Issue Notes: November 15 Monitoring Workshop

Sampling coordination (Stephen Clark)

Core problem statement: There is a need for effort to reduce redundant sampling at the same location and time and/or to better coordinate the timing of comparable efforts to improve knowledge of watershed condition.

The desired outcome is to reduce and/or optimize duplicative efforts where possible. All participants strongly agreed that this was a desirable outcome. However, they realized as the discussion progressed that there might not be as much strictly overlapping (i.e., same time, same place) effort as presumed and that equal attention should be given to efforts to improve the integration of related efforts. For example, different programs might be sampling at the same time (but at slightly different locations) and additional information about watershed processes could be gained if the programs spread out their sampling in time to obtain better information about temporal variability.

The discussion recognized that program-specific objectives may prevent certain kinds of standardization or coordination. For example, the holding time requirements for pathogen monitoring for ambient condition and for drinking water quality are different, as are the requirements for nutrient measurements. In cases such as these, different programs will need to continue collecting their own samples. However, there are instances where one program could gather additional samples on behalf of another program, even where the programs are using different analysis methods. Given that staging field teams is expensive, this sort of logistical partnership could improve overall efficiency, assuming adequate arrangements could be made for transporting samples to the programs' respective laboratories.

Participants highlighted two types of institutional barriers that can hamper coordination. In one case, a large agency collects real-time water quality data that is quickly made available in an online database. However, these data often modified at a later time due to routine quality control checks and occasional instrument recalibrations. If such adjustments are not widely publicized among the datasets users, the validity and credibility of analyses and decisions made with these data can be undermined. Users' knowledge that such adjustments to the data can occur without their knowledge makes them somewhat reluctant to use the data, even though it is readily available. In another case, participants described a situation in which two agencies maintain gaging stations on opposite banks of a tributary. Though the stations are close together, they produce flow estimates that are somewhat different. For a variety of reasons (e.g., concern for historical data record, need to maintain funding) neither agency has been willing to remove its station and shift its resources elsewhere. While the participants recognized issues such as these as barriers, they also considered them as opportunities for achieving improved coordination and efficiency.

The groups discussed several types of potential structures or processes for improving coordination. For example, they considered regulatory mandates, voluntary collaboration, and combinations of these. Although no one of these was considered the preferred option, they did agree that collaboration cannot simply be mandated and that successful efforts were likely to involve a combination of both top-down, or government directed, and bottom-up, or grassroots, activities. Government can provide incentives (e.g., by encouraging collaboration in permit language), performance targets, and funding, while those most knowledgeable about monitoring and analysis efforts can help ensure that efforts focus on feasible opportunities and account for real-world circumstances. In particular, non-governmental agencies often have the flexibility to pursue collaboration in ways that are not directly available to government agencies.

There was no solid agreement about what sort of organization would be best suited to manage a coordination effort. Suggestions ranged from a czar to a nonprofit to a decision to leave coordination completely up to individual monitoring entities. In the last case, the inventory being developed by SFEI would be instrumental in helping each monitoring group identify its likely partners. This would involve information gathering about monitoring plans, using the inventory, followed by a more detailed assessment of overlaps and other opportunities to coordinate or collaborate. This process could be enhanced by linking information in the inventory to program QAPPs, reports, and datasets. For the inventory to be a useful tool in such coordination efforts, there must be some provision for its ongoing maintenance and updating.

During the discussion it became clear that not all participants clearly understood how these different approaches might work in practice. It might therefore be useful to prepare a description of several successful efforts in California to illustrate the ways various approaches have been applied in different situations.

Methods and QA standardization (Emilie Reyes)

Core problem statement: there are discrepancies among the field and laboratory methods used by different programs, as well as in their respective QA requirements, and a need for greater transparency in and access to information about these methods and requirements.

The desired outcome is to reduce discrepancies where feasible, clarify the underlying program-specific data quality objectives that give rise to these differences, and achieve broader access to data and metadata. All participants agreed that these outcomes were desirable in general. However, many pointed out that not all method discrepancies are "bad". Some quite legitimately reflect different data quality objectives, such as those related to effluent and ambient monitoring efforts. Thus, it would be useful to more clearly identify data quality objectives and to develop ways to make data equivalent for assessment purposes even if they have been gathered to meet different data quality objectives.

The participants identified three approaches for dealing with the existing situation. The first was to provide more information to members of the monitoring community, including QAPPs, standard operating procedures (SOPs), and more detailed metadata about program goals, designs, methods, and products. Many participants believed that this approach would enable users of the data to make their own judgments about the suitability of data from various sources. This information could be posted online on program websites or as part of the inventory demonstrated by SFEI. However, making QAPPs broadly available might be an issue for some analytical laboratories.

The second suggested approach involves a number of specific methods to better define and then improve the comparability of data from different sources. This could involve performance evaluations and audits, laboratory intercalibration studies, and side-by-side comparisons of different methods. This last approach was considered to be particularly helpful in improving the utility of legacy data collected with methods that are no longer being used. Such efforts could result in certain methods being replaced, combined, or standardized. For example, participants identified a need to define a standard policy for dealing with non-detects and to develop tools for lowering the costs and increasing the reliability of data entry. However, in general, most participants believed that such opportunities would be relatively limited and that the most likely and useful result would be a much better technical understanding of the ways in which data differ and the implications of these differences for attempts to combine and compare data from different sources.

The third approach focused on defining multiple tiers, or categories, of field, laboratory, and QA requirements suitable for different purposes and/or programs. In principle, such tiers would define

progressively detailed or rigorous criteria for detection limits, QA procedures, metadata, and other issues. Despite virtually unanimous agreement that such a tiered approach would be very useful, participants were not able to describe specifically how these tiers would be defined in practice.

There are a number of barriers that could impede the implementation of the approaches suggested above. The first is simply the great number and variety of monitoring programs in the region. Combined with the large amount of legacy data that are still of interest, this makes it challenging to collect, organize, and maintain current information on the details needed to assist data users in making appropriate judgments about the utility of datasets for different purposes. The second barrier is determining who would be responsible for developing and maintaining the mechanisms (e.g., online inventory, data center) intended to improve the transparency of information about monitoring programs and access to the data they produce. While there was broad support among the participants for the concept of a data center, they did not provide specific suggestions for how a data center should be structured and managed. The third barrier is finding the resources for supporting programs in developing QAPPs, training in procedures to improve comparability, maintaining the SFEI inventory, and expanding and maintaining the existing prototype data center. Supplemental Environmental Projects (SEPs) were identified as one potential source of funding.

Funding and institutional arrangements

Core problem statement: there is a lack of stable funding to support long-term data collection and to support administrative needs (e.g., data management, annual workplans). There is no organizational structure to organize the multiple organizations active in monitoring and there is no statewide framework to ensure consistency in methods across all these organizations. Finally, there is no regional entity in position to implement any coordinating mechanisms that might be proposed.

The desired outcome is to identify potential institutional and funding structures that could fill these gaps.

In terms of funding, all participants believed that no substantial new sources of funding would be available to support the coordination efforts being discussed (e.g., methods standardization, data center). Instead, funding would have to come from using existing efforts to leverage stable funding. These existing efforts fall into four broad categories, including the regulatory community (e.g., TMDLs, SWAMP, Irrigated Lands Regulatory Program (ILRP)), the regulated community (e.g., municipalities, POTWs, agricultural coalitions), resource agencies (e.g., US Fish and Wildlife, California Department of Fish and Game), and private institutions (e.g., universities, non-governmental organizations). While the participants did not agree on any specific funding priorities, potential sources discussed included finding resources by streamlining and better coordinating existing programs, using new and/or improved technologies to improve the efficiency of sampling and data entry, and grant programs.

Participants also identified a number of institutional constraints to sharing data and coordinating monitoring efforts. Regulatory and resource management agencies have concerns that the integrity of their monitoring programs, and the long-term consistency of historical data records, could be damaged. In addition, differences in program objectives can require different monitoring designs and methods, which lessens the payoff from sharing and coordination efforts. All parties have concerns about the expense and effort involved in establishing the protocols needed to support effective data sharing. Even where such efforts can result in longer-term savings, parties may not be able to sustain the near-term costs involved. Regulated entities have a particular concern that their data, if shared with others, could then be incorrectly or unfairly interpreted with negative consequences for their program's credibility. Finally in some cases, access to data is restricted by prior agreements.

In addition to these constraints, participants identified a number of others. Contracting procedures may be so complex and/or time consuming that the opportunity for coordination has passed by the time one entity has established a mechanism to compensate another, for example, for collecting and then sharing data they both need. In many cases, cost sharing arrangements are impeded because it can be difficult to determine how much shared data are worth, especially when different parties have different cost bases or calculate their costs in different ways. New arrangements (e.g., data transfer protocols, data center) may require that staff be trained in new procedures and it can be difficult to find the time and/or money to support such training efforts. Finally, participants noted that time can be as important a constraint as funding, since both technical and management staff for all parties are typically very busy.

In terms of the institutional arrangements needed to foster and maintain monitoring coordination and data sharing, participants agreed on the need for a global or overarching statewide framework combined with regional or local structures that would ensure that specific, real-world needs were being met. Although participants did not make a firm decision about what institutional structures would be best, they did suggest a number of potentially useful alternatives. At the statewide level, the Water Quality Monitoring Council being established under SB 1070 was an obvious candidate to take on a central role in developing the infrastructure to support data sharing and coordination. It was suggested that the Monitoring Council could use aspects of monitoring within the San Joaquin Valley as a pilot project. Participants also thought that the California Water Plan Steering Committee, operating within DWR to develop statewide plans dealing with water supply issues, could be an appropriate partner for the Monitoring Council. Together, these two entities would span nearly the full range of water quality and water supply monitoring efforts, something the participants believed would be useful because of the often close functional connection between water quality and water supply.

At the regional level , participants thought that the integrated regional water quality management process (IRWMP) could provide a framework for addressing coordination at the local and regional level, especially if they implement joint powers authorities or other management structures. In addition, the Prop. 84 funds going to support IRWMPs might provide a funding source for coordination. However, IRWMPs have been implemented across the state in a variety of ways and there does not appear to be a standard template for their development and implementation. In addition to IRWMPs, participants mentioned a number of third-party arrangements that could help to organize and manage efforts at the local/regional level. Examples of these include SFEI in northern California and SCCWRP in southern California. A regional entity could fulfill a number of functions, including managing planning processes and maintaining communication among partners.

Data analysis and interpretation (Sam Harader)

Core problem statement: data are currently not assessed, interpreted, or displayed in a manner accessible and understandable to multiple users. In addition, there are poor linkages between data collectors and users.

The desired outcome is to improve the access to and usability of monitoring data.

The discussion among the participants focused primarily on impediments to data access and integration. Data formats are often not compatible with analytical software or with broader analyses. For example, inconsistent parameter names (e.g., site, constituent, method) and the absence of generalized cross reference lists make it difficult to combine data from multiple studies. In addition, different programs may use different geospatial reference systems and/or different basin or watershed boundaries and study designs that require stations to be moved periodically make it difficult to conduct longer-term analyses for trends. The lack of a centralized source for reports, QAPPs, and other documentation also hinders the

broader use of monitoring data. Most of these issues make it more difficult for users who are not directly involved in a particular program to obtain timely access to monitoring data and results.

Participants identified a number of solutions to this set of problems. Many problems would be solved by a centralized database for reports, QAPPs, program descriptions, and data documentation, along with data integration tools such as inclusive cross reference lists. Other problems would be solved by encouraging or requiring standardization of certain parameters used for integrating disparate data. Participants also agreed on the value of creating some institutional infrastructure to promote and support data integration and analysis, such as working meetings, email newsletters, and broader distribution of reports and other analysis products that could stimulate further effort. Finally, some users suggested that it would be useful to include some analysis tools in a centralized regional database, though no specific tools were identified.

While most of the discussion focused on logistical barriers described above, some participants also noted that most monitoring programs are not explicitly hypothesis or question driven, which makes it difficult to identify objectives at the beginning of a program and thus to apply the most effective data analysis tools. A broader issue also mentioned is the need for more analyses that integrate across larger space and time scales and/or disciplines in order to expand understanding of complex processes and improve predictive capabilities. However, these issues did not gain much traction with most participants, perhaps because resolving the logistical issues described above is a necessary prerequisite for undertaking more complex, multidisciplinary analyses.

Data center (Mike Johnson)

Core problem statement: the dispersion of monitoring data across multiple programs which use a wide variety of data storage approaches makes it difficult to identify, locate, access, and use monitoring data for a broader set of uses.

The desired outcome is to design and implement a data center that would provide a single point of access for all users of monitoring data.

Participants identified substantial barriers to this solution. These fell into three categories: technical, institutional, and financial. Technical barriers are related to the challenges of designing and implementing the database systems needed to centrally store and/or provide remote links to an extremely wide variety of datasets and data types. Technical barriers also relate to the often substantial effort involved in identifying datasets, documenting their characteristics and levels of QA, and standardizing key aspects of data formats to streamline transfer across data systems. Institutional barriers stem from concerns about ownership of data, including intellectual property rights related to publications based on monitoring data. These barriers also stem from concerns about loss of control over how data are used and interpreted and about how agreements about data use will be developed and enforced. The level of effort involved in creating and then maintaining a data center results in financial barriers since funding can be difficult to obtain.

Participants' familiarity with existing centralized databases and/or data centers made it relatively straightforward to identify the major issues involved in planning and implementation. For example, the types of data to be included, and their location, must be carefully defined and agreements about access to the data completed. The data center's core structure must also be defined, as either a centralized repository or a clearinghouse that links users remotely to a distributed network of smaller data centers and databases. Rules for obtaining or linking to datasets must be established, along with explicit procedures for data and metadata formats, user access, and data transfer. User guides interfaces must be designed and tested and any needed training implemented.

Despite their broad agreement on the key issues that must be resolved, participants differed widely in their opinions about what approaches were best suited to resolving these issues. For example, some participants thought that the data center should include extensive metadata, including QAPPs, while others felt strongly that QAPPs were not appropriate for inclusion. Similarly, some participants believed the data center should provide access to all available monitoring data and others that only data that exceeded certain quality thresholds should be included. At a more detailed level, some participants felt the data center should track usage and include user logs while others believed the center's function should simply be to provide ready access. Finally, the estimated annual cost to maintain the data center ranged from $1 – 2 million to $10 – 12 million.

This range of opinion about key design and functional issues reflected participants' very different ideas about what a data center could or should be. Thus, the next step in planning for any data center should focus on achieving a better definition and a broader consensus about the specific needs it is intended to meet and what functions are essential for meeting those needs.