

谷物联合收割机测产数据中的误差分析与处理

张 漫, 邝继双

(中国农业大学现代精细农业系统集成研究教育部重点实验室, 北京 100083)

摘 要: 产量分布图作为实施精细农业的起点与农业生产管理决策的基础, 其精度至关重要, 而产量数据误差分析与处理则是提高产量分布图精度的关键。通过连续两年使用带有测产系统的联合收割机进行收获试验, 并对得到的产量数据进行分析, 发现产量数据中主要存在三类误差, 即割幅宽度设置误差、填充时间误差和延迟时间误差。讨论了这三类误差的产生原因、识别和处理方法, 并对小麦和玉米在不同收获条件下的误差进行了比较。分析结果表明, 割幅宽度设置误差数据所占的比例一般小于 6%, 填充时间误差数据大于 10%, 延迟时间误差数据小于 1%。

关键词: 测产系统; 误差分析; 产量分布图; 谷物联合收割机

中图分类号: S127; S225.3

文献标识码: A

文章编号: 1002-6819(2003)03-0144-05

1 引言

作物产量作为农业生产劳动的成果是评估农业生产效率的一个重要因素^[1]。而获取农作物小区产量信息, 建立小区产量空间分布图, 则是实施“精细农作”的起点, 也是实现作物生产过程中科学调控投入和制定管理决策措施的基础^[2]。所以在实施任何田间定位管理措施之前, 进行作物产量监测, 生成高精度的产量分布图是十分重要的^[3]。但是目前安装在谷物联合收割机上的测产系统大多数比较复杂, 一般由中央控制系统、全球定位系统、谷物流量传感器、谷物含水率传感器、速度传感器、割幅宽度传感器、割台高度传感器、升运器转速传感器、显示与操作终端、数据存储设备等部分组成。在实际田间收获过程中除了标定以外, 测产系统均处于自动运行状态进行产量数据采集, 工作过程当中基本不能进行人工干预。另外由于田间的工作环境比较复杂, 产量数据采集过程必然会受到测产系统本身以及外界环境当中多种复杂因素的影响, 这样一来产量数据当中就会不可避免的存在一些误差。所以从产量数据生成产量分布图之前的数据误差分析与处理就成为了保证产量分布图精度的一个重要环节。

国外开展精细农业研究工作较早, 在许多国家, 已经积累了 7~8 年的产量数据, 已有一些研究者致力于这方面的研究, 有一些误差被识别, 但还没有进一步得到量化。例如: Murphy et al(1994)指出, 当收割机粮仓空着时脱粒后的谷物到达传感器的填充时间与谷物传输系统的滞后时间是不同的, 但目前许多商用系统并没有考虑到填充时间与延迟时间的区别, 引起产量数据误差^[4]。Moore(1998)对产量数据中包含的面积测量误差、延迟时间误差、引导时间误差、位置滞后误差等进行了分析^[5]。Andy D. Beck(1999)开发了一种识别产量

数据中的误差数据点的滤波函数, 主要依据是设定可接受的谷物产量值、含水率和收割机的行使速度等^[6]。国内在这方面起步较晚, 近几年先后在北京、上海、新疆、黑龙江等地建立了精细农业示范区, 引进国外商品化的产量监测系统收获试验, 并开始进行这方面的研究工作。

2 试验材料与方法

北京精准农业示范区位于北京市昌平区小汤山镇, 是由国家计委立项, 国家计委、北京市和北京市农林科学院共同投资建设的国家级高技术产业化示范区。示范区总面积 2500 亩(1 亩= 1/15 hm²), 划分为 23 个地块, 种植的作物有小麦、玉米和大豆。示范区在 2001 年 5 月, 引进了美国 CASE 公司带有测产系统 AFS (Advanced Farming System) 的 2366 谷物联合收割机。AFS 系统的关键部分包括美国 Trimble 公司的 AgGPS 132 差分式 GPS 接收机, 美国 CASE 公司的冲击式谷物流量传感器和配套的产量图生成软件^[7]。

在示范区中分别于 2001 年 6 月和 10 月对 100 hm² 的冬小麦和玉米进行了测产试验, 得到了产量数据。收获小麦使用的是 CASE 公司割幅为 6 m 的 1020 割台, 收获玉米使用的是 CASE 公司 6 行的 1083 割台。在收割之前 AFS 测产系统按照使用要求进行了设置, 并在收割过程当中分别对谷物质量、含水率、温度, 以及收割机的行走距离和收获面积进行了严格标定, 使系统测量误差处在了允许范围以内^[8,9]。

3 数据分析与处理

3.1 产量数据预处理

3.1.1 产量数据文件输出

AFS 系统获得的产量数据文件是扩展名为 YLD 的特殊格式文件, 这种文件必须使用 AFS 系统配套的 Instant Yield Map 软件来读取并生成产量分布图。但 Instant Yield Map 软件不包含误差处理功能, 而且用户也不能选择产量分布图生成的方法与参数。所以要想使用其它工具软件进行数据处理、误差分析与产量分布图的生成, 必须首先使用 Instant Yield Map 软件的数

收稿日期: 2002-09-02

基金项目: 国家发展计划委员会高技术产业化示范工程项目(北京精准农业示范工程 A 00300100584)

作者简介: 张 漫, 博士研究生, 北京清华东路 17 号 中国农业大学(东校区)63 号信箱, 100083。Email: cauzm@sohu.com



据输出功能将 YLD 文件中的产量数据输出为 TXT 文本格式的产量数据文件。

输出的产量数据文件可以为基本产量数据文件,也可以为详细产量数据文件。基本产量数据文件中的每条记录包括 8 个字段,分别为产量数据点的经度、纬度、作物产量、谷物含水率、监视器序列号、地块编号、车次编号和作物种类。详细产量数据文件中的每条记录包括 17 个字段,分别为产量数据点的经度、纬度、谷物流量、GPS 时间、采样周期、采样周期内收割机的行驶距离、割幅宽度、谷物含水率、系统工作状态、数据通道、监视器序列号、地块编号、车次编号、作物种类、GPS 状态、PDO P 值和高程。本文输出的产量数据文件为基本产量数据文件。

3.1.2 坐标系变换

使用以经纬度和高程表示的大地坐标系会给距离与面积的计算造成困难,所以产量数据文件中的大地坐标系需要转换为以 XYZ 表示的空间直角坐标系。本文中的坐标系变换使用美国 Macro Image 公司的 TN TL ite 6 3 软件进行,投影方式为 U TM (U niversal T ransverse M ercator),转换后的坐标系统为 W GS84 (W orld G eodetic S ystem 1984),坐标单位为 m。

3.1.3 产量单位换算

产量数据文件中作物产量的单位为蒲式耳/英亩 (bu/ac),而国内习惯使用的作物产量单位为吨/公顷 (t/hm^2)或斤/亩 (jin/mu),所以需要作物产量的单位进行换算^[7]。

3.2 误差分析与处理

通过对产量数据的处理与分析,并结合对联合收割机与测产系统田间作业过程的观察,本文发现产量数据中主要存在 3 种类型的误差,即割幅宽度设置误差、填充时间误差和延迟时间误差。

3.2.1 割幅宽度设置误差

带有测产系统的联合收割机工作时,在每个记录周期内系统计算并保存一次产量数据。产量数据值等于记录周期内谷物流量的累积值除以记录周期内所收割的小区面积。而小区面积则等于割幅宽度乘以记录周期内收割机的行走距离。目前,谷物流量的累积值与收割机的行走距离可以由系统自动测量计算得到,精度比较容易得到保障,而割幅宽度只能根据收割机操作者的目测来设置。

为了得到准确的产量数据,设置的割幅宽度与实际进入割台的作物宽度应时时保持一致。每块农田开始收割时,收割机大部分时间工作在满割幅情况下,这时将割幅宽度设为割台宽度即可。收割块边的剩余作物时,作物宽度一般达不到满割幅,这时需要根据作物的实际宽度设置割幅宽度。如果地边剩余部分的形状为规则的矩形,则割幅宽度一次设置后不需要反复改动,这时割幅设置的精度比较高。如果地边剩余部分的形状为梯形、三角形或不规则形状时,则进入割台的作物宽度是不断变化的。操作人员不可能对割幅宽度进行连续设置,而只能以一个平均的割幅宽度值来表示某段时间内

所收割的作物宽度。在这种情况下,由于割幅宽度设置而引起的产量数据计算误差较大,具有明显割幅宽度设置误差的数据点在产量分布图中常常呈条状分布,其产量值明显高于或低于周围其它数据点的平均产量^[8]。

本文对小麦和玉米的产量数据进行了分析,结果见表 1。从表 1 中可以看出,玉米收获中割幅宽度设置误差数据的比例明显高于其在小麦收获中所占的比例,这主要因为玉米是行播作物,收割时割幅宽度以行数计算,如果设置不当就会相差若干行,这样对产量值就会产生较大影响。另外,在小麦产量数据中,误差数据点在地块 F12、F23 中占的比例较小,分别为 1.28% 和 1.26%,而在地块 F15 中却达到 3.76%。这是因为地块 F15 与地块 F12、F23 相比形状很不规则。由此可见割幅宽度设置误差的产生与地块形状和收获的作物种类有关。

表 1 割幅宽度设置误差数据比例
Table 1 Proportion of raw data recorded
in cutter width setting error

| 地块名称 | 作物种类 | 产量数据点个数 | 误差数据点个数 | 误差数据点百分比/% |
|------|------|---------|---------|------------|
| F12 | 小麦 | 4448 | 57 | 1.28 |
| F15 | 小麦 | 3747 | 141 | 3.76 |
| F23 | 小麦 | 5323 | 67 | 1.26 |
| F19 | 玉米 | 4602 | 243 | 5.28 |
| F20 | 玉米 | 5637 | 257 | 4.56 |

3.2.2 填充时间误差

每次收割机开始收割时,谷物流量数据和谷物含水率数据都有一个逐步增加达到相对稳定状态的过程,这一过程经历的时间就叫做填充时间 (Filling Time)^[5]。在填充时间内所记录的产量数据是无效的,这些数据称为填充时间误差数据。通过对谷物流量与谷物含水率填充时间的研究,可以掌握其产生与变化规律,从而达到识别与补偿填充时间误差数据的目的。

下面分别对谷物流量和谷物含水率的填充时间,在小麦与玉米收割时,以及收割机空车收割与地头调头再继续收割时的不同情况进行讨论。

1) 小麦含水率的填充时间

收割机空车收割情况下,小麦含水率数据的变化曲线如图 1 所示,其中包括 5 组数据,坐标纵轴表示小麦含水率,横轴表示时间,单位为秒。从图 1 可以看出,小麦含水率经过一段填充时间达到稳定值后,其变化范围很小。小麦含水率在空车情况下达到稳定状态,需要的平均填充时间为 48.8 s。

收割机地头调头再进地收割的情况下,小麦含水率数据的变化曲线如图 2 所示。达到稳定状态需要的平均填充时间为 10.8 s,远远小于收割机空车收割情况下的平均填充时间。

2) 玉米含水率的填充时间

收割机空车收割情况下,玉米含水率数据的变化曲线如图 3 所示。达到稳定状态需要的平均填充时间为 54.4 s。

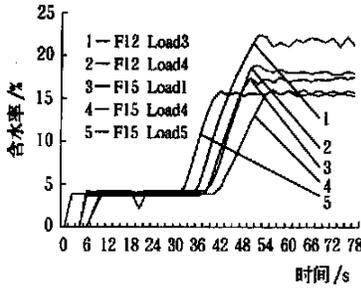


图 1 空车开始收割时小麦含水率的变化曲线

Fig 1 Winter wheat moisture curve when empty combine beginning to harvest

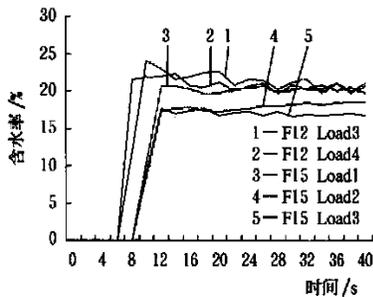


图 2 收割机调头后再收割时小麦含水率的变化曲线

Fig 2 Winter wheat moisture curve when combine turning around at headland and beginning to harvest

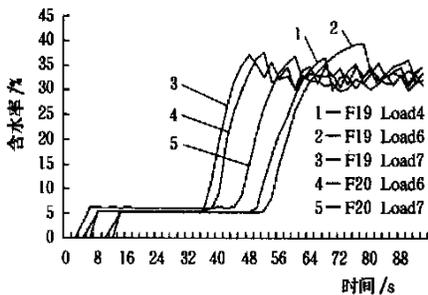


图 3 空车开始收割时玉米含水率的变化曲线

Fig 3 Maize moisture curve when empty combine beginning to harvest

收割机地头调头再进地收割的情况下, 玉米含水率数据的变化曲线如图 4 所示, 达到稳定状态需要的平均填充时间为 22 s。与收割小麦时的情况相同, 收割机空

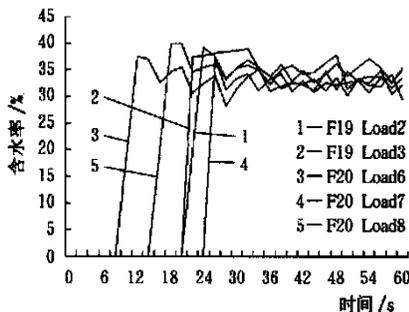


图 4 收割机调头后再收割时玉米含水率的变化曲线

Fig 4 Maize moisture curve when combine turning around at headland beginning to harvest

车收割情况下的玉米含水率填充时间, 要比收割机地头调头再进地收割情况下的玉米含水率填充时间长一些。

3) 小麦产量的填充时间

收割机空车收割情况下, 小麦达到平均产量值时, 需要的平均填充时间为 23.2 s。收割机地头调头再进地收割的情况下, 小麦达到平均产量值时, 需要的平均填充时间为 33.6 s。

4) 玉米产量数据的填充时间

收割机空车收割情况下, 玉米达到平均产量值时, 需要的平均填充时间为 13.2 s。收割机地头调头再进地收割的情况下, 玉米达到平均产量值时, 需要的平均填充时间为 26.4 s。

5) 填充时间误差数据的分析比较

不同情况下的平均填充时间如表 2 所示。通过比较与分析不同情况下的填充时间, 并结合 2366 收割机的机械结构与田间作业情况可知:

表 2 不同情况下的平均填充时间

Fig 2 Average filling time under different conditions

| 作物类型 | 收割情况 | 水分数据平均填充时间/s | 产量数据平均填充时间/s |
|------|-------|--------------|--------------|
| 小麦 | 空车收割 | 48.8 | 23.2 |
| 小麦 | 调头后收割 | 10.8 | 33.6 |
| 玉米 | 空车收割 | 54.4 | 13.2 |
| 玉米 | 调头后收割 | 22 | 26.4 |

1) 无论对于玉米还是小麦, 收割机空车收割情况下谷物含水率的填充时间, 比收割机地头调头再进地收割情况下谷物含水率的填充时间长。这是因为空车作业时, 收割的谷物必须先经过较长时间, 对收割机的凹板、滚筒与筛子等位置的空隙进行填充; 而收割机地头调头再进地收割时, 收割机中的各处空隙已填充有一些谷物, 此时需要的填充时间就相对短一些。

2) 收割机空车收割与收割机地头调头再进地收割情况下的玉米含水率填充时间分别为 54.4 s 和 22 s, 均分别大于两种情况下小麦含水率的填充时间 48.8 s 与 10.8 s。这是因为玉米的颗粒比小麦大, 进入谷物水分传感器比较困难, 所以需要较长时间才能达到稳定状态。

3) 无论对于玉米还是小麦, 谷物含水率经过填充时间达到稳定状态后, 其变化范围很小。这是因为在同一块农田中作物品种单一, 并且耕作方式相同, 所以田间不同位置作物的含水率差别不大。

4) 与谷物含水率填充时间的情况不同, 对于玉米和小麦收割机空车收割情况下作物产量的填充时间, 比收割机地头调头再进地收割情况下作物产量的填充时间短。这是因为作物产量与谷物流量以及谷物含水率的关系如下式(公式(1))所示, 其中 y_i 正比于 F_i , 并且 y_i 正比于 $(1 - M_i)$ 。在空车开始收割时, 虽然谷物流量累计值 F_i 较小, 但由于谷物含水率传感器还没有经过填充时间, 未进入正常工作状态, 所以此时谷物含水率的测量值 M_i 也很小, 比如: AFS 系统中的水分传感器读数在开始收割的一段时间内一直保持在 4% 左右。此时计

算出的产量值 y_i 就比较高, 从而造成了产量数据迅速增加, 并经过较短填充时间很快达到稳定值的假象。但是由于此时计算作物产量所使用的谷物含水率数据为填充时间误差数据, 所以作物产量的计算结果是不可信的。而收割机地头调头再进地收割情况下, 谷物含水率的填充时间较短, M_i 很快达到稳定值, 此时作物产量的计算值 y_i 只能随谷物流量累计值 F_i 的增加而逐步增加。所以在这种情况下经历的填充时间较长。

$$y_i = F_i \cdot (1 - M_i) \quad (1)$$

式中 y_i ——数据点 i 的产量值; F_i ——数据点 i 的谷物流量累计值; M_i ——数据点 i 的谷物含水率值。

5) 收割机空车收割与收割机地头调头再进地收割情况下的玉米产量的填充时间分别为 13.2 s 和 26.4 s, 均分别小于两种情况下小麦产量的填充时间 23.2 s 与 33.6 s。这是因为玉米产量一般高于小麦产量, 所以收割玉米时只需要经过较短填充时间作物产量即可达到稳定值。

6) 确定填充时间误差数据, 需要同时考虑谷物含水率与作物产量的填充时间, 并以同一情况下较长的填充时间为准。例如: 在表 2 中, 小麦空车收割情况下, 谷物含水率的填充时间为 48.8 s, 而作物产量的填充时间为 23.2 s, 此时应将前 48.8 s 内的产量数据作为填充时间误差数据。而玉米调头后再收割情况下, 应将前 26.4 s 内的产量数据作为填充时间误差数据。

不同情况下的产量数据中填充时间误差数据所占的比例情况如表 3 所示。其中地块 F12、F15 与 F19 的产量数据中填充时间误差数据的百分比均在 11% 左右, 而地块 F20 的填充时间误差数据却占 14.61%。这是因为在收割地块 F20 时, 收割机临时停车与调头次数较多所致。

表 3 填充时间误差数据比例

| 地块名称 | 作物类型 | 产量数据点个数 | 误差数据点个数 | 误差数据百分比/% |
|------|------|---------|---------|-----------|
| F12 | 小麦 | 4448 | 471 | 10.59 |
| F15 | 小麦 | 3747 | 428 | 11.42 |
| F19 | 玉米 | 4603 | 478 | 10.38 |
| F20 | 玉米 | 5638 | 824 | 14.61 |

3.3 延迟时间误差

当收割机停止收割后, 作物产量逐渐减小到零所需的时间称为延迟时间, 这段时间可分为两部分: 一是滞后时间, 即从割台升起后谷物流量仍保持稳定的时间; 二是清空时间, 即作物产量值急剧下降直到读数为零的时间。滞后时间内所记录的产量数据是有效的, 因为正常情况下作物收割后都需要经过一段时间才能从割台到达产量传感器。滞后时间的长短除了受作物类型等固定因素的影响外, 还受作物产量与谷物含水率等变化因素的较大影响。在清空时间内, 车内余粮逐渐到达产量传感器, 产生产量数据。但这部分数据并不能代表田间作物的产量, 所以是无效的。在作物产量开始急剧下降的时刻, 滞后时间结束, 测产系统应停止记录数据。目前

有些测产系统中, 例如 CASE 公司的 AFS 系统中, 已经设置了固定滞后时间, 但用户不能对其进行修改。

表 4 为不同情况下延迟时间误差数据所占的比例。从表中可以看出不同情况下, 延迟时间误差数据在产量数据中所占的比例都很小, 均不足 1%。这主要是因为 AFS 测产系统中已经设置了固定滞后时间, 对延迟时间误差进行了处理。但是由于滞后时间受许多因素影响, 而系统中设置的滞后时间为固定值, 用户不能根据不同的实际使用情况对其进行修改, 所以造成原始数据集中仍存在少量的延迟时间误差。

表 4 延迟时间误差数据比例

Table 4 Proportion of raw data recorded in delayed time error

| 地块名称 | 作物类型 | 产量数据点个数 | 误差数据点个数 | 误差数据百分比/% |
|------|------|---------|---------|-----------|
| F12 | 小麦 | 4448 | 38 | 0.85 |
| F15 | 小麦 | 3747 | 29 | 0.77 |
| F19 | 玉米 | 4603 | 33 | 0.72 |
| F20 | 玉米 | 5638 | 47 | 0.83 |

4 误差数据处理对产量分布图的影响

对地块 F14 的小麦产量数据, 不进行误差数据处理, 采用距离反比加权方法进行插值, 栅格大小与搜索半径取为 20 m, 产量数据单位为 t/hm^2 , 所生成的产量分布图如图 5a 所示; 而对同一组数据, 使用 EXCEL 软

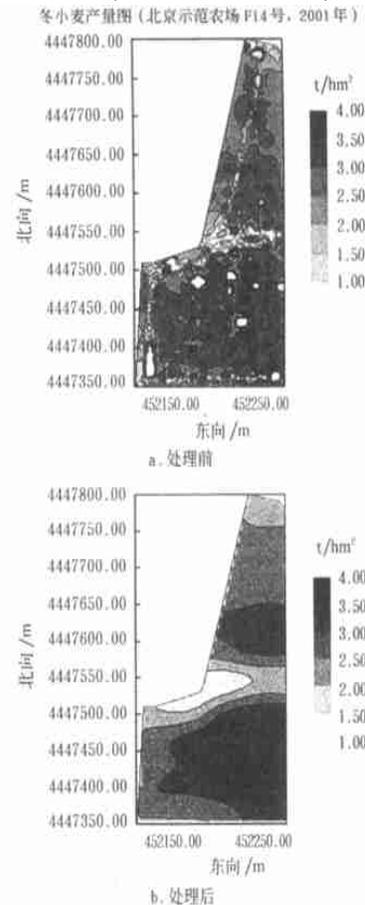


图 5 误差数据处理前、处理后的产量分布图

Fig. 5 Yield map before and after error data processing

件进行了割幅宽度设置误差、填充时间误差与延迟时间误差处理之后,再采取同样的方法进行插值,所生成的产量分布图如图 5b 所示。比较两图可以看出,误差数据处理后,产量分布图中的奇异点显著减少,数据的平滑性与聚类性得到了明显改善。

5 结 论

通过对产量数据进行分析和处理,识别并剔除了产量数据中包含的三类主要误差。其中,割幅宽度设置误差数据所占的比例,与地块形状和收获的作物类型有关,一般小于 6%;填充时间误差数据所占的比例最大,一般大于 10%;延迟时间误差数据所占的比例小于 1%,因为制造商已经对延迟时间误差进行了处理。在剔除了包含割幅宽度设置误差、填充时间误差和延迟时间误差的产量数据点后,生成的产量空间分布图质量有了显著的提高。

[参 考 文 献]

- [1] Panneton B, Brouillard M, Piekurowski T. Integration of yield data from several years into a single map[A]. Proc of the 3rd European Conf on Precision Agriculture[C], S. Blackmore, G. Grenier (Ed), June 18~ 20, 2001, Montpellier, France
- [2] 汪懋华“精细农作”技术发展与创新系列讲座,第六讲:“精细农作”的主要支持技术(四),田间信息采集与处理技术[J]. 农业机械化, 1999, 7.
- [3] Missotten B, Strubble G, J. De Baerdemaeker. Accuracy of grain and straw yield mapping[A]. Proc of the 3rd Int Conf on Precision Agriculture[C], P. C. Robert, R. H. Rust, and W. E. Larson (Ed), 23~ 26 June, 1996, Minnesota, France
- [4] Murphy P, Schnug E, Haneklaus S. Yield mapping-A guide to improved technique and strategies[A]. Proc 2nd International Conference on Site-Specific Management for Agricultural Systems[C]. University of Minnesota, March 27~ 30, 1994
- [5] Mark Moore. An investigation into the accuracy of yield maps and their subsequent use in crop management[D]. Denmark: Department of Agriculture and Biosystems Engineering, Cranfield University, 1998
- [6] Andy D Beck, Jody P. Roades, Stephen W. Searcy. Post-process filtering techniques to improve yield map accuracy[Z]. ASAE meeting presentation, July 18~ 21, 1999, Toronto, Ontario, Canada, Paper No. 991048
- [7] CASE IH Corporation (2000): AFS Universal Display and Display Plus Operators Manual Harvesting Operations[Z]. Version 3, 2000
- [8] 张漫,李相平,邝继双等. CASE 2366 谷物联合收割机 AFS 系统的使用及数据处理研究[A]. 浙江省第二届青年学术论坛文集[C]. 北京: 中国科学技术出版社, 2001.
- [9] Kuang Jishuang, Zhang Man, Li Xiangping, et al. Yield Mapping Using Case 2366 Harvester with AFS2000[A]. International Conference on Agricultural Science and Technology, Session 6: Information Technology of Agriculture[C]. November 7~ 9, 2001, Beijing, China, 342~ 344

Error analysis and process of yield data collected by yield monitoring system of corn combine harvester

Zhang Man, Kuang Jishuang

(Key Lab of China Ministry of Education for Modern Precision Agriculture System Integration, China Agricultural University, Beijing 100083, China)

Abstract: A yield map represents the basic information for precision agriculture and a site-specific crop management. Its accuracy is extremely important. Yield data processing and error analysis are key approaches to improving the reliability of the yield map and minimizing the influence of error on the yield data. Based on the first two years' practice on a grain combine with a yield monitoring system, this paper contributes to the yield data error analysis and improvement of the yield map. Three kinds of main errors of yield data were described: cutter width setting error, filling time error, and delayed time error. The paper discussed the error occurrence, recognition and processing. Errors were compared under different harvesting conditions of corn and winter wheat. The percentage of cutter width setting error data in the yield data set is less than 6%, the percentage of filling time error data in the yield data set is more than 10%, the percentage of delayed time error data in the yield data set is less than 1%.

Key words: yield monitoring system; error analysis; yield map; corn combine harvester