Chapter 3

What We Have Learned about Field Programs

MARGARET A. LE(MONE

National Center for Atmospheric Research,* Boulder, Colorado

ABSTRACT

Based on personal experience and input from colleagues, the natural history of a field program is discussed, from conception through data analysis and synthesis of results. For convenience, the life cycle of a field program is divided into three phases: the prefield phase, the field phase, and the aftermath. As described here, the prefield phase involves conceiving the idea, developing the scientific objectives, naming the program, obtaining support, and arranging the logistics. The field phase discussion highlights the decision making process, balancing input from data and numerical models, and human interactions. The data are merged, analyzed, and synthesized into knowledge mainly after the field effort.

Three major conclusions are drawn. First, it is the people most of all who make a field program successful, and cooperation and collegial consensus building are vital during all phases; good health and a sense of humor both help make this possible. Second, although numerical models are now playing a central role in all phases of a field program, not paying adequate attention to the observations can lead to problems. And finally, it cannot be overemphasized that both funding agencies and participants must recognize that it takes several years to fully exploit the datasets collected, with the corollary that high-quality datasets should be available long term.

1. Introduction

Because my current work (on the diurnal cycle of the fair weather boundary layer over land) seemed a bit peripheral to the rest of the talks on the program, I asked Joanne what she would like to hear about—then at least one person would be interested. She suggested that I talk about field programs, and what we have learned. I decided to take a different tack on the subject (as she might have expected) and speak more about the field programs themselves than about the science that we have learned.

2. The stages of a field program

I will describe the stages of a field program from its inception as an idea to the most important phase, the data analysis and synthesis, using cartoons and experiences drawn from several field programs.

a. Prefield phase

1) Origin

Figure 3.1 idealizes two very different origins for a field program. In the first, an august committee of scientists of impeccable reputation and immense funding decree that a field program be conducted to answer a Big Question, which will in turn guarantee them further immense funding and exposure on the Cable News Network (CNN). In this case, a group of scientists (identified here as “the usual suspects”) are called in to fill in the details (Fig. 3.2). In the second example (Fig. 3.1b) a group of scientists (some august, some only January) get together and try to attack a question that has stumped them. They develop a hypothesis and design a modest field program to address the hypothesis. In this case, reviewers often point out that either (a) the dataset that the investigators want to collect is subcritical, so additional principal investigators (PIs) and/or instrumentation are needed, or (b) the funding agencies or facilities managers decide that the facilities needed (radars, aircraft, etc) are so expensive and the opportunities so large that additional PIs are needed. Once again, the usual suspects (Fig. 3.2) are brought in.

2) Determining the scientific objectives

The field program that survives the origin/ review stage is further designed in community planning meetings where someone constantly reminds everyone else that “We can’t design the field program without a hypothesis!” During such meetings, the scientific objectives morph into something often somewhat different (read “bigger”) from the original ones. This stage is illustrated in Fig. 3.3, drawn during a planning meeting sponsored by the Stormscale Operational and Research Meteorology (STORM) program in 1988; but the au-
dience comments can be applied to most field programs. The expansion problem is compounded when the signal (real or imagined) that funding might be available gets out to the community. This leads to a snowball effect (even for tropical experiments), which increases the attendance to such planning meetings and the potential for broadened objectives.

In other cases, a field program is sufficiently similar to an earlier one that the PIs closely examine the objectives of earlier programs, to make sure everything is covered or to avoid being redundant. This is illustrated in Fig. 3.4, drawn during the early planning stages of the Monsoon Experiment (MONEX).

3) THE NAMING OF THE EXPERIMENT

Atmospheric scientists are masterful acronym creators, not only creating first-order acronyms, such as the Tropical Rainfall Measuring Mission (TRMM), but second-order acronyms, such as FIFE (First ISLSCP Field Experiment, where ISLSCP is the International Satellite Land Surface Climatology Project), and third-order acronyms such as GAUC (GATE Aircraft Utilization Committee, where GATE is the GARP Atlantic Tropical Experiment, and GARP is the Global Atmospheric Research Program). The ideal acronym itself describes an aspect of the experiment [like STORM or Boreal Ecosystem Atmosphere Study (BOREAS)] and, even before this era of political correctness, care was taken to create acronyms that offend no one. GARP for example, was selected because it was not an offensive word in any language, at least until publication and filming of The World According to Gar at years later. Thus some field program names changed (FOPS to FAPS and SCUM to...
FIG. 3.4. MONEX: “In Search of Objectives.”

SCMS), but some genuinely humorous names have survived (BARFEX). BARFEX, by the way, stands for Boardman Atmospheric Radiation Flux Experiment.

4) OBTAINING THE NEEDED INSTRUMENTATION/MONEY

This step is where reality sets in. The experiment organizers, with the help of more community planning meetings, rough out a design that describes objectives, location, timing, instrument platforms, and their use. Through this stage, the number of PIs entrained through interest in the subject matter, concentration of useful measurements, or a funding opportunity remains large, an advantage when going to the agencies ("347 scientists want to use the data from your new super airborne lidar"). Often, the needed facilities must come from more than one government agency, or even more than one government, each of which has its own goals, timetable, and application procedure. Potential participants are thus deeply involved in simultaneous applications for funding and facilities, and putting together an experiment plan despite the uncertainties. Once the applications and proposals are in and reviewed, the final deployment is outlined in an iterative procedure that usually fits the facilities to the experimental objectives, and that usually ensures that the PIs responsible for the needed platforms are participants in the experiment. Note: even though the number of PIs and instruments may be reduced in this phase, the objectives usually remain untouched.

Unexpected factors can make the planning exercise even more interesting at any time. Timing and even location can shift. Timing is usually shifted to give worthy scientists access to needed instruments. Location can be trickier. Veterans of GATE recall that it was originally planned for the tropical Pacific, until someone discovered that it was impossible to pronounce “GPTE.” Smaller experiments can also get shifted in location. Just months before the field phase of an experiment I was involved in, by which time we were well on our way to securing leased sites for surface instruments, a representative of the National Science Foundation (NSF) strongly suggested that we shift our experiment site, since a certain agency (to remain unnamed) wanted our instruments to beef up one of its field programs. We stayed where we were, because a third agency already had its instruments in place, and we weren’t prepared to think up a new acronym!

5) LOGISTICS: PRACTICE, PRACTICE, PRACTICE, AND HEALTH ISSUES

Getting people and instruments coordinated is a big task. In GATE, the Preliminary Regional Experiment for STORM (PRESTORM), and Mesoscale Alpine Programme (MAP) there were actually rehearsals to teach scientists and air crews to work together in coordinating multiple aircraft flights. Smaller field programs have a few shakedown days at the beginning of the field phase. These shakedowns are valuable, because planners addressing the Big Questions sometimes forget how much trouble the average person (and average scientist) has with minutiae that can ruin a potentially perfect intensive operations period. When there are so many things to consider, even time and place can be difficult to straighten out. In this respect, the planners of GATE were brilliant in locating the experiment in the Greenwich time zone, so that we could live and collect data in UTC, and mercifully most activity was restricted to the Northern Hemisphere. On the other hand, participants in the Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE) had to simultaneously deal with three time zones (Townsville, Australia, time; Honiara, Solomon Islands, time; and UTC) and latitudes north and south of the equator.

One of the special attractions of tropical experiments is the possibility of getting malaria and a host of other interesting and exotic diseases. Before GATE, it wasn’t certain whether the shots or the warnings were worse. Fig. 3.5 illustrates some of the potential dangers, based on a briefing before TOGA COARE. Unfortunately, there were even more dangers than we originally realized, resulting from the antimalarial medication Larium. As the experiment ensued, we became experts on the

---

1 Actually, the decision was made for political reasons. For details, see online at http://uniblab.atmos.ucla.edu/tropic/newsletters/newsletter28.html.
expected side effects ("vivid dreams,"\(^2\) and "mental instability," or curious behavior even for scientists\(^3\). If someone was being particularly argumentative on a particular day, we would ask him/her whether they had just taken their Larium tablet, suspecting them of suffering from PMS (Post Mefloquine Syndrome, which lasted 24–48 h after ingestion). Unfortunately, a new form of (fortunately mild) malaria resistant to Larium infected several TOGA COARE participants.

**b. The field phase**

1) **KEEPING TO THE SCIENTIFIC OBJECTIVES**

Anyone familiar with large field programs know that the scientific objectives become so ambitious (or so diffuse) that almost no one can remember them. This inspired the cartoon drawn by Josh Holland during GATE (Fig. 3.6), probably after a debate about priorities.

2) **RUNNING THE FIELD PROGRAM**

Field programs, unlike armies, are run by consensus, which is sometimes difficult to obtain, particularly with ambitious objectives and instrumentation that can interfere with other instruments or flight plans that can interfere with other instruments or other flight plans. Developing a consensus becomes particularly difficult for experiments run using more than one operations cen-

\(^2\) Some of the dreams were so entertaining, the name "hi-larium" might be more appropriate.
\(^3\) Similar symptoms can originate from trying to keep track of the three time zones and northern vs southern latitudes while conducting aircraft missions.
CHAPTER 3

LEMONE

This was done in the Taiwan Area Mesoscale Experiment (TAMEX) (1987; i.e., pre-internet), with communication by phone or fax between Taiwan and Okinawa, or by radio between Taipei and the P3 aircraft. TAMEX had fairly straightforward objectives, making multiple centers doable. When field programs have complex objectives and multiple centers [e.g., STORM Fronts Experiment Systems Test (STORM FEST) and TOGA COARE], the lack of face-to-face communications leads to misunderstandings and even resentment.

Figure 3.7, which shows a difficult COARE conference call between Townsville and Honiara illustrates the “hearing but not listening” phenomenon that could not be solved even through the internet. The problem in GATE and COARE was compounded by the isolation of ships, which had different instruments and operational considerations from the aircraft. Television communication, used in STORM FEST, worked slightly better.

Lack of food, lack of sleep, and sometimes illness or simply the wrong kind of weather can also lead to conflict, as illustrated by Fig. 3.8, drawn during TOGA COARE. Figure 3.9 was prepared during STORM FEST, during which a stretch of adverse weather (i.e., “good weather” to anyone else) and lack of sleep led to such tensions that one student filled his journals with observations of human interaction instead of the pearls of wisdom he expected from map discussions.

In the Cooperative Atmosphere–Surface Exchange Study (CASES-97), we had a computer-generated sampler on the wall with words of wisdom on how to survive a field program. Things like “assume the other person is an idiot,” and its companion “don’t be offended if other people assume you are an idiot,” and “redundancy is good.” Watching out for one another and retaining our humility helped prevent forgetting the many details important to operating a field program.

Joanne Simpson taught me another good rule—“Don’t sweat the small stuff”—during GATE. This is illustrated by the following story, which requires a little background. GATE was in 1974, before the “politically correct” era. The walls of the radio room in the GATE Operational Control Centre (GOCC, another third-order acronym) were covered with Playboy centerfolds. The analysis room had fewer but more select nude women, including Chesty Morgan, whose endowment matched her name (think “watermelons”). One day, a colleague returning from the States gave a secretary (female—this was 1974) and me a copy of Playgirl. During a quiet moment, we examined the male centerfold (one African-American football player) and showed him to the Senegalese women who worked in the analysis room. Having already noticed the imbalance in female (many) versus male (zero) nudes and European (many) versus African (zero) nudes, they decided to rectify the situation by posting the Playgirl centerfold on the wall. This

4 This program had an interesting evolution, even though it was conducted in Kansas.
situation was most unacceptable to Colonel (Ret.) Barney (informally known as the GATE operations director) who immediately chewed out my supervisor (Ed Zipser), who was ordered to chew me out in turn for this flagrant violation of the rules. If the humor of this story is now obvious, it wasn’t at the time, and when I told the story to Joanne while walking out to the airplanes, she just said to ignore the small stuff and focus on the real problems, which I have tried to do since. (I must admit, though, that the women in the centerfolds were always better posed than the men.)

3) MAKING DECISIONS BASED ON DATA

We have more data now than we did three decades ago, but we probably made better use of the data back then, because more of us understood the limitations of the instrumentation, and more of us knew how to use our eyes. Today, with a proliferation of instruments and the use of computers, many are unaware or even uninterested in where the data come from. This led to a ludicrous situation, illustrated by Fig. 3.10, which really occurred in TAMEX.

TAMEX was conducted out of the Central Weather Bureau building in Taipei. TAMEX had the best ground–aircraft communication of any experiment I have participated in, truly remarkable given the mountains in Taiwan. But there was a bit too much focus on computers. We were on the sixth floor, which would have afforded a good view if the shades weren’t always drawn to make it easier to read the computer screens.

FIG. 3.9. Two scientists deciding on strategy and priorities for next intensive observing period. The reader is invited to fill in favorite field program, and two favorite scientists. From STORM FEST, 1992; based on figure generously provided by William Blumen.

To see outside, we had to walk to the stairwell. The decision to have an intensive observing period in TAMEX was based partially on the wind direction and speed—a good strong low-level jet slamming into the Island of Taiwan was ideal for setting up strong storms from the forced uplift or maintaining them through low-level shear as discussed by Rotunno, Klemp, and Weisman (1988). Weaker winds would result only in shorter-lived convectively generated storms. One day a colleague came into the TAMEX headquarters, very excited about the strong winds: cumulus were moving rapidly across the sky from the southwest and winds at the surface were gusty. However, the people inside denied the strong winds, saying that a radiosonde had just been sent up and the computed winds were light! I cannot recall whether eyeballs won out over the computer screen, but the eyeballs were right: it did rain. And rain.

All field programs are rich with examples of data misinterpretation, particularly when we are trying to use new types of data. In PRE-STORM (1985), we vectored the National Oceanic and Atmospheric Administration (NOAA) P3 to an apparent storm based on a cold cloud shield in the satellite infrared image. After all, several articles had been published showing strong precipitating convection under such cloud shields. A rather angry airborne scientist radioed us that there wasn’t even any rain under the cloud. We had a similar experience in TOGA COARE with the National Center for Atmospheric Research (NCAR) Electra.

c. After the field program

The most important phase of the field program is the combination and synthesis of the resulting data, a process that takes continuing coordination and funding.
The arguments for expecting fast data processing and analysis (not unique to COARE) are typically that faster computers can make data analysis faster. This is false for many reasons. First, new and more complex instruments with greater data volume have increased the time required for data processing and analysis. Second, focus on new instruments sometimes leads to less attention to older, “proven” instruments, as in the case of the COARE radiosondes. Third, datasets from diverse sources must be understood and made compatible, a problem compounded by broader mix of instruments, objectives, and people. Blending datasets is not always straightforward, as illustrated by Fig. 3.13, drawn while validating GATE data. Fourth, the increased sophistication of numerical models has led to a new sort of study that integrates model and data in exercises with the twin objectives of model verification and extending the data to study the physics. Finally, we aren’t any smarter than we were 30 years ago. In order to tease new knowledge out of data, we need time to think.

Post-experiment funding failure is not new. Ocean Enough and Time: Discovering the Waters around Antarctica, by Gorman (1995), describes a six-ship expedition led by Charles Wilkes that headed to Antarctica in 1839 to survey the coast (they established that Antarctica was a continent) and collect specimens. Only two ships returned, laden with specimens that were delivered to Louis Agassiz at Harvard. Agassiz could not complete analysis of the specimens, because there wasn’t enough money.

An additional problem, referred to earlier, is that the objectives of an experiment may have been frozen at their most optimistic maximum, before the funding, number of PIs, or instrumentation was finalized at a reduced level, the largest field programs probably have the most significant problem—objectives may have been determined in large, international meetings that are hard to repeat. Objectives outstripping resources leads to stretching resources in the field and in post-experiment analysis and synthesis. With smaller field programs, adjusting objectives is easier.

Finally, scientists share some of the blame for inadequate analysis of field data. Shortly after GATE, Joanne Simpson (Simpson 1976) noted the following.

Unfortunately, my experience has been that where most field experiments fail is in their follow-up, namely analysis and publication. The fault in the past has lain with both management, who finds other pressures on funds, and scientists, who find more glamour in running out to the field again, than in gluing their bottoms to a desk chair to carry through the painful and laborious scrutiny, corrections analysis, and reanalysis of data.

NASA has bucked this trend, perhaps because long time lags associated with even getting data from space probes requires longer-term thinking. Funding for analysis of FIFE (1987, 1989) has been at such a sufficient level that a special FIFE edition of the “Journal of Atmospheric Sciences” was published in 1998. Improvement of the land-surface parameterization in the ECMWF model can be directly traced to what was learned from FIFE (Viterbo and Beljaars 1995). Much effort was made to make the data available in usable form, so that even nonparticipants could benefit from the data. The data management plan for TRMM reflects a similar philosophy. One hopes that sustained funding for TRMM PIs is maintained for a sufficient time so that the data can be used to its fullest.5

5 Note that the author recommends this even though she is not a TRMM PI.
3. The use of models

I separate out the use of models because they are now used in all phases of a field program, from planning through execution and in postanalysis and synthesis. GATE was the first field program for which I saw participation of scientists involved in modeling and parameterization in the planning. While it was a struggle to bring these diverse groups together, GATE was designed to collect data for testing of both numerical models and convective parameterization schemes. Today, numerical models are used in deciding where to put instruments, how many instruments are needed to achieve important objectives, and what measurements are needed. Models are tested with the results of field programs. Once verified using observational data, models can extend the domain and fill in data gaps, making possible detailed studies of important physical processes.

Models can also be misleading. Here, I use the term “model” to refer to conceptual or theoretical models as well as numerical models. One of my favorite examples is in the book The Rejection of Continental Drift by Naomi Oreskes (1999), which relates several reasons why scientists in the United States were so tardy in accepting continental drift. In this case, the problem was accepting as true an assumption made by William Bowie [of American Geophysical Union (AGU) Bowie Medal fame] in a calculation. The calculation demonstrated to reasonable accuracy that land floating on a substrate could explain observed gravitational anomalies. The assumption was that the base of the earth’s crust was flat. The success of the calculation led American scientists to begin thinking that continents did have flat bottoms, rather than “roots.” It was hard to conceive of mantle currents dragging flat-bottomed continents around to produce continental drift! This was one of the factors that delayed the acceptance of plate tectonics in the United States.

The plate-tectonics episode is an example of an erroneous assumption taken too seriously. A more common problem occurs when models are taken too seriously, their assumptions and limitations ignored, misunderstood, or forgotten. My favorite example in meteorology relates to people’s conception of the behavior of the wind profile in the convective boundary layer. When I was a student in the late 1960s and early 1970s, the wind in the planetary boundary layer (PBL) was characterized by an Ekman spiral (Fig. 3.14), epitomized by the famous Leipzig wind profile (Mildner 1932). Much work at that time was devoted to estimating the appropriate eddy-exchange coefficient (e.g., O’Brien 1970), studying instabilities for the Ekman spiral (Faller and Kaylor 1966; Lilly 1966; Brown 1970), or estimating the departures from the Ekman spiral created by thermal wind (Gray and Mendenhall 1973).7

Around 1970, perhaps partially inspired by the temperature and moisture “mixed layers” Malkus (1958) identified in the tropical oceanic subcloud layer, Geisler and Kraus (1969) found it useful to assume a mixed layer for wind as well. This idea was reinforced by Deardorff’s (1972) early large eddy simulations (LESs), and such well-mixed layers are “observed” in current LESs of convective boundary layers (e.g., Moeng and Sullivan 1994). Thus boundary layer meteorologists changed from characterizing the daytime boundary layer wind in terms of the Ekman spiral to assuming that the wind is nearly constant with height except near the top and bottom (Fig. 3.14), as a result of the convectively driven mixing processes that similarly affect the temperature and mixing ratio. How accurate is this mixed-layer concept? Small vertical shears are reported by

---

7 One of my favorite “informal” field programs was described to me by a professor in England. Convinced by the Ekman-instability work that Ekman spirals don’t exist in the convective atmosphere, he launched and tracked a balloon for his dynamics class to show that the Leipzig wind profile was a rare occurrence. Unfortunately, for whatever reason, the data showed a perfect Ekman spiral!
Riehl et al. (1951) and Nicholls and LeMone (1980) in fair weather over the tropical oceans, where the diurnal variation is small. However, this was not the case for Pennell and LeMone (1974). Over land, the situation is more problematic, as illustrated in Fig. 3.15, which shows the wind shear vector magnitude as a function of height and time of day. Only between about 12:00 and 16:00 local time is the vertical shear small in the interior of the boundary layer. This also happens to be the time of day when many overland boundary layer experiments have been conducted to ensure that things are steady state enough for aircraft to collect a reasonable statistical sample for flux estimates. Similarly, LES modelers like to perform averages over a long enough period to ensure a good statistically significant ensemble average, leading to a preference for simulating near steady-state conditions.

A fortuitous combination of conditions led to documentation of significant boundary layer shear even during the early afternoon during STORM FEST and a reasonable hypothesis for its origin. (This is thanks to that string of “good” weather that led to a larger number of boundary layer studies than expected.) Figure 3.16 compares the wind profiles during the early afternoon for two days, 27 February and 10 March. While both $U$ and $V$ wind profiles on 10 March and the $U$ profile for 27 February have the expected near-zero vertical shear, the $V$ profile on 27 February has considerable vertical shear. Comparison of the two days revealed that the synoptic and turbulence characteristics were virtually identical except for one very important feature: the temperature inversion was weaker on 27 February, enabling continued growth of the boundary layer during the afternoon, and with it, entrainment of air with higher northerly momentum. Thus, even during the hours when the boundary layer wind over land “should” be constant with height, it doesn’t always happen.

4. Conclusions

I have followed the stages of field programs from their inceptions to the publication of results. A field program can be divided into three stages: the planning stage, the field stage, and most importantly, the data-analysis and synthesis stage. Although science and instruments are important, human interactions can make or break a field program. Humans determine the objectives, what is measured and when, and how much attention is given to the data afterward. To this, we have to work as a team. I tried to reproduce some of the feelings and difficulties we experience through the three phases. Even though field programs can be frustrating, we keep going back again. The important thing
is to keep one’s sense of humor, remember that tired people (including experimentalists) are fallible, and to maintain the experiment team long enough after any experiment to reap the results and answer the Big Questions.

Acknowledgments. Many people contributed to the ideas in this manuscript. Above all, I acknowledge my colleagues in the many field programs in which I have participated, and thank them for mostly maintaining their sense of humor. I also acknowledge that I share the flaws attributed to anonymous colleagues. Many of the ideas for the cartoons were suggested to me by colleagues, who also suggested adding a few details. Among these are Ed Zipser, Julie Lundquist, Bob Grossman, Greg McFarquhar, Gary Barnes, Garpee Barleszi, David Jorgensen, and Brad Smull. Also, two anonymous reviewers provided or reminded me of additional insights, sometimes made while still under the influence of Larium in Honiara. I would also like to acknowledge the staff of UCAR’s Joint Office of Scientific Support, who have helped me and a number of colleagues through tight spots on all three phases of many field programs, especially Jim Moore, Steve Williams, Dick Dirks, Nielmal Gamage, and Greg Stossmeister. I also gratefully acknowledge the support and sacrifice of the staff of NCAR’s Atmospheric Technology Division, and NOAA’s Office of Aircraft Operations. Both groups could write volumes about field programs. Kyoko Ikeda proofread the manuscript.

Many of the cartoons are from two unpublished manuscripts prepared by the author after GATE and TOGA COARE. The GATE manuscript is titled “Senegal’s chief import is nuts which (sic) are used to feed the mosquitoes.” The COARE manuscript entitled “What REALLY happened in Honiara: An off-white paper on the TOGA COARE folks on the turboprops . . . and the mosquitoes loved them.” Both are available from the NCAR archives (P.O. Box 3000, Boulder, CO 80307).

REFERENCES


