



SIGAR

Office of the Special Inspector General
for Afghanistan Reconstruction

January 29, 2013

The Honorable Hillary Rodham Clinton
Secretary of State

Dr. Rajiv Shah
Administrator
U.S. Agency for International Development

The Honorable James B. Cunningham
U.S. Ambassador to Afghanistan

Dr. S. Ken Yamashita
Mission Director, Afghanistan
U.S. Agency for International Development

I am writing to share with you the results of an analysis we recently conducted—in conjunction with both the Criminal Investigation Task Force (CITF)¹ and the National Geospatial-Intelligence Agency (NGA)²—on data contained in the Afghanistan Infrastructure and Security Cartography System (AISCS). AISCS is a web-based geospatial database intended to provide the U.S. Agency for International Development (USAID) a comprehensive and accurate picture of infrastructure development in Afghanistan. Because of the importance and potential value of this database as an oversight and planning tool, we tested the reliability of AISCS data. Our analysis showed that, while the database could be a useful tool for mapping infrastructure in Afghanistan, there are some problems with the data's accuracy and completeness. Most notably, the geographic coordinates for a number of the sites appear to be incorrect.

Background

AISCS was developed by International Relief and Development, Inc. (IRD), a USAID contractor, through its Afghanistan Infrastructure Data Center.³ According to the USAID Office of Inspector General, AISCS was designed to include geospatial information on development activities, including construction of roads, schools, clinics, hospitals, and public buildings such as courthouses and district centers. IRD was to obtain infrastructure project site information from the USAID mission in Afghanistan, as well as the Department of

¹ The Department of Defense created CITF in 2002 to conduct worldwide criminal investigations of alleged or suspected war crimes and acts of terrorism against U.S. persons, property or interests in conjunction with the Army Criminal Investigation Command.

² NGA is the lead federal agency for geospatial intelligence. Its mission is to provide timely, relevant, and accurate geospatial intelligence in support of national security.

³ This data center, including the staff, hardware, and software used to create AISCS, cost an estimated \$4 million.

Defense, and verify the accuracy of that information.⁴ According to IRD, it has “stringent ongoing multi-tiered quality control protocols in place to insure accuracy and precision collected data [sic].”⁵

What We Did

In 2012, SIGAR requested that USAID provide a list of all USAID-constructed infrastructure projects in Afghanistan. USAID’s Office of Economic Growth & Infrastructure responded to this request by providing a spreadsheet of 5,648 construction project records.⁶ Records in the dataset contain information such as project name, location, status, and donor.

Our initial assessment of the dataset raised some concerns about its completeness and reliability. For example, we identified duplicate records—entries with an exact match by project name and location (latitude and longitude). We also found instances of missing data elements within records, including project type, implementing partner, and actual cost information.

We determined that the dataset had been pulled from AISCS and decided, therefore, to assess that database more carefully. We observed that, of the 33,000 records in AISCS, only 16 percent were shown as having confirmed locations. We selected 10 projects from AISCS and asked CITF to corroborate the projects’ stated locations using a variety of unclassified and classified sources. These 10 projects were spread geographically throughout Afghanistan and represented various sectors, including health, education, government, transportation, and hydroelectric power.

CITF found that location coordinates for 4 of the 10 projects were highly accurate, but were significantly inaccurate for 3 of the 10. For the remaining 3 projects, CITF determined that insufficient corroborating information was available to assess the spatial accuracy of the coordinates. In other words, CITF was unable to locate the project.

Because this initial test raised our level of concern about the reliability of AISCS, we chose to perform a more methodologically rigorous analysis of the data. We selected one category of buildings from the data set provided—those listing “education” under the “building type” category. We chose this category, in part, because location information (both latitude and longitude) was available for all buildings in this group. We then isolated those that were shown as “completed” in the “status” field and also removed duplicate entries we had identified.⁷ This filtering process resulted in a total of 549 completed USAID-funded schools⁸ in Afghanistan.

We then stratified the 549 project records into urban and rural locations.⁹ We found that 83 percent of the schools were located in rural areas, while 17 percent were located in urban areas. Using these two strata, we

⁴ USAID Office of Inspector General, Audit of USAID/Afghanistan’s Human Resources and Logistical Support Program, Audit Report No. 5-306-10-007-P, March 31, 2010.

⁵AISCS Newsletter Issue No. 1

⁶ USAID omitted from this list large energy infrastructure rehabilitation projects, including existing hydroelectric plants, substations, transmission lines and some distribution networks.

⁷ Of the 613 buildings in this category, 53 were not listed as completed and 11 were duplicates.

⁸ For the purposes of this letter, we are using the term “schools” to describe a variety of education-related facilities, also including dormitories, provincial training centers, and kindergartens.

⁹ Due to the disparity in access to and resources available in urban and rural locations within Afghanistan, we determined that the rural/urban classification of a school could have a significant impact on a building’s record accuracy with respect to location. We sought to increase the precision of our findings by matching the proportion of rural/urban locations in our sample to the larger population. To determine whether a school would be considered urban or rural, we used population data from the Afghanistan Central Statistics Organization for the school’s civil division. If the school was located in an area

sampled 189 school records in rural locations and 38 in urban locations for a total of 227 school records in our final stratified statistical sample.

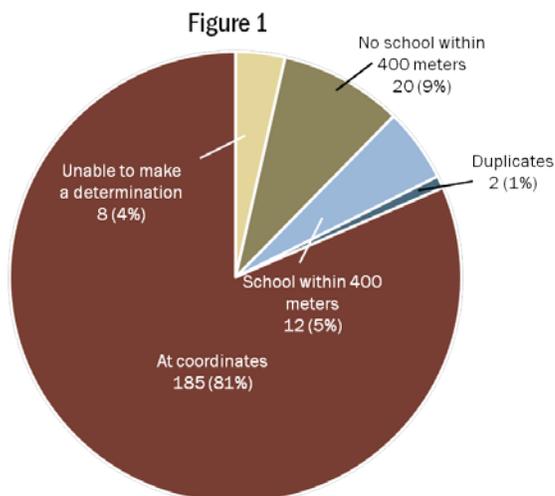
We submitted the 227 sample school records to NGA. Using the latitude and longitude provided within each record, NGA analyzed available imagery and returned an analysis of the sample records that listed location confidence as “school at coordinates,” “school within a 400 meter radius,” “no school within a 400 meter radius,” “unable to determine,” or “duplicate.”¹⁰

What We Found

NGA determined that the stated location for most of the schools in our sample was accurate. Specifically, of the 227 schools, NGA found that 185 (81%) of them were at the given coordinates.

However, NGA could not confirm the coordinates for the remaining 42 records (19%). Specifically, for 12 of the records (5%), NGA determined that there was a school within 400 meters of the given coordinates; for 20 (9%), there was no school within 400 meters; and for 8 (4%), NGA was unable to make a determination. NGA considered two of the schools (1%) to be duplicates.

SIGAR considers the coordinates for these 42 records (19%) to be questionable. Based on our sample, we project with 95% confidence that between 74 and 129 of the total 549 completed, USAID-funded schools in Afghanistan have location coordinates that cannot be verified using geospatial means or are incorrect.¹¹



Conclusion

Complete and accurate data on the location of U.S.-funded infrastructure in Afghanistan is critical for successfully managing reconstruction projects and is relied upon by SIGAR and others as an important oversight tool. While most of the AISCs data on school location appears to be accurate, we are also concerned by the problems we found, particularly given IRD’s statement that it applies “stringent ongoing multi-tiered quality control” to the data. Incomplete and/or inaccurate data limits our ability to verify the effectiveness of expended reconstruction funds, validate that reconstruction projects are being used for their intended purpose, and analyze the efficiency and effectiveness of U.S.-led reconstruction efforts in Afghanistan. It will also hinder the ability of the Afghan government to sustain these projects after it assumes responsibility for them. We therefore encourage you to carefully assess how IRD verifies and maintains this database and to hold IRD accountable if shortcomings in this process are identified.

with a majority urban population, it was categorized as “urban;” if the majority of the local population was rural, we categorized it as rural.

¹⁰NGA considered a record to be a duplicate if it contained coordinates that were at or near another provided set of coordinates in the sample, even if the project names differed. By contrast, SIGAR considered duplicates to be those for which there was an exact match by both building name *and* location.

¹¹ At a 95-percent confidence level, our sample proportion fell within 5.1% of the actual population parameter. We can say with 95-percent confidence that the total percentage of questionable project locations in the population falls within 5.1% of the total observed in the sample (18.5%), and is therefore between 13.45% and 23.55% of the 549 records in the population.

We conducted our work in Arlington, Virginia, from July 2012 to January 2013 in accordance with SIGAR's quality control standards. Those standards require that we plan and perform the review to obtain sufficient, appropriate evidence to meet our stated objectives and to discuss any limitations in the work. We believe that the information and data obtained, and the analysis performed, provide a reasonable basis for our findings and conclusions. This work was conducted under the authority of Public Law No. 110-181, as amended; the Inspector General Act of 1978; and the Inspector General Reform Act of 2008. Major contributors to this report were Derek Benekebove, Mia Bonarski, and Ben Tangpricha. Please contact my Assistant Inspector General for Audit, Mr. Ryan Coles, at [REDACTED] if you have any questions or concerns regarding this work.

Sincerely,

A handwritten signature in black ink, appearing to read 'J. Sopko', with a long horizontal flourish extending to the right.

John F. Sopko
Special Inspector General
for Afghanistan Reconstruction



March 9, 2013

The Honorable John F. Sopko
Special Inspector General
for Afghanistan Reconstruction (SIGAR)

Dear Mr. Sopko:

Thank you for your letter of January 29, 2013, sharing the Afghanistan Infrastructure and Security Cartography System (AISCS) analysis that SIGAR conducted with the Criminal Investigation Task Force (CITF) and the National Geospatial-Intelligence Agency (NGA). Based on this analysis, SIGAR raised concerns on the accuracy and completeness of AISCS data, specifically citing errors with the geographic coordinates for a number of the sites.

Comparison of SIGAR/NGA and USAID/IRD Reviews

We understand that SIGAR identified a sample of 227 school records from AISCS, which NGA analyzed using available imagery. NGA found 185 of those schools (81%) at the given coordinates, but could not confirm the coordinates for the remaining 42 records (19%). As a result, SIGAR considered the coordinates to be questionable.

To address SIGAR's concerns, USAID directed the contractor that developed AISCS, International Relief and Development, Inc. (IRD), to conduct an immediate review of the database, with particular focus on the 42 records (19%) that SIGAR/NGA identified as having either incorrect or inconclusive coordinates in AISCS. IRD reviewed the coordinates using a combination of aerial imagery and site visit reports. The review findings are summarized below and in Enclosure 1, and supported by the enclosed slides showing imagery for each location.

- For the 12 schools (5%) NGA found within a 400-meter radius of the given coordinates, IRD's review located 11 structures within the 400-meter radius. One (1) structure was located beyond 400 meters from the AISCS

coordinates. The AISCs coordinates for this school have since been updated. Imagery for the 12 schools is provided in Enclosure 2.

- For the 20 schools (9%) NGA could not locate within a 400-meter radius of the given coordinates, IRD's review located 11 structures within 400 meters of the AISCs coordinates. Using coordinates found in site-visit reports, IRD located eight (8) structures outside the 400-meter radius from the AISCs coordinates. Coordinates for one school were incorrectly recorded in both AISCs and the field trip report, and will require a field visit for verification. Please refer to Enclosure 3.
- For the eight schools (4%) that NGA was unable to make a determination, IRD's review located all eight structures at the given AISCs coordinates. Enclosure 4 shows the imagery for each school.
- For the two records (1%) that NGA considered duplicates, IRD's review located one (1) school at the AISCs coordinates, with the other school located 225 meters from the AISCs coordinates. Please refer to Enclosure 5. With regard to SIGAR's concern with duplicate records, it should be noted that each record in AISCs represents a project or assistance, rather than a structure. Some structures received assistance more than once for different purposes (i.e., non-structural renovation, structural renovation, etc.). A structure would be listed in AISCs as many times as it received assistance.

In summary, of the 42 school records that SIGAR/NGA found questionable, IRD located all schools, with the exception of one structure. Thirty two (32) schools were located within 400 meters of the AISCs coordinates, in addition to the 185 schools that NGA already confirmed as having accurate coordinates. This would bring the total number of schools located within 400 meters of the AISCs coordinates to 217, or 96% of the total sample. Nonetheless, nine (9) schools (4%) were found to have the wrong coordinates recorded in the system. The lone school not located was completed in 2004 and was last visited on November 4, 2007 as part of a post-occupancy evaluation. While the structure and condition of this school were documented in the site-visit report, the location coordinates were erroneously recorded, and will thus require a re-visit for verification.

In the same manner, USAID and IRD would also like to address CITF's findings that "location coordinates for 4 of the 10 projects were highly accurate, but were significantly inaccurate for 3 of the 10." Without more information about the sample data, the assessment methodology employed, and

any other parameters that would help us identify the project records CITF assessed, we cannot respond to the SIGAR/CITF findings.

Challenges of Infrastructure Data Collection and Location Verification

To understand the challenges of verifying location data for the schools, it is important to note that most of these projects were completed prior to 2008, when IRD, under the Human Resource and Logistical Support (HRLS) Program, did not have the technology for data verification. Location data were collected using hand-held Global Positioning System (GPS) devices manufactured for recreational purposes, such as those from Garmin, and data manually recorded on paper. Between 2004 and 2007, the HRLS Program relied largely on supervisors' review of location data submitted by the engineering monitors. With the availability of improved technology, i.e., NGA-supplied imagery, IRD has added the use of aerial imagery verification to its location data-confirmation process.

Recognizing the need to develop a comprehensive list of infrastructure projects in Afghanistan for effective planning, USAID's former Office of Infrastructure, Engineering and Energy (OIEE) tasked IRD in 2008 with the expansion of the database to include project data from other USAID offices, the Afghan Government, and donors, including the Military. USAID's contract with IRD did not require verification of the non-OIEE data, as the cost of doing so, especially given the security and logistical conditions in Afghanistan, would have significantly outweighed the benefits. As such, IRD collected non-OIEE project data and used them in the database, but classified them as "unconfirmed." Location data for non-OIEE projects are verified as availability of human and logistical resources and security conditions allow. To date, IRD has confirmed location data for 1,555 buildings, or 52%, of the 3,003 health, education and government building records in the database.

Conclusion

USAID welcomes feedback on AISCs, as it will help us further improve the database. AISCs evolved from a list of infrastructure projects that IRD was monitoring on behalf of OIEE under the HRLS Program. Its expansion into a database that includes infrastructure projects of other Mission offices, the Afghan government, donors and the military was borne out of the need to know what other players in Afghanistan are doing in order to improve infrastructure planning and development. We recognize the expansion of AISCs has its

trade-offs on data accuracy, but we also believe that this can be mitigated by joint data reviews, consultations and data sharing. The assistance of SIGAR, NGA and CITF towards this end is greatly appreciated.

Sincerely,



Brooke Isham
Acting Mission Director

Enclosures:

1. Summary of IRD Review of the 42 School Records
2. Imagery for NGA Classification "Schools within 400 meter radius"
3. Imagery for NGA Classification "No school within a 400 meter radius"
4. Imagery for NGA Classification "Unable to determine"
5. Imagery for NGA Classification "Duplicate"