Using Personal Digital Assistants and Database Forms as Tools for Field Data Collection: Practical Experiences from Research Trials

Arthur Villordon¹,
Jason Franklin², and
Don LaBonte³

ADDITIONAL INDEX WORDS. handheld computing devices, desktop database, nonvolatile memory

SUMMARY. Handheld computing devices, such as personal digital assistants (PDAs), can potentially reduce repetitive tasks that pervade data collection activities in horticultural research. PDA-collected records are electronically transferred to a desktop computer, eliminating manual reentry as well as the need of reviewing for incorrect data entries. In addition, PDAs can be enclosed in protective cases, enabling data collection in inclement weather. Visual CE-generated database forms installed on PDAs were used to electronically collect data from research trials conducted in 2003. The records were subsequently transferred to Microsoft Access desktop database tables for archiving and subsequent statistical analyses. Data for certain trials were also manually collected

using paper forms to facilitate comparison between manual and PDA-assisted data collection methods under controlled conditions. Using paired samples analysis, we determined that electronic transfer of records reduced the time required to store the records into desktop computer files. Manual and PDA-based recording methods did not vary in the time required to enter numerical measurements. Our experience demonstrates that off-the-shelf software and consumer PDA devices are viable options for data collection in research. PDA-assisted data collection is potentially useful in situations where remote, site-specific records need to be merged into a central database and where standardized measurements and observations are essential for performing analysis.

Handheld computing devices can automate repetitive tasks that permeate various phases of data collection in horticultural research. Specifically, measurements and observations from field research trials are traditionally recorded onto paper data sheets followed by the transfer of these records into a computer file for archiving and statistical analysis. The latter step involves the reentry of previously recorded information from data sheets and reviewing for missed or incorrect entries, an inefficient process that can be time consuming and inaccurate (Fletcher et al., 2003; Spain et al. 2001). Lehman-Salada et al. (1995) described how keyboard-equipped palmtop computers installed with spreadsheet software were useful alternatives to notebook computers and other data acquisition devices for gathering plant data in the field. Recent technological innovations have led to more compact computing devices as well as the availability of electronic forms software that facilitate data collection. These devices include PDAs, sub-notebook computers, and tablet computing devices. The use of portable touch-screen or pen-based PDAs is already widespread in industry, particularly in the healthcare sector (Karampelas et al. 2003; Torre and Wright, 2003). Regardless of the operating system that runs these devices, it has been shown that database entry forms simplify the data collection interface by providing data entry screens to specific records and fields in databases or spreadsheets. For example, Porter (2001a, 2001b) tested an elec-

tronic field data collection program and demonstrated how data can be rapidly summarized using a spreadsheet once the PDA is synchronized with a desktop computer. This system used a consumer PDA and off-the-shelf software to develop forms that facilitated field data collection.

This report documents the use of PDAs in horticulture field data collection. We describe the procedures and the computing hardware and software components that we used to integrate PDA-assisted data collection in an active research program. In this report, we define data collection as comprising two phases. The first phase involves the recording of numerical and descriptive observations (REC). This step can be performed in the field, greenhouse, orchard, remote site, or in the laboratory. The second phase involves the transfer or encoding of previously collected data into computer spreadsheet or database formats for storage and archival purposes (TRA).

Materials and methods

Microsoft Access 2002 (Microsoft, Redmond, Wash.) desktop database software was used to: 1) create tables that define the fields required for data collection; 2) ascertain that synchronization between the desktop and PDA tables has been successful; and 3) compile data for archiving and statistical analysis. Data entry forms for each table were created using Visual CE Professional Edition (version 7; Syware Inc., Cambridge, Mass.), a PDA database and forms development software (FDS) for the Windows CE operating system (Microsoft). The database fields, e.g., U.S. #1 weight, row number, or entry name, defined for the electronic data collection forms (EDF) were identical to the corresponding printed paper data collection forms (MDF) that have been previously used for field data collection. Defining the field type (e.g., alphanumeric, numeric, date and time) is an important step in creating database tables as it enforces data entry rules as well as the calculations possible for a specific field. Briefly, each FDS was created using the following steps: 1) the FDS "form wizard" automatically generated data entry boxes and corresponding labels based on the previously defined database fields; 2) if necessary, the default data boxes and labels were modified; and 3) addition of navigation buttons (e.g.,

^{1,2}Louisiana State University Agricultural Center, Louisiana Agricultural Experiment Station, Sweet Potato Research Station, P.O. Box 120, Chase, LA 71324-0120.

³Department of Horticulture, Louisiana State University, Baton Rouge, La 70803.

Approved for publication by the director of the Louisiana Agricultural Experiment Station as manuscript no. 04-66-0128. Mention of a trademark, proprietary product, or vendor does not imply endorsement by the Louisiana State University Agricultural Center nor its approval to the exclusion of other suitable products or vendors. We thank Paul Wilson for reviewing an early draft of the manuscript and three anonymous reviewers for providing helpful comments. We acknowledge Syware Inc. for donating a Visual CE Personal Edition license for initial evaluation and testing.

¹Assistant professor; to whom reprint requests should be addressed. E-mail: avillordon@agctr.lsu.edu

²Research associate.

³Professor.

“previous record,” “next record,” and “jump to”). The modifications included revising descriptive data entry boxes to “drop-down” list boxes, font size and color, and background color. EDFs were subsequently downloaded to a secure digital (SD) memory card for data collection. The SD card retains data even when the device is turned off, providing us an option of transferring the card to a secondary PDA unit in case the main unit failed. We collected data from 13 research trials that were conducted in 2003 at the Sweet Potato Research Station, Chase, La. These trials involved measurements or descriptions of vine traits (vine length, diameter, internode length), petiole nitrate and potassium readings, and storage root yields, skin and flesh color, as well as other observations. The number of database fields and tables created per trial varied, depending upon the measurements or observations planned for each respective trial.

A Dell Axim Advanced (Dell Inc., Round Rock, Texas) PDA with Pocket PC 2002 (Microsoft), was designated as the primary data collection unit. The specifications of this PDA included the following: 400 megahertz (mhz) Intel XScale processor (Intel Corp., Santa Clara, Calif.), 64 megabytes (MB) synchronous dynamic random access memory (SDRAM), 48 MB read-only memory (ROM), color 16-bit trans-reflective liquid crystal display (LCD), and one slot each for a compact flash (CF) Type II card and an SD card. All record synchronization activities were performed with the built-in synchronization software (ActiveSync; Microsoft). A second PDA (Compaq Aero 1550; Hewlett-Packard Co., Palo Alto, Calif.) was designated as a backup unit. A protective case (Armor 3600; Otter Products, LLC, Fort Collins, Colo.) that allowed the use of the PDA through a clear plastic screen was used in all field data collection activities. The cost of the primary PDA, FDS, and protective case was about \$700. All synchronization activities were performed with a host desktop computer running Windows 2000 Professional (Microsoft) with the following specifications: dual Intel Pentium III 1.0 gigahertz processors, small computer system interface (SCSI)-based system drive, and 512 megabytes PC133 SDRAM system memory. Before using the PDA, we simulated all phases of data collection activities to ascertain

that all records were transferred during the synchronization process.

Comparison between “pen-and-paper” (PPR) and pda-based (PDA) data collection methods was facilitated by using preprinted MDFs from representative trials. To test for differences in REC time between PPR and PDA collection methods, five users recorded a representative data set (RDS, 88 records, five database fields) using both methods in a paired samples analysis. Each user recorded the RDS three times for each data collection method. The means for each user were used in the analysis. The length of time for PPR and PDA to complete TRA was also compared. For PDA TRA, synchronization of other items like calendars, contacts, tasks, offline files, and other database tables were turned off during the file uploads. This ensured that only the test database table’s synchronization time was measured. For PPR TRA, users recorded the RDS onto MDFs. All REC trials were repeated three times and the means were used for analysis. Analysis of REC and TRA time differences were performed using Wilcoxon matched-pairs signed-ranks test. For all comparisons, we assumed that the test subjects accurately entered all records. All computations were performed using the SAS System for Windows (Version 8.1; SAS Institute, Cary, N.C.). Users were given instructions on how to use the PDA, navigate the EDF, and perform data entries.

Results and discussion

A total of 1809 records were generated for 13 field trials, with the number of database fields ranging from 7 to 17. Overall, the PDA-assisted data collection process reduced the time required for making field-collected data available for analysis, with TRA accounting for much of the difference (Table 1). Under controlled conditions, PDA significantly differed from

PPR-based TRA (mean of 22 s vs. 16 min 28 s, respectively) ($P < 0.01$). PDA and PPR methods did not exhibit differences in REC time (mean of 31 min vs. 17 min, respectively) ($P = 0.25$). Assuming that we only used five database fields for our trials and a proportional increase in PDA TRA time, PPR-based data collection would have required about 7 h or more to complete transfer of records into a spreadsheet or database table. The length of time required to develop, install, and test the EDFs ranged from 30 min to 4 h, depending upon the number of fields and whether external records or tables needed to be referenced. For example, the EDF shown in Fig. 1A was ready for data collection in about 30 min. This EDF’s underlying database comprised five database fields (row, replication, U.S. #1, canner, and jumbo). In contrast, a linked system of three EDFs designed to collect data from a relatively large breeding trial (over 100 entries with three replications) required about 4 h of development and testing (Fig. 1C). Due to the size of the trial, we used one EDF for each replication. Each EDF represented a database table with 15 database fields. All EDFs that we created are potentially reusable, eliminating development and testing time for future data collection activities.

We did not detect any missed or dropped records resulting from PDA synchronization. Similar results were reported by Fletcher et al. (2003) and Porter (2001b) when they compared paper-based and PDA-based data collection methods. Most of our database tables were exported as text files for statistical analysis purposes. Alternately, current versions of desktop statistical software [e.g., SAS (SAS Institute)] can directly import database and spreadsheet formats for analysis. PDA-assisted data collection enforces data integrity due to the fact that data are recorded only once followed by direct transfer

Table 1. Comparison of recording (REC) and data transfer (TRA) times using personal digital assistant-assisted (PDA) and paper-based (PPR) data collection methods.^a

	Time (min:s)		<i>P</i>
	PDA	PPR	
REC	31:01 ± 2:01	17:01 ± 0:38	0.25
TRA	0:22 ± 0:45	16:98 ± 1:20	<0.01

^aPDA and PPR times are means of three replications and $\pm$ SE of five test users. Tests for differences were performed using the Wilcoxon matched-pairs signed-ranks test.

row 1

rep 8

us1 53.500000

can 22.000000

jum 11.000000

<< >>

Record Edit Option

53.500000

7	8	9
4	5	6
1	2	3
+ / =	0	.

Clear

<< >>

Record Edit Option

Test Loc DPlanted DHarvest

B03 MLU 9/10/03 9/10/03

Entry 2-1 Rep 2

US#1 Canner Jumbo

0.00000 0.00000 0.00000

keep shape size skin flesh

☒ medic small coppe orange

Rep 2

Rep 3

short

Rep 1

Rep 2

< >

Record Edit Option

Fig. 1. Sample personal digital assistant (PDA) database forms that show labels, data entry boxes, pop-up numerical keyboard, and various navigation controls. PDA database forms were created using Visual CE (version 7; Syware Inc., Cambridge, Mass.).

to desktop database tables, without further need of manual reentry or encoding. Furthermore, full-featured FDS development applications can be configured to delete data from the PDA following synchronization, preventing accidental overwriting of previously recorded data.

A few challenges exist for using mass-market or consumer PDAs in field data collection. These include functionality under harsh field conditions, small screen size, and the potential for data loss due to battery or equipment failure (Fletcher et al., 2003). Although "ruggedized" or "field" PDAs are commercially available, these are usually at least four times more expensive than their consumer versions. We used a commercially available and relatively inexpensive protective case that allowed us to operate the PDA in extremely dusty conditions. To address the perceived disadvantage of the small LCD screen size, we used combo boxes (a data entry box with a drop-down list containing predefined values), "lookup tables," and context sensitive virtual keyboards that helped promote efficient user interaction (Navarette, 1999). Figure 1A shows how we designed an EDF to simplify storage root weight data collection. The layout of the data entry boxes, along with the numeric pop-up keyboard (Fig. 1B), enabled users to perform data entry and navigate to the next record intuitively. Figure 1C shows how we designed an EDF that permitted us to look up records from another replication or block. The use of paper records in this example would have required the researcher to manually locate the records on the field plot plan or tabular data that contained over a hundred entries. To address equipment failure issues, we designated a second PDA as a standby unit in case our primary device failed. Had our primary device failed, the nonvolatile storage card containing the EDFs could be reinserted into the secondary PDA. We did not experience battery-related problems during our data collection activities. Spare batteries can be purchased in environments that require prolonged data collection activities. During our data collection activities, we occasionally observed glare from the protective screen, especially during bright light conditions. Although this condition did not prevent data entry, it had the potential to slow down the

data collection process. We reduced the potential for glare by changing the EDF background color from the default setting of white to grey.

Our experience demonstrates that relatively inexpensive off-the-shelf software and consumer PDA devices are viable options for data collection in research. The selection of a PDA operating system [Pocket PC/Windows CE (Microsoft) vs. Palm OS (Palm-Source Inc., Sunnyvale, Calif.)] and forms development software for use in data collection will largely depend on prevailing institutional computing support policies. Based on our trials, the minimum criteria for a PDA to be a viable substitute for paper forms include the following: availability of non-volatile memory for storing data, synchronization with various desktop operating systems either through built-in or third-party software, and at least 3 h of sustained operation on battery power. Current PDA processor speeds range from 16 to 400 MHz. Models suitable for data collection are in the 100–300 MHz range (McCombs, 2004). Although the majority of current PDA models use SDRAM as the main device memory, some models have built-in non-volatile memory (NVM)-based flash memory that enables stored data to persist without power. The FDS applications available for the Pocket PC and Palm OS operating systems do not require users to be proficient in computer programming. However, some knowledge of desktop database forms design will be useful. In particular, experience with creating, navigating, and using forms within the Access (Microsoft) software or similar desktop database applications will provide sufficient skills in creating and maintaining PDA-based database forms.

PDA-assisted data collection is potentially useful in data collection activities that involve cooperators in several locations. EDFs can be configured to standardize data entry through "drop-down" lists, numerical data range constraints, and "check boxes." Database tables and EDFs can be generated in a central location, installed on PDAs, and then deployed to remote locations. The central database can then be updated by directly importing remote tables or using remote synchronization through built-in functions or third-party applications. The FDS that we used, as well as similar applications that are currently

available, interoperate with Microsoft Access as well as other open database connectivity (ODBC)-enabled software and has built-in functions for accepting inputs from barcode readers, global positioning system (GPS) devices, as well as other inputs via the serial port. These features allow the incorporation of electronic measuring devices to further enforce data integrity. For instance, the availability of wired or wireless weighing scales, barcode readers, calipers, data loggers, probes, and similar devices allow the direct recording of measurements into the PDA-installed electronic forms, eliminating the need for any type of manual entry.

PDA-assisted data collection reduces repetitive tasks associated with manual data collection, enforces data integrity, and significantly decreases the time the data are made available for analysis and archiving. PDA-based electronic forms are a viable alternative to paper-based forms for improving efficiency in data collection, most especially in collaborative research activities that require several cooperators to perform repeated data collection in several locations.

Literature cited

Fletcher, L.A., D.J. Erickson, T.L. Toomey, and A.C. Wagenaar. 2003. Handheld computers—A feasible alternative to forms for field data collection. *Evaluation Rev.* 27:165–178.

Karampelas, P., D. Akoumianakis, and C. Stephanidis. 2003. User interface design for PDAs: Lessons and experience with the WARD-IN-HAND prototype. *Lecture Notes Computer Sci.* 2615:474–485.

Lehman-Salada, L., K. Hickey, and T. Salada. 1995. Using palmtop computers to collect plant data in the field. *Hort-Technology* 5(2):159–160.

McCombs, B. 2004. Personal digital assistants. *Can. J. Rural Medicine* 9:45–46.

Navarette, G. 1999. In the palm of your hand: Digital assistants aid in data collection. *J. Mgt. Eng.* 15:43–45.

Porter, M. 2001a. Field testing HanDBase 2.5 on the Palm III for data collection. *Computer User Section, Amer. Fisheries Soc.* 15:2–3. 21 Apr. 2004. <<http://www.fisheries.org/cus/newsletters/CUS-news-S2001.pdf>>.

Porter, M. 2001b. Summarizing field data rapidly using a spreadsheet. *Computer User Section, Amer. Fisheries Soc.* 15:3–4. 21 Apr. 2004. <<http://www.fisheries.org/cus/newsletters/CUS-news-S2001.pdf>>.

Spain, K.A., C.A. Phipps, M.E. Rogers, and Barbara S. Chaparoo. 2001. Data collection in the palm of your hand: A case study. *Intl. J. Human-Computer Interaction.* 13:231–243.

Torre, D.M. and S.M. Wright. 2003. Clinical and educational uses of handheld computers. *Southern Medical J.* 96:996–999.