphical evaluation

由于缺乏对系统真实过程的完全理解,作物模型是建立在许多假设和经验基础上,导致了观测结果和模拟结果的差异,因此,对模型模拟结果和测量数据进行校验是必要的。模型校验可以通过两种途

收稿日期:2012-01-04　　**接受日期**:2012-05-29
基金项目:吉林省科技厅项目(20040548-2, 20100581)资助。
作者简介:杨靖民(1970—),男,吉林省榆树人,博士,副教授,主要从事作物生长模型与模拟技术研究。E-mail: yangjingmin2001@163.com
*通讯作者 E-mail: yang_jingyi@hotmail.com

径: 一是用一组测量数据比较不同模型; 二是用不同测量数据比较同一种模型。不论取哪种方法,比较的原理是一致的,既寻求发现模拟结果和测量数据的差异是来自于实验误差还是模型的失真。

模型校验可分为程序测试和输出结果测试两步,对程序进行测试是分析模型的输出值和其假设理论的逻辑关系是否一致,通常这项工作是建立模型的关键步骤,也是在建模时完成的。例如,对一个即将发行的新版模型,这个阶段的目的在于通过正确的编程以准确表达模型的数量关系,提高模型准确性。对模型模拟结果的测试则是检验模型的假设和数量关系是否真实地反映了所模拟的过程(如生长过程、土壤养分循环过程等)。这是通过一系列校验方法比较模型输出与实验数据或者输入参数的吻合性。模型输出结果校验和软件编写校验相比是更难的一步,Knepell 和 Arangno 1993 年对模型校验给出了最好的概括[1]。

在模型的建立和应用中, 模型的敏感性分析和统计校验是两个基本的模型校验方法; 前者是检验模型参数对模型输出的影响, 具体方法在前文中有详细报道[2-3]。统计校验分析是指用测量数据对模拟数据的比较分析。测量数据分为时间序列观测数据和阶段数据两种; 时间序列可以按小时(例如土壤水、温度)、天(土壤氮)、周(生物量、土壤氮、水)等划分,阶段数据一般以作物生长季节为阶段单元,例如收获时期的考种数据、产量、土壤养分分析数据等。由于作物模型模拟每天生长和土壤水、温度和营养变化,所以输出的变量也可以分为时间序列输出和阶段输出。一个好的校验分析应该既包含时间序列校验又包括阶段校验。这样就对测量数据提出要求,既要有时间系列的观测值,又要有作物生长阶段观测值。如果没有足够的测量数据,统计校验分析会失真或无法进行。

在农业系统中,测量数据除了能从前人的科学试验得到,也可以从普查、调查或者年鉴汇总资料中得到。另外,也可以为模型校验而设计新的试验,通过取样来获得数据。统计校验的目标在于通过统计量化来判断模型残差是否来自随机误差或者模拟过程的失真,从而去修改假设,完善模型。

1　统计校验方法

使用数理统计方法去校验模拟数据和测量数据间的差异,称为统计校验。常用的统计方法有两类: 一类是应用经典统计方法,如回归分析和假设性检验(t,F 测验)来进行统计校验[4-6]; 另一类是应用近年来新发展起来的统计方法(例如 RMSE)进行校验,这类方法侧重于对残差 $d_i = y_i - x_i$ 进行分析,$i = 1,2,\cdots n$, y_i 代表测量值,x_i 代表模拟值。下面分别介绍两类统计方法的计算公式[7-8]。

1.1　经典统计方法

1.1.1 回归分析($y = \alpha + \beta x + \varepsilon$)　在一元线性回归模型下,检测无效假设 H0: $\alpha = 0$, $\beta = 1$。如果无效假设成立,既 $y = x + \varepsilon$,ε 为随机误差,表明模拟数据和测量结果统计上没有显著差异。如果 $\alpha = 0$,$\beta \neq 1$,说明模型过高或过低估计了实际结果; 如果 $\alpha \neq 0$,$\beta = 1$,说明模型有一个系统误差。这些结论有助于查找原因,为修正模型提供依据。

1.1.2 决定系数检验　$R^2(0 < R^2 \leqslant 1)$ 可以用于检验回归关系的拟合度。在 R^2 显著的情况下,还要检验回归系数 α、β[4],否则会得到不正确的结论。

1.1.3 *Paired-t* 检验　*Paired-t* 检验是用来检验模拟值和测量值是否来自于同一总体,即无效假设 H0 为: $\bar{d}_i = (\bar{y}_i - \bar{x}_i) = 0$, $\bar{y}_i$ 和 $\bar{x}_i$ 分别为 y 和 x 的算数平均数[5]。

$$Paired\text{-}t = \bar{d}/s_{\bar{d}}$$

根据 $|t| \geqslant t_{0.05}$ 或者 $t_{0.01}$,来判断无效假设是否成立。如果无效假设成立,说明模拟值和测量值来自同一总体,二者无显著差异。

回归分析、R^2 和 t 检验方法由于依赖于统计模型的三个基本假设(即误差的正态性、独立性和方差齐性),近年来在模型校验上的应用受到了限制。其原因是作物生长状态变量(例如干物质重、作物氮吸收等)都和时间有关系,并且每次取样值和以前的测量值有自相关,导致了误差方差随时间而增加, 即破坏了正态性、独立性和方差齐性的原则[5,8-9]。所以建议在回归分析之前,都要对误差方差进行检验,然后进行适当转换再进行回归分析等等[6]。

1.2　残差统计方法

鉴于经典统计方法存在的问题,一些其它统计方法应运而生,以下几个统计量都是近几十年来常用的残差统计量(difference measures)。这些统计量不是建立在理论分布上的统计变量,而是直接对残差进行估计,所以无需考虑以上回归分析的三个假设[8-9]。

1.2.1 均方根差 *RMSE*(root mean squired error)　*RMSE* 是衡量模拟精度的统计量,由于它的计算和标准差相近,近年来应用较多。

$$RMSE = \sqrt{\sum (y_i - x_i)^2 / n}$$

式中：$i = 1,2,\cdots n$，y_i代表测量值，x_i代表模拟值，以下同。

如果要比较不同输出变量(生物量、叶面积指数 LAI)的统计校验效果。要用无量纲的统计量，因为带量纲的统计量受单位影响，不能进行此类比较。

RMSE 有时表达为百分数，称为标准均方根差(the normalized root mean square error)，记为 *nRMSE*(%)。

$$nRMSE = RMSE/\bar{y} \times 100$$

和 *RMSE* 相比，*nRMSE* 无量纲，可以用于不同变量之间的比较。当应用 *nRMSE* 时，一般都是根据经验给定差异是否显著的界线，例如 $nRMSE < 10\%$ 为无差异，$10\% \leqslant nRMSE < 20\%$ 为较小差异，$20\% \leqslant nRMSE < 30\%$ 为中等，$nRMSE \geqslant 30\%$ 为较大差异[10-14]。必须指出，这种划分不是统计学中概率意义下的划分，具有很大的主观性。应用时还要根据其它统计量（如 *EF* 和 *d* 值）来作出模拟结果是否与测量结果相吻合的结论。

1.2.2 平均误差 *E* *E* 可以检测模型模拟值与测量值的平均偏离程度，*E* 为正值时表明模拟值偏高，反之偏低。

$$E = [\sum (y_i - x_i)/n]$$

1.2.3 平均绝对误差 *MAE* *MAE*（mean absolute error）克服了 *E* 统计量的正负相互抵消效果，因此取平均绝对离差。所以 *MAE* 是指模型模拟值偏离测量数据的绝对平均程度。*MAE* 越大，模拟平均值偏离测量平均值越大。

$$MAE = \sum |y_i - x_i| / n$$

RMSE，*E* 和 *MAE* 都和 $d_i = y_i - x_i$ 取相同单位，经常用来比较模拟值和测量数据的符合程度。

1.2.4 预报效率 *EF*(efficiency of forecasting) 预报效率 *EF* 被 Willmott 于 1980 年用于模型校验后[15]，被许多模型专家用于模型校验，作为测量残差的相对统计量。*EF* 与 R^2 相似，但二者不等，例如 $EF \leqslant 1$，即 *EF* 可以小于 0，但是 $0 \leqslant R^2 \leqslant 1$[8]。

$$EF = 1 - \sum (y_i - x_i)^2 / \sum (y_i - \bar{y})^2$$

1.2.5 一致性指数 *d*(index of agreement) *d* 统计量被 Willmott[15] 于 1980 年用于环境模型校验后，也一直被许多学者关注和引用，被认为是一个和 R^2 相似的有界相对统计量：

$$d = 1 - \sum (y_i - x_i)^2 / \sum (|y_i - \bar{y}| + |x_i - \bar{y}|)^2$$

EF($EF \leqslant 1$) 和 *d*($0 \leqslant d \leqslant 1$) 都是无量纲统计量，主要用于检测在多大程度上残差 $d_i = y_i - x_i$ 趋近于零，描述模型模拟的不同输出变量的准确性。

统计方法的优势在于它量化了模型模拟和观测数据的差异，并指出何时、何地模型导致了不能接受的误差。统计校验要求有足够的样本容量，当数据少的情况下统计校验精度差。对于用以上统计校验方法校验模型，国外学者做了大量工作，例如 Willmott 等系统介绍了模型统计校验的特点和统计方法[9,15]；Reckhow 等系统地比较了经典统计校验方法[6]；Smith 等应用 6 个统计量对 9 个土壤有机质模型在 12 个长期定位试验上进行了系统的统计比较[16]；Yang 等系统地介绍了模型校验的统计方法，并且对 7 个统计量进行了相似性聚类分析，得出该 7 种统计量分为 4 类：1) *E*；2) *EF*；3) *t*，*F* 和 *d*；4) *RMSE*，*MAE*[8]。因此在实际应用上，如果选出 4 个统计量，应该从以上各组选一个即可以，以避免重复计算。

国内近年来也有许多应用实例，邱建军等[17]利用回归分析和 R^2 检验方法，对 GOSSYM 棉花模型中的土壤参数进行了校验，根据土壤含水量的实测值和模拟值建立一元回归方程，通过计算 R^2 值，进而判断方程的差异显著性。姜志伟等[10]采用 *nRMSE* 来度量小麦和花生模拟值与实测值的相对差异程度，用一致性指数 *d* 来检验模拟值和实测值的吻合度。戴明宏等[18]利用 *RMSE* 校正了 CERES－Maize 模型的品种参数。

2 统计校验实例

以上统计量计算都是常用的模型统计校验方法，在实际应用中，大多数研究者都是利用不同软件或者自己编写的程序来计算。另外，有些统计量直接标记到图形上方便分析。因此，图形分析和统计分析都是对模型结果进行校验分析不可缺少的方法。目前较好的作物模型，都有图形支持软件和统计分析结果。EasyGrapher 是 Yang 和 Huffman[19-20] 为 DSSAT v4 提供的图形和统计分析软件。近年来，EasyGrapher 不断更新，Yang 等为 DSSAT v4.5 推出了 EasyGrapher v4.5[21]，目前 EasyGrapher 软件已经成为 DSSAT 软件系统的支持软件，被世界许多模型专家应用[20]。下面的统计校验和图形校验分析是应用 EasyGrapher v4.5 对玉米、大豆、马铃薯和小麦田间试验的统计校验分析。

DSSAT 模型是集作物系统模型、土壤模型和土壤水动态模型为一体的软件[22-24]。它的另外一个特点是为使用者提供许多实际模拟试验范例,为用户使用 DSSAT 软件模拟田间试验提供帮助。本文从 DSSAT 软件中选取玉米、大豆、小麦和马铃薯 4 个试验,用 EasyGrapher v4.5 进行图形和统计分析示范,以便于读者查对和使用。选定这 4 个试验的原因是,以上所选定的 4 个作物代表了 DSSAT 作物系统模型中 4 个发展和应用都比较完善的模型,即 CERES-Maize (玉米), CERES-Wheat (小麦), CROPGRO-Soybean (大豆), SUBSTOR-Potato (马铃薯)。其主要步骤为:

第一,用 DSSAT 模型运行下列田间试验,以获得输出模拟数据。所选定的田间试验包括: 1999 年美国 Iowa 的玉米试验(氮肥和密度 2×2 设计),该试验是专门为校验 CERES-Maize 模型在 Iowa 所设计的玉米试验; 1981 年美国 Florida 的专门为校验 CROPGRO-Soybean 模型设计的大豆试验 (3 水平灌溉试验); 1980 年美国 New York 的马铃薯试验(为模型校验设计的只有一个处理的 N275 + 灌溉试验)和 1981 年美国 Kansas 的冬小麦试验(氮肥和灌溉 3×2 设计),该试验是 Waggger 的博士论文试验,后被 DSSAT 收录为示范试验。田间试验的地理位置、气象、田间管理和土壤基础数据见表 1。表 1 所给出的参数是模型模拟的最基本参数,但不是全部参数,目的是介绍每个作物的试验方案及其土壤、地理和气候条件。以便为下一步统计分析和图形分析提供实验模拟研究的示范数据。

表 1　用于 DSSAT 模型统计校验分析的田间试验参数

Table 1　Field experiments and parameters for statistical validation methods of DSSAT model

试验地点及参数 Experimental sites and parameters		作物 Crop			
		玉米 Maize	大豆 Soybean	马铃薯 Potato	冬小麦 Winter wheat
作物文件 Crop file (X)		IUAF9901	UFGA8101	NYCU8001	KSAS8101
处理 Treatments		N×密度(2×2)	灌溉 3 水平	N275 + 灌溉	N×灌溉 (3×2)
试验地点 Field location		美国 Iowa	美国 Florida	美国 New York	美国 Kansas
纬度 Latitude		42.02N	29.63N	42.70N	37.18N
经度 Longitude		93.75W	82.37W	97.00W	99.75W
生长期内作物生长参数 Crop parameters during growing season	生长期 Growth period	155	130	114	249
	密度 Population (plant/m^2)	4.7, 7.5	35.9	5.1	162
	氮肥 Fertilizer-N (N kg/hm^2)	56, 224		275	0, 60, 180
	灌溉量 Irrigation (mm)		315, 158, 264	209	0, 213
	降水量 Percipitation (mm)	535	270	296	579
	蒸发量 Evaptransporaiton (mm)	516	484	373	388
土壤参数 Soil parameters	土壤质地 Soil texture	细砂壤 Fine loamy	细砂 Fine sand	砂壤 Sandy loam	壤质 Loam
	容重 Bulk density (g/cm^3)	1.45	1.40	1.20	1.10
	田间持水量 Field capacity (cm^3/cm^3)	0.30	0.09	0.22	0.23
	剖面深度 Profile depth (cm)	152	180	65	150

第二,编辑田间试验测量数据用于校验分析。以上的测量数据文件都已经编辑好,并且分别存放到和管理文件相同的目录下,用相同的文件名,扩展名结尾用"T"标记的文件中,例如 IUAF9901. MZT 是玉米试验 IUAF9901. MZX 的测量数据文件。

第三,分别运行 EasyGrapher v4.5 软件,对以上 4 个试验的生长输出文件(PlantGrow. out)画图和统计分析。EasyGrapher 软件先进行数据处理和统计分析,然后调用 MS Excel 进行自动画图和统计输出,把每类输出存放为一页。主要输出包括: 1)时

间系列图形　以X轴代表生长期,每个变量画一个图形。玉米、大豆、马铃薯、小麦生长输出分别有46、42、30和45个时间系列图形(图1);2)统计校验图形　以X轴代表模拟值,Y轴代表测量值,每个测量数据变量输出一个图形。玉米、大豆、马铃薯、小麦的XY统计校验图形数分别有10、14、6、6个,校验统计也输出到XY图形上,以便于分析(图2);3)数据输出　EasyGrapher软件输出3种重要数据(测量数据、模拟数据和统计数据)。

目前,EasyGrapher v4.5软件能够在大多数语言环境下的MS Window系统下运行,并且在很短时间内(大约每个文件0.5～1分钟,取决于计算机速度),对DSSAT模型的15个输出文件进行图形和统计分析。这些文件大多数都是时间序列输出;包括作物生长(PlantGro. OUT)、作物氮(PlantN. OUT)、作物碳(PlantC. OUT)、土壤水(SoilWat. OUT)、土壤温度(SoilTemp. OUT)、土壤有机质(SoilOrg. OUT)、土壤无机氮(SoilNi. OUT)、天气(Weather. OUT)、植被蒸腾(ET. OUT)、覆盖(Mulch. OUT)和管理措施(MgmtOps. OUT),也包括生育期结束后的阶段输出(Summary. OUT)和阶段校验分析(Evaluate. OUT)等。为此,大大加快了DSSAT模型的图形和统计校验速度。下面以地上干物质重和产量为例,讨论统计分析和图形分析结果(图1～图3)。

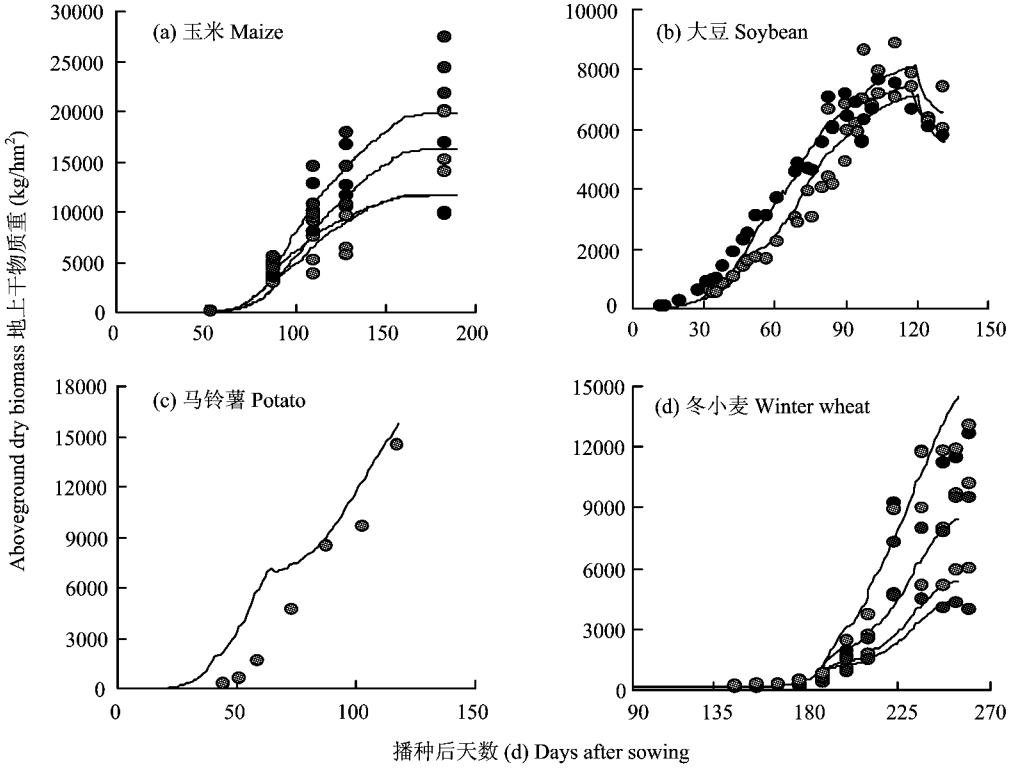

图1　用EasyGrapher软件对DSSAT作物生长曲线校验分析

Fig.1　Evaluation of DSSAT crop growth curves using EasyGrapher program

[注(Note):图(a)、(b)、(c)、(d)中的线为不同处理模拟值,点为不同处理测量值 Lines with different colors represent the simulated aboveground biomass and symbols represent the measured data in the Fig(a),(b),(c),(d).]

3　结果与分析

3.1　统计校验

统计分析结果量化了残差的大小和方向(表2),现分别讨论如下:*RMSE*、*E*和*MAE*所校验的变量(例如产量)取相同单位,可以用来和常规的变量平均相互比较以判别其差异的大小。例如,在收获时,玉米、大豆、马铃薯和冬小麦的地上平均干物质重分别为17738、6111、14440和8751 kg/hm²,*nRMSE*分别为13.9%、8.5%、18.8%和15.0%。从产量看,玉米,大豆,马铃薯和冬小麦的平均产量分别为7402、3131、11020和2688 kg/hm²,*nRMSE*分别为14.5%、8.6%、8.1%和20.6%。因此可以得出大豆的标准误差较小(<10%)。从*E*值可以看出,玉米和小麦的模拟值低于测量值,马铃薯的模拟值高于测量值,而大豆的模拟地上干物质重略

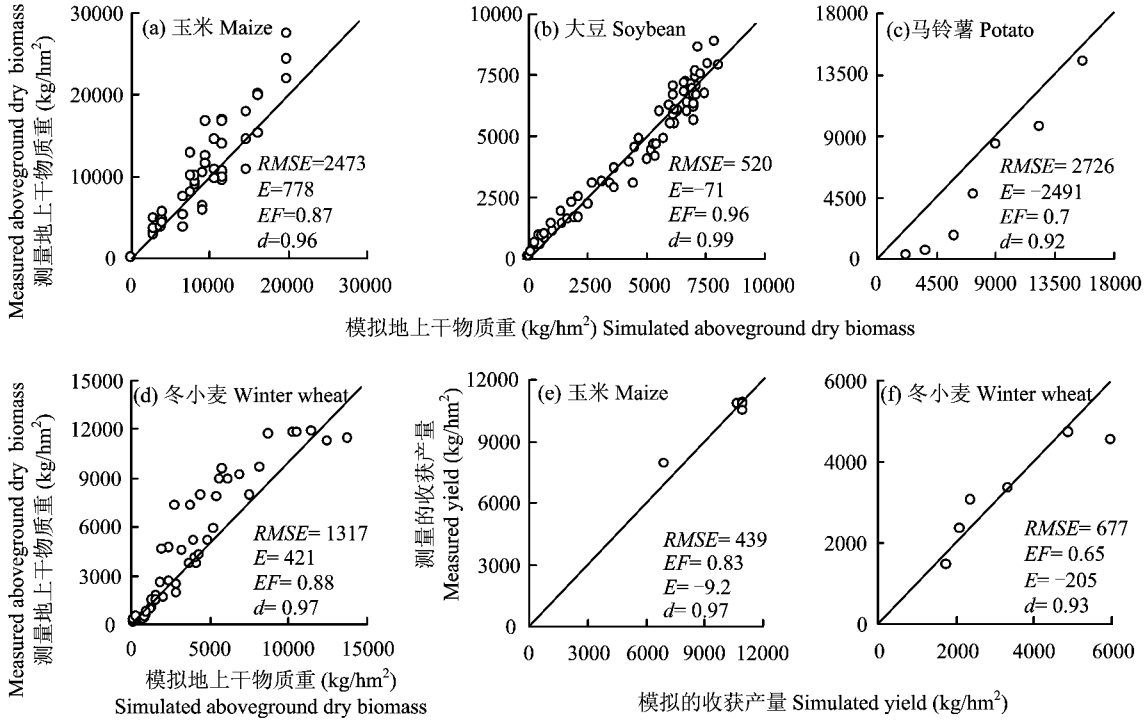

图 2　用 EasyGrapher 软件对 DSSAT 作物地上干物质重和产量的统计校验分析

Fig. 2　Statistical evaluation of DSSAT simulated aboveground dry biomass and yields using EasyGrapher program

[注(Note): 图(a),(b),(c),(d),(e),(f)的对角线是理想曲线,点表示测量值和模拟值的位置,凡是远离对角线的点,模拟的残差越大。*RMSE*,*E*,*EF* 和 *d* 和表 2 的统计量相同,由 EasyGrapher 软件自动绘制到统计检验图中 1∶1 lines represents 100% agreement of the simulation, symbols represent the positions of each simulated and measured value, the far the symbol's position from the line, the larger the residue error in the Fig (a), (b), (c), (d), (e), (f). *RMSE*, *E*, *EF* and *d* values are the same as in table 2; they were copied to the figure by EasyGrapher program automatically.]

表 2　DSSAT 模拟的玉米、大豆、马铃薯和冬小麦地上干物重和产量的统计校验分析

Table 2　Statistical evaluation of aboveground dry biomasses and yields of maize, soybean, potato and winter wheat by DSSAT model

项目 Item	统计量 Statistics	玉米 Maize	大豆 Soybean	马铃薯 Potato	小麦 Wheat
地上干物质重 Aboveground dry biomass (kg/hm^2)	*RMSE*	2473	520	2726	1317
	E	778	-71	-2491	421
	MAE	1696	402	2491	746
	EF	0.87	0.96	0.70	0.88
	d	0.96	0.99	0.92	0.97
	Sample No.	59	91	7	72
产量 Yield (kg/hm^2)	*RMSE*	1074	268	898	554
	E	617	110	-101	152
	MAE	933	220	642	485
	EF	0.72	0.95	0.94	0.47
	d	0.92	0.99	0.99	0.91
	Sample No.	10	37	5	12

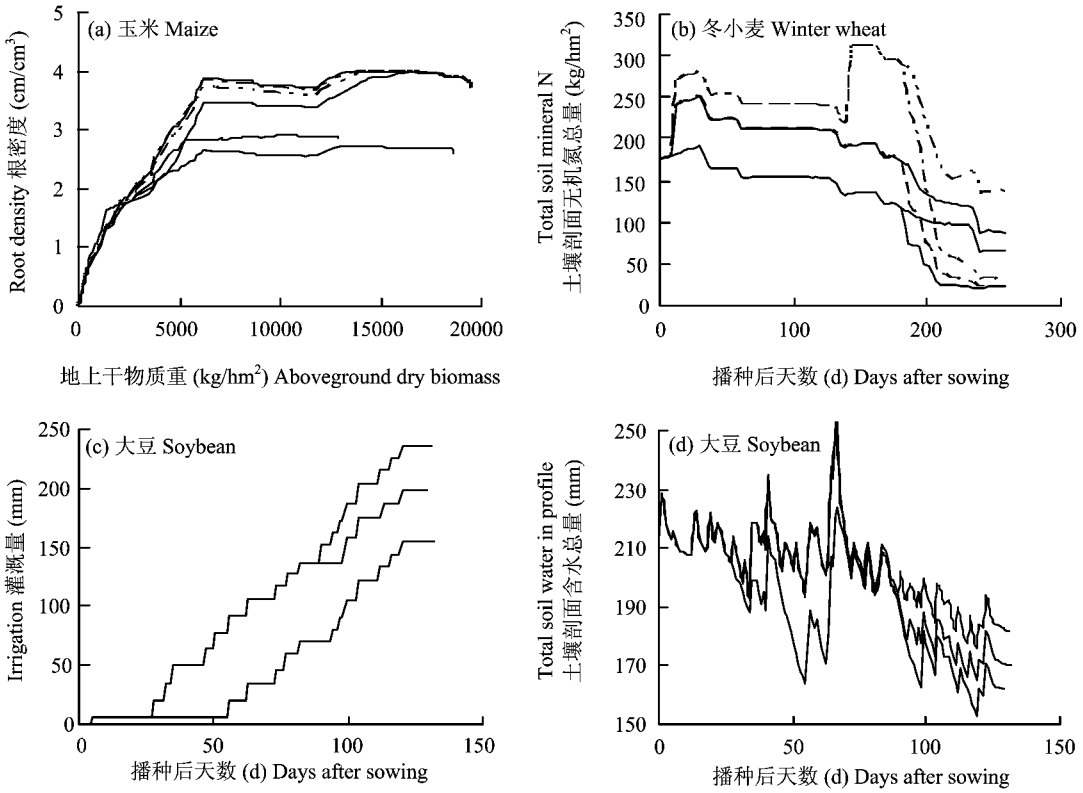

图3 用 EasyGrapher 软件对 DSSAT 模拟的根系生长和土壤无机氮、灌溉量、土壤含水量动态变化的模拟图形

Fig. 3 Examples of simulated graphical displays on DSSAT simulated root growth, soil mineral N, irrigation and soil water dynamics using EasyGrapher program

低于测量结果,但是模拟的产量高于测量产量。从 *EF* 和 *d* 值看出,除小麦产量外,其它统计值介于 0.70～0.99 之间,说明模拟值和测量值之间的复合度较高。

测量样本的大小对统计值有很大影响。一般而言,测量样本越大,所计算的统计值就越具有代表性,可信度也就越高,如玉米试验在生育期有 5 次采样,共 4 个处理,每个处理取 3 次重复,由于第 3 个处理后期缺少 1 次取样,因此一共有 59 个测量样本用于校验。相比之下,玉米产量取样在生育后期,只有 1 次取样,有 2 个处理只取 2 次重复,因此只有 10 个测量数据。马铃薯试验只有 1 个处理,一共取 7 次生育期样本,每次记录一个平均值,因此,只有 7 个地上干物质重的测量样本和 5 个块根样本。

如果只对模型最后结果校验,测量数据等于处理数。例如对以上的 4 个作物收获时期的干物质重和籽粒产量的统计校验,每个处理只有一个最后测量数据,玉米、大豆和小麦的测量数据只有 4、3 和 6 个,计算的 *d* 值分别为 0.97、0.79 和 0.93。马铃薯试验只有 1 个处理,导致 $RMSE = E = 1216$, $d = 0$, EF 值无法计算。玉米、小麦的校验图形和统计列于图 2e、f。这个例子指出,校验分析要尽量增加测量数据,例如增加处理数或者增加重复测量数,才能有效地运用统计方法进行校验分析。

3.2 图形校验

图形校验是用图形来表达模拟数据和测量数据相互吻合的一种直观方法,常用于模型校验中。有许多图形表达方法,这里重点介绍两种常用的图形,时间序列图形和 XY-图形。时间序列图形一般以生育期为 X 轴,以输出变量为 Y 轴,模拟数据用连续性曲线表示,测量数据用点表示(图 1)。XY-图形以 X 和 Y 轴分别代表模拟和测量数据(或者相反),然后用 1∶1 对角线为基准线,判断模拟数据和测量数据的吻合度(图 2)。有时研究者把线性回归和 R^2 也放到图形中,来表示 X、Y 之间的回归关系。在这种情况下,应该测验是否回归系数 $\alpha = 0$, $\beta = 1$,如果 $\alpha \neq 0$, $\beta \neq 1$,即使 R^2 值很大,只能说明回归关系拟合得很好,但是没有说明 X、Y 数据之

间的残差 $d_i = y_i - x_i$ 的大小。近年来,关于校验生物量和产量,经常用 *RMSE* 和 d 统计量来直接检验残差 $d_i = y_i - x_i$,以避免 R^2 给模型校验带来的偏差。

和统计校验比较,图形校验还有它的优点,即能够判定在什么时候模拟数据偏离测量数据,这样有助于对作物生长曲线的正确校验。例如,表 2 所显示的玉米,大豆和小麦的地上干物质重的 d 统计量没有本质区别(即 0.96、0.99 和 0.97),但是从图 2 的图形校验清楚看出,只有大豆的模拟值和测量值在整个生育期内相互吻合;模拟的玉米地上干物质重在生育后期显著比测量值低;模拟的冬小麦地上干物质重在生长中期(7000 ~ 10000 kg/hm^2 范围内)明显比测量值低。又例如,作物的肥料效应曲线有多种,它们的统计拟合结果基本没有差异,但是模拟的曲线走向却不尽相同。图形分析有助于选择在取值范围内的最优拟合曲线,即模拟曲线要在整个预测范围内通过测量值中心曲线(图 1b、图 2b)。关于作物生长曲线的模型校验和肥料效应曲线的比较研究,可参阅文献[25 -26]。

只有模拟数据而没有测量数据的图形简称为模拟图形。这类图形大多数以时间序列图形为主,如果要分析输出变量之间的关系,可以选用 XY 图形。EasyGrapher 软件特别提供了这种图形输出功能。例如,要研究植株氮含量和干物质重的关系,可以用 Y 轴代表氮含量,X 轴代表地上干物质重[2]。本文列举玉米根系生长密度与地上干物质重的关系,用 Y 轴代表根密度,X 轴代表干物质重 (图 3a)。其它例子都以 X 轴为时间,例如,冬小麦试验表层土壤无机氮随不同处理在生长期的分布(图 3b),大豆试验不同时期灌溉量(图 3c),大豆试验不同灌溉量下土壤含水量的分布(图 3d),可以用模拟图形清楚地表示出来。模拟图形被广泛地应用到模型的模拟研究中。

3.3 模型预报

利用作物动态模型进行预报,前提是要对模型进行如上的统计校验和图形校验后,得到了配参后具有准确模拟功能的模型。预报分为两个方向:第一是对同一作物的不同田块、地域进行预报,比如应用作物模型对区域的作物进行预报;第二是对作物生产力的未来预报,比如气候变暖预报和农业政策模拟。第一种预报实际上是从点到面的过程,要求所预报的区域在作物生长的同一个生态带,否则就要重新校验品种参数。在预报过程中,要尽量应用所预报田间和区域的土壤类型及其稳定的理化数据和天气数据,以便反映当地的水热和营养条件,使预报更准确。第二种预报是对未来气候或政策进行预测模拟,因此要建立一个未来的"假设"的天气数据。比如说,未来 50 年气候变暖对作物产量的影响预报,要建立一个气候变暖相适应的天气数据库。有关这方面的内容,已经超出本文的研究范围,此处不多述。

4 小结

作物模型如今和其它模型一样,被我国农业科学家和推广部门广泛用于生长模拟和环境保护研究。但是关于模型校验统计方法还没有一个统一的认识和使用标准。本文通过应用 DSSAT 作物模型和作者为其开发的图形和统计校验软件 EasyGrapher 为范例,系统总结了模型的统计校验研究和图形校验的经验。关于模型校验的统计方法国外早有评价,图形支持软件也有很多不同特点,但是都还在发展时期,有待进一步完善。本文的统计和图形校验方法是在 DSSAT 模型基础上所做的部分研究,其统计方法适合其它模型校验。EasyGrapher 软件是一个独立无偿使用软件,目前被 DSSAT 软件兼容在世界各地使用,也可以用于支持其它模型的统计校验和图形分析[19]。作者希望借助本文,把模型统计校验方法和图形校验方法介绍给我国的模型研究专家,为更好地普及,学习和推广作物模型在我国的应用做出贡献。

致谢: 真诚感谢 DSSATv4.5 模型所提供的田间试验和测量数据。EasyGrapher 软件研究和开发得到加拿大农业与食品科学部的研究资金支持。

参 考 文 献:

[1] Knepell P L, Arangno D C. Simulation validation: A confidence assessment methodology [M]. Los Alamitos, California: IEEE Computer Society Press, 1993. 123-131.

[2] 杨靖民, 刘金华, 窦森, 等. 应用 DSSAT-CSM Maize 模型对吉林省黑土玉米最佳栽培技术的模拟和校验研究 I-模型品种参数校验和产量的敏感性分析[J]. 土壤学报, 2011, 48 (2): 366-374.

Yang J M, Lui J H, Dou S *et al*. Evaluation and optimization of best management practices of maize for black soil in Jilin China using the DSSAT model I- cultivar calibration and sensitivity

analysis of maize yield parameters[J]. Acta Pedol. Sin., 2011, 48(2): 366-374.

[3] 杨靖民，窦森，杨靖一，等. 应用DSSAT模型对吉林省黑土作物-土壤氮循环和土壤有机碳平衡的敏感性分析[J]. 应用生态学报, 2011, 22(8): 2075-2083.
Yang J M, Dou S, Yang J Y *et al*. Crop-soil nitrogen cycling and soil organic carbon balance in black soil zone of Jilin Province based on DSSAT model[J]. Chin. J. Appl. Ecol., 2011, 22(8): 2075-2083.

[4] Aigner D J. Basic econometrics[M]. Englewood Cliffs, NJ: Prentice-Hall, 1971.

[5] Snedecor G W, Cochran W G. Statistical methods(6th Edn.)[M]. Ames, IA, USA: The Iowa State University Press, 1976.

[6] Reckhow K H, Clements J T, Dodd R C. Statistical evaluation of mechanistic water-quality models[J]. J. Environ. Eng., 1990, 116, 250-268.

[7] Willmott C J. Some comments on the evaluation of model performance[J]. Bull. Am. Meteorol. Soc., 1982, 63: 1309-1313.

[8] Yang J, Greenwood D J, Rowell D L *et al*. Statistical methods for evaluating a crop nitrogen simulation model, NABLE[J]. Agric. Syst., 2000, 64(1): 37-53.

[9] Willmott C J, Ackleson S G, Davis R E *et al*. Statistics for the evaluation and comparison of models[J]. J. Geophys., Res., 1985, 90: 8995-9005.

[10] 姜志伟，武雪萍，华珞，等. 洛阳旱地夏玉米生产潜力长周期定量模拟与评价[J]. 生态学报, 2009, 29(1): 315-324.
Jing Z W, Wu X P, Hua L *et al*. Long-period quantitative simulation and evaluation of summer maize potential productivity in Luoyang dryland[J]. Acta Ecol. Sin., 2009, 29(1): 315-324.

[11] Timsina J, Humphreys E. Performance of CERES-Rice and CERES-Wheat models in rice-wheat systems: A review[J]. Agric. Syst., 2006, 90: 5-31.

[12] Biernath C, Gayler S, Bittner S *et al*. Evaluating the ability of four crop models to predict different environmental impacts on spring wheat grown in open-top chambers[J]. Eur. J. Agron., 2011, 35: 71-82.

[13] Liu H L, Yang J Y, Drury C F *et al*. Using the DSSAT-CERES-Maize model to simulate crop yield and nitrogen cycling in fields under long-term continuous maize production[J]. Nutr. Cycl. Agroecosyst., 2011, 89: 313-328.

[14] Rinaldi M, Ventrella D, Gagliano C. Comparison of nitrogen and irrigation strategies in tomato using CROPGRO model. A case study from Southern Italy[J]. Agric. Water Manag., 2007, 87: 91-105.

[15] Willmott C J. On the validation of models[J]. Phys. Geog., 1981, 2: 184-194.

[16] Smith P, Smith J U, Powlson D S *et al*. A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments[J]. Geoderma, 1997, 81: 153-225.

[17] 邱建军，肖荧南，胡锡宁. 作物生长模拟模型参数校正与有效化的理论和实践[J]. 应用生态学报, 1999, 10(6): 679-682.
Qiu J J, Xiao Y N, Hu X N. Theory and practice of calibration and validation of crop simulation model[J]. Chin. J. Appl. Ecol., 1999, 10(6): 679-682.

[18] 戴明宏，赵久然，Wilhelm Claupein，王璞. 基于CERES-Maize模型春玉米水分优化管理决策[J]. 水土保持学报, 2009, 23(1): 187-192.
Dai H M, Zhao J R, Wilhelm C, Wang P. Decision for optimized water management based on CERES-Maize crop model[J]. J. Soil Water Conserv., 2009, 23(1): 187-192.

[19] Yang J Y, Huffman E C. EasyGrapher: software for graphical and statistical validation of DSSAT outputs[J]. Comput. Electron. Agric., 2004, 45: 125-132.

[20] Yang J Y, Huffman E C. EasyGrapher v1.0 help manual[A]. Jones J W, Hoogenboom G, Wilkens P W *et al*. Decision support system for agrotechnology transfer version 4.0[M]. Honolulu: University of Hawaii, 2003. V3: 41-80.

[21] Yang J Y, Drury C F, Johnston R *et al*. EasyGrapher v4.5: software for graphical and statistical evaluation of DSSAT v4.5 outputs[Z]. Lang Beach CA: Poster presentation in annual meeting of ASA-CSSA-SSSA. 2010.

[22] Tsuji G Y, Uehara G, Balas S. DSSAT v3[M]. Honolulu: University of Hawaii, 1994.

[23] Hoogenboom G, Jones J W, Porter C H *et al*. Decision support system for agrotechnology transfer version 4.0, V1 Overview[M]. Honolulu: University of Hawaii, 2003.

[24] Jones J W, Hoogenboom G, Porter C H *et al*. The DSSAT cropping system model[J]. Eur. J. Agron., 2003, 18: 235-265.

[25] Yang J, Wadsworth G A, Rowell D L, Burns I G. Evaluating a crop nitrogen simulation model, N_ABLE, using a field experiment with lettuce[J]. Nutr. Cycl. Agroecoyst., 1999, 55: 221-230.

[26] 杨靖一，Wadsworth G A，Greenwood D J. 三种蔬菜氮肥效应曲线的比较研究[J]. 植物营养与肥料学报, 1995, 1(1): 71-78
Yang J Y, Wadsworth G A, Greenwood D J. A comparison study of three types of N-fertilizer response curves on vegetables[J]. Plant Nutr. Fert. Sci., 1995, 1(1): 71-78.