



INTERNATIONAL
INSTITUTE FOR
ANALYTICS™

RESEARCH BRIEF

RESEARCH & ADVISORY NETWORK

Leveraging Analytics in Environmental Compliance

SAS EXPERTS: DAN CHILDERS, JULIE ESPY, AND JEREMY RAY

IN DISCUSSION WITH ROBERT MORISON, SENIOR ADVISOR, IIA

NOVEMBER 2024



Discussion Overview

Across the country, state environmental compliance agencies are dealing with growing amounts of data, reporting requirements, and responsibilities, often with limited staff and resources. Their landscape is in constant motion—regulations and policies change, public emphasis on the environment grows, and states try to adjust the balance between enabling economic activity and safeguarding the environment. More advanced data and analytics offer opportunities to streamline permitting, focus compliance efforts, mitigate workforce shortages, and advance the core mission of ensuring a safe and sustainable environment. To explore these challenges and opportunities, IIA interviewed SAS experts Dan Childers, Solutions Architect, Julie Espy, Senior Industry Consultant, and Jeremy Ray, Senior Industry Consultant.

What are some of the biggest challenges facing environmental compliance agencies today?

Julie Espy: Agencies are deluged with data, often including permit applications and permittee reports that are still on paper. The data comes from many different sources, the amount of data keeps growing, and agencies' data management systems often need updating. Most agencies work with a variety of data structures and systems that have been pieced together over time and don't communicate well with one another.

So people spend a lot of time and effort dealing with document overload and trying to pull information together, and they still can't get a complete and timely picture of what's happening. Lack of strong data management makes it difficult to do the job—permitting takes longer and non-compliance is harder to uncover.

Dan Childers: Oftentimes environmental agencies are resource constrained in terms of both budget and staff. They don't have enough people to handle growing workloads and reporting responsibilities, and this new work requires new information and analyses. Work processes are too manual and labor-intensive, which makes them too time-consuming and error-prone. So it's no surprise that many agencies face constant and lengthy backlogs. Many agencies are also seeing more staff turnover and retirements than they are accustomed to, so institutional knowledge, capability, and experience are being lost. All of these factors slow processes down, and manual audits and inspections often aren't able to detect non-compliance or fraud in a timely fashion. That's frustrating for agency staff and management. They simply need better tools to do the job.

Jeremy Ray: Here are two more complications. First is that environmental policy and priorities keep changing over time. It's a moving target. As rules and metrics are captured in databases and implemented in applications, those data structures and systems need to be flexible enough to cope with change, so you don't have to keep going back to the drawing board. Second is the high degree of reliance on self-reporting by permittees whose own data and systems may not be in the best shape. The information they provide is often incomplete or inaccurate. To screen for completeness and accuracy, spot simple errors, and surface instances of non-compliance or deception, agencies can use more advanced and automated analytical capabilities.



How can analytics put data to better use and improve agency performance?

Dan: By detecting patterns and putting that insight to work. With better analytics we can notice things sooner and, more specifically, what's unusual, what's changed, who the outliers are. We can spot more subtle changes and indicators of non-compliance than people can on their own. Seeing the patterns gives us the power to better automate routine actions and common decisions, thus the power to free up and focus staff attention on what's unusual, important, and most worth their effort. So we're improving both the simplest and the most challenging activities and processes. We're triaging and prioritizing work, expanding the capacity of the organization, and enhancing the work of the individual. And greater productivity within the agency translates into less regulatory burden and cost for permittees.

What are some specific use cases in the permitting process?

Dan: In many agencies, permitting is ripe for more automation. Some permit application types are high volume but very simple and not very variable, like small-scale construction permits, for example. When we recognize the simple cases, they can be processed for completeness by a decisioning engine with little human interaction—perhaps just a check of the result. Permit history can be regularly analyzed to provide a baseline. When a new permit application comes in, it's assessed and assigned a complexity or difficulty score and routed accordingly. Renewals may follow simplified steps. All of that frees staff time to handle the more complex permits. They can do less of the small stuff.

Julie: Looking at the patterns of permits and approvals reveals a lot. How long does the process take? What kinds of permits prove time-consuming or troublesome, and why? What are the common sticking points and how are they commonly resolved? The answers form the basis for process improvement and staff training or knowledge exchange. Over time, the process becomes more repeatable, case management becomes more focused, and performance improves.

Dan: If you're building business rules and analytical knowledge into triaging work and automating decisions, you're codifying that institutional knowledge and experience we mentioned. Better analytics makes people smarter in their jobs, but also preserves and shares those smarts.

Jeremy: We've been focusing on using analytics-based automation to help staff in their work and performance. We should keep in mind that this also greatly benefits the permit holder. Each permit gets the attention and resources it truly needs. Simple permits, complicated permits, and those in between can all be handled more quickly, consistently, and equitably. That's good for permittee service, relationships, and mutual trust.

How do analytics play a role in compliance detection and assurance?

Dan: Spotting non-compliance relies on two fundamental analytical techniques, peer grouping and time series analysis. In both cases, we're looking for outliers in permittees' reported performance relative to their permitted levels. When we find outliers, we flag them and assign them a risk score, say from zero to 100, indicating likelihood and potential magnitude of a problem. Significant differences from the norm, hence high-risk scores, are most worth investigating.



With simple time series, we can see variations and trends in an individual permittee's history. Are there sudden changes in reported emissions, or changes that persist? With peer grouping, we can also cluster and compare permittees of similar size and operating characteristics. Compared to the group's averages, who are the outliers in what they report? With the EPA's ECHO database, we can form groups to analyze within and across states.

Julie: Over time, the agency accumulates a library of sorts of common scenarios that suggest something may be going on in permittees' operations or reporting. Investigators can use risk scores to determine and prioritize their case management next steps, including when to go into the field. They can focus on the hot spots rather than trying to look at every piece of data and treating all permittees and their reports the same.

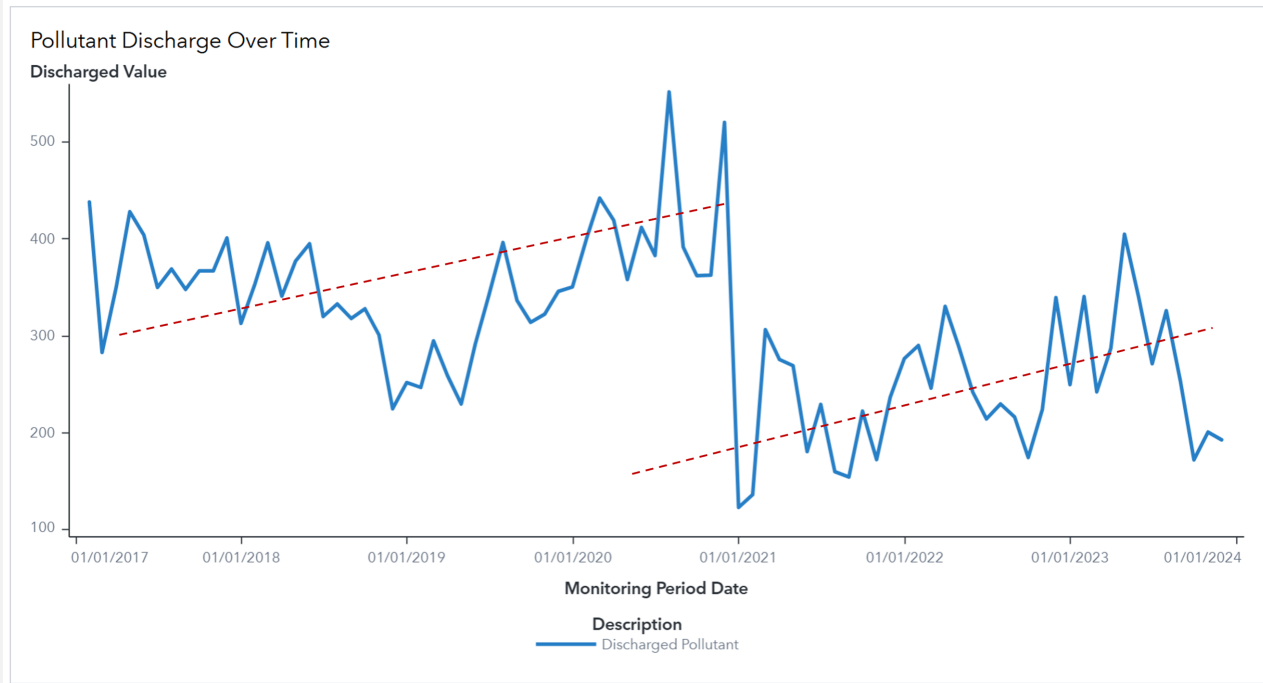
Jeremy: These basics of compliance investigation are still largely manual in many agencies. A better data and analytics platform enables them to be done more quickly, consistently, and effectively. Analytical tools can be essential for detecting something that's truly underhanded or fraudulent. Bad actors try to take advantage of agencies' reliance and trust in their self-reporting, and sophisticated bad actors might be able to disguise their activities in ways that are difficult to uncover. Agencies need to be proactive in detecting subtle or unusual changes.

Dan: An example of outright fraud that made the news was Amphenol Aerospace, fined \$945,000 by the New York Department of Environmental Conservation for consistently under-reporting pollutant levels for over a year. The company was removing nitric acid from their wastewater treatment process prior to compliance sampling, then returning it into the wastewater treatment stream after samples were collected. They were cited for a total of 83 violations of their permit.

Amphenol's activities and fraudulent reporting were revealed by an anonymous whistleblower. But in retrospect, the story was in the data and might have been surfaced in a matter of a few months, enabling much earlier direct investigation. Instead of waiting for or relying on a whistleblower, it's often possible to blow the whistle with analytics.

Julie: At the other end of the spectrum, sometimes significant deviations from the norm can be resolved easily because they represent sudden but temporary operational quality problems or even just reporting errors. In one case, a permittee inadvertently swapped reported levels of nitrate and nitrite, a simple mistake that jumped out on a time series graph.

Casting the Right-Sized Net



The chart depicts a wastewater facility's reported pollutant levels in a time series, including an instance of an unusual dip. We call these sudden changes "structural breaks," because relatively stable trends are separated by large deviations, which may indicate a new trend taking shape.

The structural break in this time series seems obvious to the human eye. But behind the scenes, automatically detecting these relatively large changes, as well as smaller but still-significant changes, involves a complex and difficult balancing act. An agency could apply simple (but naïve) business rules to operators as a group, flagging periods where changes cross a fixed threshold (e.g., +/- 20%). The big question then becomes: what should the threshold be?

Whatever is chosen, the rule would likely work to some degree. However, if the threshold is set too low (casting a wide net), the algorithm will flag a lot of false positives. And if it is set too high (casting a small net), it will fail to detect many true positives. Both cases are problematic, because there will always be a tradeoff between wasting agency resources through unproductive investigation efforts and failing to detect non-compliance.

Thankfully, there is a better way. Using advanced forecasting techniques, we can statistically determine whether a period-over-period change was large enough to be considered improbable *for any facility's time series*, considering the idiosyncrasies of its individual history. We can reconcile the tradeoff between casting a large or small net by dynamically customizing the size of the net.



Please say more about addressing the underlying data challenges.

Dan: The objective of a robust data platform is to support business operations and analytics as effectively and flexibly as possible. The platform has two major layers. One is the databases associated with business processes and systems that store the core information of the enterprise. Atop that is a data integration layer that assembles data from a variety of sources, both inside and outside the agency, then reconciles and prepares the data for analytical algorithms and tools to use. The first layer enables everyday operations. The second layer gets relevant data into one view that's analytically ready to go, and so enables insight about those operations.

Jeremy: The integration layer is essential. No matter how cleanly architected and modernized core systems and databases are, there are still going to be legacy systems data and data from external sources to incorporate. And you have to be able to look at data broadly, to pull in data from disparate sources, and create an integrated view in order to make new connections, discover new patterns, and derive new insights. We don't know what we don't know, so we need to analyze data at scale.

Dan: Many environmental agencies have efforts under way to modernize their technologies, some with funds from the Inflation Reduction Act. Most are focusing primarily, even exclusively, on their core systems right now, literally trying to get their base data together. But it's short-sighted to get data in order without looking ahead to how it will be used and actually starting to use it.

Julie: That's exactly what we're finding. Agencies really focusing on organizing their data, and not what they can do with the data once they have it in order.

Agencies may find core systems modernization daunting enough and be tempted to postpone data integration. However, we strongly recommend building out the data integration layer in close concert with core systems modernization. That way agencies can develop and exercise the interfaces, pilot new analytics, do more sophisticated data analysis sooner, and deliver value as they go.

Jeremy: Keep in mind that an overarching objective is to keep the technology platform agile, flexible enough to change. That way agencies can deal with new data and increasing demand for information, constantly improve their operations, and respond to regulatory and policy changes. A flexible architecture can also embrace complementary technologies. OCR can be used to scan documents still in paper form, natural language processing can help interpret and evaluate the documents, and mobile technology can enable investigators to search data, do basic analytics, and refine their inquiries and methods on-site.

Dan: As another example, generative AI is maturing fast. It may have practical use cases helping staff keep pace with regulatory change, examining new permit applications, or even serving as a “co-pilot” interpreting analytics and making suggestions on next steps to take.

Julie: A high performance information systems and data analytics platform improves agencies' performance in fundamental ways. It supports repeatable processes for permitting and compliance, as well as regular improvement of those processes. Through both automation and analytics, it enables agency staff to accomplish more with less time and effort. It enables the agency to communicate consistently—internally, with permittees, and with the public—and create more accountability, transparency, and trust all around.



What summary advice do you have for agency leaders seeking to capitalize on better analytics?

Julie: Enable and encourage your staff to adopt better automation and more analytically informed decision making. We read a lot these days about technology, particularly generative AI, taking people's jobs. That's not the case here, especially if the agency is short handed. As we mentioned, a more robust data and analytics platform helps automate routine tasks and free staff for more challenging and value-added work. Better information and analytics enrich their work and enable them to perform better and accomplish more. I know from my experience in state government that most people there like being public servants, they have a passion to contribute, and they want to get better at what they do. Embracing change can be difficult, but the changes we've discussed are worth the effort.

Jeremy: If you're not already improving your data and analytics platform, get going. If you are, redouble your efforts and realize more value sooner. There's never going to be a perfect time to modernize core systems or implement the data and analytics solutions we've discussed. And as mentioned, it's a mistake to wait for core systems modernization to conclude before getting serious about analytics. So build out applications, data management, and analytics capabilities in concert. Take a modular approach, governed by a technology architecture that enables you to add on and scale up. Prove the value of additional analytics as you go, and gain momentum with each success.

Dan: Intensity varies by state, but the overall trend has the public paying more and more attention to the environment and exerting more pressure to protect it. The importance of environmental compliance agency work will continue to grow, as will the scrutiny on agency performance. I strongly recommend getting ahead of the game with the help of data and analytics. You don't want to be caught unprepared or placed in reaction mode when lawmakers, government officials, and the public demand more information, explanation, accountability, and transparency. Better to anticipate the demands for information and analysis and develop a capable and flexible platform to meet it. Better to be proactive than risk being scapegoated if something goes wrong.

Jeremy: If there's one agency within most states that is really focused on science and technology of many kinds, it's environment compliance. Scientists and technicians in the agencies should have the best tools, including data science and analytics, at their disposal. They want to use science and technology to the advantage of the agency, the permittees, and the public. How can an agency perform its most fundamental mission—protecting public health and well-being by protecting the environment—if it doesn't have the systems and technologies to keep pace? Advanced data and analytics are increasingly mission-critical.



DAN CHILDERS

After several years as a solutions architect at SAS, Dan Childers has designed and supported data-driven, analytical solutions built on top of the SAS Viya platform for government clients, including but not limited to the Delaware State Police, Florida Department of Financial Services, Florida Department of Emergency Management, and Florida Commerce.

Ultimately, Childers's mission is to help push the world forward by solving some of the world's most vexing problems through innovative, practical analytical and AI solutions. For example, he is the lead architect and designer of SAS Environmental Compliance.



JULIE ESPY

Julie Espy is a senior industry consultant with SAS, where she supports engagements with local and state governments on water quality and environmental topics, provides insight on water and environment programs and policies, and promotes the use of data analytics to solve complex issues and implement environmental policies. She has almost 30 years of experience working in environmental programs with 23 of those at the Florida Department of Environmental Protection (FLDEP). In her role at the FLDEP, she had a variety of responsibilities over her long career. They included administering monitoring and assessment programs, managing environmental data, developing water quality standards, establishing water quality restoration goals and implementation plans, and advancing environmental policies.



JEREMY RAY

Jeremy Ray is a senior industry consultant at SAS where he specializes in providing subject matter expertise related to multiple governmental focus areas that include insurance, labor and employment, social benefit programs, and environmental regulation.

Prior to his time at SAS, Ray served over ten years in various roles in state government in North Carolina, where he served as a county prosecutor, agency staff attorney, and most recently as a legislative analyst for the North Carolina General Assembly.

Ray has a BS in Criminal Justice from Western Carolina University and earned his law degree from Elon University.



ROBERT MORISON

Robert Morison serves as senior advisor to IIA. An accomplished business researcher, writer, discussion leader, and management consultant, he has been leading breakthrough research at the intersection of business, technology, and human asset management for more than 25 years. He is co-author of *Analytics At Work: Smarter Decisions, Better Results* (Harvard Business Press, 2010), *Workforce Crisis: How to Beat the Coming Shortage of Skills And Talent* (Harvard Business Press, 2006), and three Harvard Business Review articles, one of which received a McKinsey Award as best article of 2004. He holds an AB from Dartmouth College and an MA from Boston University.

IIANALYTICS.COM

Copyright © 2024 International Institute for Analytics. Proprietary to subscribers. IIA research is intended for IIA members only and should not be distributed without permission from IIA. All inquiries should be directed to membership@iianalytics.com.