

A Roadmap for Enabling Integrated Science: The USGS Experience in Open Data

Twentieth International Conference on Grey Literature "Research Data Fuels and Sustains Grey Literature"

> Viv Hutchison US Geological Survey Science Analytics & Synthesis (SAS)

> > December, 2018





Viv Hutchison

- USGS employee since 2002
- Background in Library and Information Science (MLS from University of Maryland)
- Current Position(s):
 - Science Data Management Branch Chief
 - Acting USGS Library Director



US Geological Survey

Science for a Changing World

The USGS serves the Nation by providing reliable scientific information to describe and understand the Earth; minimizing loss of life and property from natural disasters; managing water, biological, energy, and mineral resources; and enhancing and protecting our quality of life.

≥USGS US Geological Survey: Organization science for a changing world

7 Mission Areas

- **Core Science Systems** ٠
- Ecosystems
- **Energy and Minerals**
- **Environmental Health**
- Land Resources
- Natural Hazards
- Water
- 7 Regions
- Over 90 Science Centers + Research Stations (totaling over 400 facilities)
- Science Support Offices

Organization

iew Region Ma

We employ the best and the brightest experts who bring a range of earth and life science disciplines. By integrating our diverse scientific expertise, we understand complex natural science phenomena and provide scientific products that lead to solutions. Every day the 10,000 scientists, technicians, and support staff work for you in more than 400 locations across the United States

Mission Aroos

Our Science Strategy outlines the major societal issues that USGS science is poised to address. We've also created pecific strategies for each of those areas to expand and advance the actions we can take in the next decad

Science Centers

am more about the research and monitoring that are conducted throughout the country by an individual Science nter. Find out more about what science is being done in your area.

Regions

We employ over 9,000 people who work in every State across the nation with Regional Executives in the Southeas Northeast, Midwest, Southwest, and Alaska. Throughout the regions across the United States, there are science centers that monitor, assess, and conduct targeted research on a wide range of natural hazards, ecosystems, and

Science Support

The amazing science we do at the USGS couldn't be done without many people handling the support functions of our nission areas and their programs

US Geological Survey Core Science Systems – *Mission Overview*

CSS continuously strives to better understand, characterize, synthesize, and apply context to the complex processes and interactions taking place in the **Critical Zone**.

U.S. Department of the Interior U.S. Geological Survey America relies on our broad range of nationally-focused Earth system science, geospatial research, and foundational data to remain informed, safe, and healthy.

- Provide baseline geospatial, geological, and biogeographical data for the Nation's topography, natural landscape, and human-constructed environment.
- Conduct science, surveys, and research on the Nation's geological and biological resources and offer data, map services, and decision-support tools for the public to easily discover/use for local, regional, large-scale, national, or continental analyses.
- Coordinate effective and economical use and management of geospatial data assets for the government and the Nation.
- Conduct data synthesis and analysis across disciplines to enable data driven science in support of energy and mineral assessments, infrastructure modernization, natural hazards mitigation, land and water management and conservation, and national security.
- Connect USGS scientists to published literature, data, and collections to support evidence-based practices.
- Inventory and preserve geological and geophysical data collections to provide a framework for geoscience data and information sharing.
- Improve Federal-State cooperation and collaboration by effectively leveraging partnerships.
- Provide computational capabilities through high—performance computing infrastructure, workshops, and expertise.

Science Analytics & Synthesis (SAS)

SAS Mission

- Accelerate research and decision making through data science, information delivery, advanced computing and biodiversity analytics.
- **Key Science Themes**
 - Data Science
 - Biodiversity Science
 - Advanced Research Computing
 - USGS Library
 - Science Data Management

SAS Science Data Management Branch

providing tools, best practices, and education for data management

Science Data Tools

USGS Science Data

- Multiple disciplines of Earth and biological sciences
- Distributed environment in many Science Centers
- Labs, field collection, instrument collection, mapping, remote sensing, etc
- Many data formats
- Wide variety of data sizes
- Data integration needs

Modeling wave heights due to surface winds in Hurricane Sandy

Large River monitoring & mapping

Example of USGS science data in action

The Event(s)

- Lava flows travelling at 3 ft/s (2 mph)
- Ash plumes to 10,000 ft from main vent
- 24 fissures, 19 active (mid August 2018)
- Major Damage to Leilani Estates
- Encroached on a hydrothermal power plant

Fissure 8 - May 28, 2018

Science for a changing world

Rapidly producing new Coastline

Damage in Volcanoes NP

Before...

Science for a changing world Lava Flow Hazards & Damage

Hannah Dietterich, USGS Volcano Hazards Program, Alaska Volcano Observatory

- Tracking new vent formation on Kilauea
- Estimating potential damage impacts due to lava flow and inundation
- Multiple-agencies involved FEMA, USGS, BLM, NPS, NASA

Rapid data analysis using:

- Lidar and Unmanned Aerial Systems (UAS) data collected to update Digital Elevation Models (DEM)
- High Performance Computing

Lava Flow Hazards & Damage

Angie Diefenbach, USGS Cascades Volcano Observatory

- Developed models based on imagery captured with Ricoh GII camera using Firefly Pro Vertical Take off and Landing (VTOL) fixed wing UAS
- 4,008 images (25GB) from UAV mounted camera generated several "products"
- Dense Point Cloud with 1,091,421,418 points (27GB)
- Orthomosaic (2.1GB)
- DEM (3.0GB)

≋USGS

science for a changing world

- Data products used to understand the extent of the caldera collapse at the summit that occurred during the 2018 eruption
 - As much as -518m change in elevation (slumping) has been observed at the summit

USGS Publications Warehouse

The Pubs Warehouse provides access to over 150,000 publications written by USGS scientists over the century-plus history of the bureau.

Science for a changing world

Federal Lands Greenhouse Gas Emissions and Sequestration in the United States: Estimates for 2005–14

Scientific Investigations Report 2018-5131

U.S. Department of the Interior U.S. Geological Survey

USGS Publications Warehouse

- Authoritative resource and public interface to USGS authored publications
- Provides metadata and links to USGS funded and authored publications

https://pubs.er.usgs.gov

• USGS Library leads Pubs Warehouse

Publications Warehouse Content

Published by USGS	Published by external entity
 USGS numbered series reports Other USGS reports 	 Journal articles Conference proceedings Books Book chapters Other government reports

Publications Workflow

U.S. Department of the Interior U.S. Geological Survey

≊USGS

Science Publishing Network: Services

Science Publishing Network (SPN)

- Product planning and consultation
- Publication policy advising
- Map editing and production
- Scientific text editing
- Illustration
- Graphic design and layout
- Online publication
 support

U.S. Department of the Interior U.S. Geological Survey

Use of Science Publishing Network

Requires approval by Regional Communications Chief (Office of Communications and Publishing) at design and concept stage. Pinformation Periodicals, Annual Financial Report of the U.S. Geological Survey, National Water Conditions, Minerals Information Reports, and Preliminary Datermination of Epicenters Monthly Listing. "Other publishing services available. "Requires justification form.

Publications + Data Release Process for Authors

Overview of the Publishing Process for USGS Series Publications Accompanied by a Data Release

[Any dataset used to support the conclusions in a series publication must be released simultaneously with the publication, unless the data have been previously released. Abbreviations: BAO, Bureau Approving Official; DOI, Digital Object Identifier; FSP, Fundamental Science Practices; GPO, Government Publishing Office; IPDS, Information Product Data System; mgmt, management; OSQI, Office of Science Quality and Integrity; PDF, Portable Document Format; PW, Publications Warehouse; QC, quality control; SPN, Science Publishing Network; USGS, U.S. Geological Survey]

U.S. Department of the Interior U.S. Geological Survey

²For an administrative report, this workflow is followed to the "Address BAO comments" step. Administrative reports document studies that are released to other Federal agencies. Administrative reports are considered to be unpublished. ³The report DOI is assigned automatically in IPDS after the author approves the "Prepare for SPN Production" step.

⁴Only interpretive products go to the BAO; noninterpretive products are granted Bureau approval by center directors or their designees.

Drivers: Government Initiatives

Broader Ecosystem of Government Mandates:

- Open Government Initiative (2009)
 - transparency, participation, and collaboration
- Digital Government Strategy (2012)
 - "Digital Government: Building a 21st Century Platform to Better Serve the American People"
- Open Data Initiatives (2013)

Open Data Initiatives – 2013

Executive Order Increasing Public Access to the Results of Federally Funded Scientific Research Feb. 22, 2013

Executive Order Making Data Open and Machine Readable: the New Default for Government Information May 9, 2013

M-13-13 Open Data Policy

• Managing Information as an Asset

Highlights:

- Data must be accessible (in open formats)
- Data must be described
- Data must be reusable
- **Complete** (published in "primary form" with finest level of granularity)
- Timely
- Managed post-release

Science for a changing world

Becoming a Global Open Data Movement: Governments around the world

Department of the Interior: Open Data Coordination Efforts

Coordinated Interior metadata contributions to data.gov

Sponsored Cross-Agency Working Groups:

- Data Governance Board
- Data Services Team
- DOI policies
- Information Management & Technology Strategic Priorities Groups (e.g., "Managing Data as a Strategic Asset";
 "Risk Based Portfolio Management"...)

M-13-13 Open Data Policy

USGS Response, published in 2016:

"Public Access to Results of Federally Funded Research at the U.S. Geological Survey: Scholarly Publications and Digital Data"

Public Access to Results of Federally Funded Research at the U.S. Geological Survey: Scholarly Publications and Digital Data

USGS Public Access Plan

- Regarding data, the Plan introduces a "new normal":
 - USGS scientists must release the data upon which their scientific publications are based.
- This requires, for data:
 - A detailed **data management plan** for all research projects
 - Data in machine readable, open formats
 - Standard metadata describing the data
 - A digital object identifier for the data, recorded in the metadata record
 - Approvals: requiring data review, metadata review, and either Center
 Director or Bureau Approving Officials
 - Hosted and shared from a reliable, **repository** location
 - Metadata shared through the USGS Science Data Catalog

U.S. Department of the Interior U.S. Geological Survey Data

Pub

Review & Approval

Publish

- Implementation of the Public Access Plan in the USGS
- The "new normal" was a surprise to some.
- It requires a new workflow for scientists.
- It represents a culture shift for USGS.

Building on Foundations

Fundamental Science Practices (FSP)

Fundamental Science Practices

- Established 2006
- FSP is a set of consistent practices, philosophical premises, and operational principles to serve as the foundation for research and monitoring activities related to USGS science
- FSPAC: FSP Advisory Council provides guidance/direction for FSP implementation

FSP includes USGS Policies

Fundamental Science **Practices**

⊠US6S	18-38-2		Real Property in the second second		URGI Notice Concess UPDB Bearth VIDD
G.S. destegical S	urvey Nanual				
0.0. Geological Servey No: Dr Diddi 2015.41	Debuctional Renormalism				
Imuteria Data Petroa	ry 14, 3815				
Experiation Date: Relati	- Delli Raspendial				
Subject Scientific but	a Runayanant Foundation				
Auropenet reported	≥USGS		A REAL PROPERTY AND A REAL		UTION Name Contact UTION React UTION
preside by operant 3. Pellos, 1003 exert	U.S. Geological Survey I	Nanual I			
proceeduries that support (H (MSQ) 2015-62, (H (SR)	U.S. Seclepical Burvey Instruct	tend Removandum			
4. Elemente el ber 10	New 104 00001 2055-02	3211			
ana tompe (002.1m)	Second Date Petruary 16, 20 Second Second Second Second	its			
 Algaria Reliada en Agarante reliada, ap 	Automit Matadata for Scientifi	to Data, Balturara, and Other Information Produ-			
C. Provis. Data provis	L. Purpose. This Instructional To mandemix ensured by the Palenal	encrustum (IM) provides interim palso on metablish I Geographic Tetra Committee (NGOC), the interapero	regumenterite for 3.3. Geological Survey (UADE) scientific data, adfessor, and after internation (committee that provide metodate guidance for all federal Ecveniment scientific data (propola	products. 3 area provides guidance for completing with appropriate URGE at and new properties. This meaning policy is security before the time new	and other Holenz standards, such as the metadate and the USES science activities to fully implement the
Per 646 (28 502 2)	A Backwood Core Da	11000			utidi Rene
Aptrovativities are resp	Rundation.	USUS ET MAL		Contract of the second s	Cardial Data
 Autoritoria, 1003 bando USB interdate, metadate 	A. Patrip.	Geological Survey Hanual			
E. Describe (Minister), The back, and Information	A Relation had accent probably and pipeling with the probably from the cost of the De-	alogical Survey Instructional Hemorendum 0901 2015-43			
n. Altrian Quality. Date	S. J. Names of starts Disaster	in Date: Petroary 18, 2015			
Ropes of the scance of 1. Review and Servers	Data Autom (PDS) mat. Engine	ran Date: Autain Until Surgended			
Mode as accedences	A. Creating Matadata: Bullyon A. Data. Matadata must S. Parts	 Review and Approval of Boundity Data for Ran pees and Scope. The Instructional Networkship Directory. 	een C provides interim requirements and procedures for review and approval of scientific data provide	- research or departmentation. The Dr applearing at U.S. Santappar Survey (2)	ISSE assembly data that any released to the public locksdare, but not instant to page
X. Aqualate Directors at	or optiglied by USQS any that are a data products is uppened. Detables	 main analisis in Delevato, Antohuses, Solo services, an. The part of SA 300 34 that addresses Computer 	and publications, an well as model subjects and derived products. The DF replaces the period to be regresse will be addressed arguments. This DF is interve policy to allow the time massied for USD	rves Namual (3H) chapter 3H 300.24 - Holly for Netwook of Computer Dec 3 science activities to fully implement the fails release requirements here	alkaan and Computer Programs, dated April 9, 0402, that abbrears Computer in and it will be retained until supersected by a permanent Sumay Menual (SM) policy
8. Office of Assess put	the data and information.	spreame, to work new country to an senar any	1000 are the property of the Federal Doversment. 1000 scientific data are made available aut	thing and have of conductions, second in rank cases, where access must be a	sended because of security, privacy, confidentiatio, or other conditions. 3 is the
Nordentet with the Alex Nordentet mit alex	COMPARIANT INCOME NO. OF STREET	The USOS to provide their public arrows to undertain a information of a way that autopic's internationant only and Tachesters Patry 2.	data, influenzation, peri lasheningan deveraped ita the Bureau's influenzation and reasons program manue processing and magnetization activities. This includes using machine reachable and pairs for	re. The Ray 8, 2013, Office of Nanagareest and Subject (2082) memorant formate, data attendants, and common core and extransitie metadate for at	Sure on "Open Data Policy- Ranaging Data as an Asat, "regures agencies to unlest I nave of prevalent counting and collection efforts. The Fabruary 22, 2015, Office of
C. Assess Control Texas data Assessment activation 2. Approvemp Diffusion A assessment of the Astas assessment	CD Metallola annuella in CD Metallola annuella in 5 Granaure mare robuit, n. A. Cola				UBS Name Cardiat UBS Namit UBS
	(1) Metaclaria name in the	net, Autochared safest U.S. Geological Ser	rey Manual		
	the utilit Science Date C & Date	Darves. At artise rap as the cost 2015-04	dructional Reneration		
	& Software, Metadate I E. Prov	manar Data. These data beauting better Petersony	a, 2010		
	for the adhees or muro adheas, including entry 4. Peter	Expiration Date: Botale	uli Euspendud		
	C. Other Johnmatton Phy & UBDS Fi	5. releases both process underseried Science Proc. 1. Purgess and Science. 1	in Johnschmut Mennenstein (24) specifies preservative requirements that apply to all U.S. One descentions will experient the assessment Access Revue (Met and Access) File	right formy (USD) optic scentific and assessed offernation. I	The IV is interver party to allow the line method for USDS science activities to fully implement the data preservation where to ensure the science and utility database preservations. There are not ensure for the database to be intervented and intervented an
	CE Meladada recordo for E. April	mond scientific data must restance containing data or i mon of scientific data and 2, Pedage. Cata created in	experience of label) or physical samples are abbreviat in the USDE records its position schedules or an industrial the USDE are the presenty of the Asterial Soverment (refer to <u>SM SSE)</u> for prior	matter in USES and new USES cate) and the finderal beamin Ad 20.07	1102.24). 3 is the policy of the LSCE to preserve scientific data and information funded or developed by the
	are depicited in and due 10. App	meet acientific data mus Auto acientific data mus A USGS scientific data data	eans propuers, no science sets as a result of 1555 funding must be preserved as followed deets and information, and accessible, available, and caseline on the appropriate meths in accord	serve with the 1955 means dispersion schedules (http://www.usu.com	(and manual schedule) or as appropriate to more the flattanal Archives and Tecords Administration (MAA) reservice
	a sector	a at the USOS PSP Bed - format requirements for per	nament records (http://www.anthreel.ancheurolit.ment.battochander-publices.batted.html unbertrative or provide copy of all data for which it is reasonable and that are structured for obse	burded threaded and to comply with Persent Exveniment Court Data and	Contraction inflations, seeks of child data have be described and through sectors will be the sector and an end of the sector and
	5. The	prostures and putters an obligation of the	senter data ara pinat-car in angenetice alle non-citité existen, consente d'inte data rout b a alte non-citité antéene muit alle de lational (vier la <u>25,102,7</u> section), faitemente).	e dearly stated to a tocurrented agreement between the parties. In add	tion, other applicable requirements for scientific least performed and internation developed under various aprentents
	302.21	tese and Approval of B C. Optic data and associate such as these midded to fac	el enternation that chick is reasonable for preserving that be assed and reasoned from property Tableurose, data tribuptio, and information saturity are addressed (who to <u>this chicks and</u> sold allo	s centred and accentred internation systems following LSLB information class menuscriptically 32(1):1	increasely summa process and providens. These processes share important long-term preservation (process
	A. A. Yu (197 102	menut of two reviews a I. Preservation Demote	 Tach alement loads below represents a component of required diplot scientific data preservation option. This is an all secondaries function strategies and area to alternative and the spectrum. 	on. Detailed guidance and specific implementation totals are available or the doubled process of rate. All a monotonic has contracte comes of all loss.	n der 1925 Des Nangement Weis aller <u>Hits Communitier and Lancidiation and Tal amount info</u>
	9. As #	a Sector Specified on he is	anel in a totallo-fiel & algorith from a Science Center, which is the primary location of the sil- ter . Tota coldific and interacts economics science on a memory in addition in science and one	ng teamainte. Ann de annuels inne . The includes anticiding to compared defend and second	then consisted influence a interfaced or comparison channel to take as and as articular the comparison in
		the attenues analysis and the chemical analysis of data through tech anisotre and mails	 Longital or sthe changes over time. Scanges include unitying the imaging of head intervent. 	mantaning providents to replace or report complete data, responses or i	one, therefore a new lots to new screpcing hardware from deputing a should have and appendix sectors
	publical	tore, and so on) C. Solumeter Security: 5 D. Networks, Medicality and	lementer security includes presedunts in present human second compliant or defense of any un unsu present decumentation of Rumas, data to making contentual understanding and lines term can	sucherant assess to the data. Note based authorization applies with reg- annity traffer to 24 0500 2015 00 - Weigelaid for Sciencific Data. Software,	ant to mail, write, move, and deeter actions on inductive flee. and Other Orderspace Productor. Considers metabolic most be preserved and directly associated with the scientific
		data they describe		and all as the shift and seen an bill been to see a	
		1003 encouraged the use of antinears for eccess and use	spec the formatio, codece, compression schemes, and encapeutation schemes. These formats at	tel met the Office of Hanagement and Budget principles of ensuing mail	hte natidity. Negraphical, oper ida fernas are prevent our proprietary formationpuring quantum
U.S. Department of the Int	terior	A. Responsibilities.			
		allofas.	where exercise reasons measure here any probably present (and based manufacture	and the second sec	AND AND ADDRESS OF ADDRESS AND A DRESS AND A DRESS ADDRESS ADDRE
U.S. Geological Survey		 Office of Science Quality of USGS styles sciencile de discovery of USGS topical is 	and Jobpins and Cire Romon Evenue. The Office of Summit Quality and Integrity (35Q) and 2 assets. The CSS maintains the USOS Data Nanagament Riels also, which principle guidence an entity out.	The Care Science Builterie (232) are required in the party developing of a providure related to preservation of allocities date (relative to 2022) and 2022 are stated in the second state of the second stat	Is preservation policy. The DOC is required the intercenting and communicating policy regarding the preservation and the intercention of the intercenting of the USCS Science Over Casing, which provide policy reaction and additional and additional and additional additiona Additional additional addit
o /		 Science Cardier Checkler requirements in this patient is tarts improvident dramapy at 	Service Center Diversion in their responses as they always increasing, and advanted with their responses part of the total management prior to account for preserving the total and information and the re	ettive Alta ancise Alta to assign ar instant the antipolog of data managers attaches in accordance with this policy and records dependent authorizes.	to survey that under and affect this proprieties advices. A the start of the propert, the sciented must develop
		di Unio Menugera, Design the agentamate records ma	ted data managers are requireden to constructing and execting preservation activities for USG agained, requirements for their data are mat.	t bits. Deta menagers wit ensure that the preservation activities for who	It this are assigned are risk. This also combinate with the data produces and the USGS Records Office to ensure
		E. Assarb Office: The Ut and samplete results, man	 Neurola Office: In the Office of Enterprise Internation is responsible for the policy dependence along these movies throughout the science total illeptice, and legally transferring permanent as 	rc. continuous, and oceral management of the USOS Recents Nanageme antific late in accordance with applicatile USOS and SMAI records management in accordance.	ent Program and amounts potoles, italitants, and processes are in place that provide guidance on creating accurate primetil requirements.

USGS released 4 new policies for data management & public access

- Scientific Data Management ٠ Foundation (requires DMPs)
- Metadata for Scientific Data, Software, and ٠ Other Information Products
- Review and Approval of Scientific Data for ٠ Release
- ٠ Preservation Requirements for Digital Scientific Data

Fundamental Science Practices

 The new data policies and accompanying tools are based on the model

Bureau Approving Officials (BAOs)

Bureau Approving Officials

- Team of science experts
- Office of Science Quality and Integrity
- Provide approval for release of interpretive science information products including publications, data, and others

	Approval levels for information products prepared by USGS authors								
		Bureau App	oroval						
Information Description		Science Center Director or equivalent	BAO in OSQI						
Circular	Synthesis based on approved published products. This series should not be used to present new data or new scientific findings.		х						
Data Series	Data, multimedia, and (or) noninterpretive.	х							
Open-File Report	Interpretive and preliminary or interim release pending formal series release.		х						
	Data, multimedia, and (or) noninterpretive.	х							
Administrative Report	Data (studies and findings) and (or) noninterpretive; proprietary interest to Federal agency cooperator.	x							
	Interpretive or policy sensitive; proprietary interest to Federal agency cooperator.		х						
act Sheet	Interpretive and based on approved published products, not policy sensitive.	х							
	Policy sensitive.		х						
	Noninterpretive and (or) programmatic.	х							

Science Publishing Network (SPN)

Science Publishing Network

- Distributed team of editors and publishing specialists for USGS
- Oversee internal publishing process for USGS scientists
- Coordinate with the Publications Warehouse team

Community for Data Integration (CDI)

Community for Data Integration (CDI)

Communities of practice are groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly – Etienne Wenger, 2011

U.S. Department of the Interior U.S. Geological Survey Purpose:

- Advance understanding of Earth systems through enhanced use of data and information
- Provide a forum for sharing ideas and learning new skills
- Increase the visibility of work with data and information
- Over 1,000 members
- Monthly Meetings
 - Open to all (inside and outside USGS)
- Working Groups
 - Data Management Working Group, seeking to elevate the perceived value of science data management

Executive Sponsors: Kevin Gallagher and Tim Quinn

Science for a changing world Building on Foundations: Tools and Applications

Data and Code Management Tools

Publications Warehouse

Review and Approval Tools

- Open Researcher and Contributor ID
- Unique identifier for people
 - Helps with name disambiguation
 - Per Public Access Plan, now required for all publishing authors

U.S. Department of the Interior U.S. Geological Survey ORCID

Connecting Research and Researchers

ORCID as a URL

http://orcid.org/0000-0003-2609-364X

Eighteen years (1996-2014) of channel cross-sectional measurements made in Spring Creek after the 1996 Buffalo Creek wildfire and subsequent flood John A. Moody, Deborah Martin 2017, Dataset

The consequence of the 1996 Buffalo Creek wildfire disturbance and a subsequent high-intensity summer convective rain storm (~100 mm h-1) was the deposition of a sediment superslug in the Spring Creek basin (26.8 km2) of the Front Range Mountains in Colorado. Changes in the superslug near the confluence of Spring...

Synthesis of soil-hydraulic properties and infiltration timescales in wildfire-affected soils

Brian A. Ebel John A. Moody

2017, Hydrological Processes (31) 324-340

We collected soil-hydraulic property data from the literature for wildfire-affected soils, ash, and unburned soils. These data were used to calculate metrics and timescales of hydrologic response related to infiltration and surface runoff generation. Sorptivity (S) and wetting front potential (Ψ f) were significantly different (lower) in burned soils compared with...

- Internal to USGS
- Authors, BAOs, ٠ Center Directors use IPDS to track publications;

New Product

Record

Number

Center

New Report

Saved Reports

- Per Public Access ٠ Plan, updated to include:
 - data release ٠ review and approval capability
 - dark archive for ٠ manuscripts

U.S. Department of the Interior U.S. Geological Survey

Information Product Data System (IPDS)

IPDS Authors -	New Item			
				-
IP-05123	9			
	Rank	2		
0	entributor Nole	Author	¥	
	oncio	1214		
USGS A	ther			
	Author Name	Bibhuti, Nitesh i	8,- W	
	Cost Center	Alaska Climate Science Center	¥	
Nen-US	SS Author			
Non-US	GS Contributor			
Non-I	1505 Affiliation	une ten jannar: une, sone s		
		SAVE CANCEL		
				18

ORCiD field

 Designed around science data lifecycle model

> Best practices, tools, links to policy

PUBLISH / SHARE

MANAGE QUALITY

BACKUP & SECURE TRAINING & RESOURCES finish

DESCRIBE /

METADATA

The USGS Data Management Website is organized according to the

USGS Scientific Data Lifecycle Model, which describes the stages of data management and how data flow through a research project from start to

U.S. Department of the Interior U.S. Geological Survey

https://www.usgs.gov/products/data-and-tools/data-management

Stewardship

Data Release

Manage Quality

USGS Data Management Website

Data	Data	Information Product Data System	FSP Approval
FGDC-compliant Metadata	Metadata	USGS Dataset Repository E	Reliable Location
Digital Object Identifier	DOI	Science Data Catalog	Shared & Discoverable

USGS Metadata Tools

Online Metadata Editor (OME)

 Allows scientists to create and manage metadata records that describe their datasets science

- Uses common questions about the data
- Based on FGDC standard (something hidden from scientists)
- Validate and download records
- (internal to USGS)

USGS for a changing world							U C S	SGS Home ontact USGS earch USGS	-	•	Ρlι
eological Survey - Core Science A	nalytics, Sy	nthesis, a	nd Libraries - Onli	ne Metada	ta Editor (OME)				-		All
&L / My Dashboard: vhutchison@usgs.go	ov / Logout				Ne	ed help or I	have questions? C	ME Service Desk			
nt Standard for Digital Geospa gical Data Profile	tial Metada	ata 🛛	OME performs an auto after any lengthy input	matic "Docume of information.	nt Saved" action every 10 min	nutes; howev	er, we recommend cli	icking 'Quick Save'			me
			Save and Preview Fo	rmatted	/alidate & Download as X	ML Sa	ve & Close Record	d Quick Save			the
e of the Dataset (Include "where," "wha ample: Aquifer Systems and Recharge Pot	at," and "when tential in Louisi	n" in the title iana from LDI	a): EQ source data, geogra	phic NAD83,	LOSCO (1999)						Ba
red	me & Place	Keywords	Entity & Attributes	Data Ous	lity Applytical Tools	Distributi		Extra	-	•	Pu
	ine a riace	rteywords	Entry & Autoutos	Dutu Qut	ing Penalytical Tools		Entity and Attribu	te Ruilder			
Dataset Author(s): Please use the "Add Author" button for m	oultiple authors.						,	Click Through Each Review/Update Its L	Attribute and Description	_	Attribute Defini Type of solar p
Add Author							Attribute (Field) OBJECTID	Type Long Ir	teger		
Data Point of Contact							SIZE_MW	Charac	ter string		Attribute Defini
Contact information for the individual or o	organization the	at is knowleds	geable about the data.				LOCATION	Charac	ter string		Field Type
Name:	tact						STATUS	Charac	ter string		@ En
Provide the name of the individual to com	1001.						SHAPE_Leng SHAPE_Area	Double	precision floating precision floating	р р	Unique Valu
Organization:							YearOnline SourceInfo	Integer Charac	ter string		Concentrate Parabolic T
							Polygon	Charac	ter string		Photovoltai Photovoltai
Address											Photovolta
*Street:							Overview Description	n (optional)			
*City:			*Co	untry:			Please see the deta complete description attribute table.	led Entity/Attribute se n of all the fields conta	ction for a ined in the	* H	
							Please note that two	facilities are included	I in this data set		
							Datail Citatian (antia		o deniorio notas.	*	
							See the detailed En	ity/Attribute section.		*	
										-	
								Save			

Metadata Wizard

- Plug-in tool for ESRI ArcGIS tools
- Allows scientists to create metadata records that describe their geospatial datasets
- Based on FGDC standard
- Publicly accessible

ity and Attribute Bu	ulder					
Click Revie	Through Each Attribute and w/Update Its Description	Attribute Defin Type of solar p	ition power technology employ	yed by facility		
nbute (Reld)	Type					
CILITY ZE_MW	Character string Character string	Attribute Defin	ition Source Producer	defined		
PE	Character string					
ICATION TATE	Character string Character string Character string	Field Type	numerated	Range	Codeset	O Unrepresentable
IAPE_Leng IAPE_Area	Double-precision floating Double-precision floating	Unique Val	ues	Definition of this	value	
arOnline urceInfo lygon	Integer Character string Character string	Concentral Parabolic Photovolta Photovolta Photovolta	ed PV Trough ic ic / Concentrated PV ic Storage Project	Parabolic troug	h concentrating solar power	(CSP) technology being used at site.
				Definition source	e	
				Producer define	ed	
erview Description (op	tional)					
ase see the detailed Er iplete description of all bute table.	ntity/Attribute section for a the fields contained in the	* III		"Enumerated This attribute ty entries for a fiel appear in the fiel	Domain" pe should be used when a fi d. Provide a definition for eac eld will be an entry from the e	nite set of values exist as possible ch possible entry, All values that numerated list.
do not contain information	ties are included in this data set ation for all of the attribute fields.	-		Example: a field Duty', and 'Trai	d named "RoadType". Possit I'. Three value definitions wo	ole values are "Heavy Duty", "Light uld be provided that clearly explain
ail Citation (optional)				the criteria for e	sach of the three classificatio	ns.
the detailed Entity/Att	inbute section.	*				
5	Save		Save and Close		Can	cel (Close no save)

U.S. Department of the Interior U.S. Geological Survey

sciencebase.gov/metadatawizard

Robust metadata records are produced using OME and/or Wizard:

- Metadata describes data: who produced it, how (what processes were used), data quality testing, entities and attributes, etc
- Metadata records include digital object identifiers

U.S. Department of the Interior U.S. Geological Survey

Historical methyl mercury in San Francisco Bay

Metadata:

- Identification Information
- Data Quality Information
- Entity and Attribute Information
- Distribution Information
- Metadata Reference Information

Identification Information:

Citation:

Citation Information:

Originator: Steven E. Schwarzbach Originator: Colin A. Eagles-Smith Originator: Colin A. Eagles-Smith Originator: Michael L. Casazza Originator: Julie L. Yee Originator: Julie L. Yee Originator: Alan C. Heyvaert Originator: Thuy-Vy D. Bui Originator: John Y. Takekawa Publication Date: 2015 Title: Historical methyl mercury in San Francisco Bay Geospatial Data Presentation Form: Tabular Digital Data Publication Information:

Publication Place: Sacramento, CA Publisher: U.S. Geological Survey

Online Linkage: http://dx.doi.org/10.5066/F78P5XKP

Description:

Abstract:

San Francisco Bay, California is considered a mercury-impaired watershed. Elevated concentrations of mercury are found in water and sediment as well as fish and estuarine birds. Sources of mercury to the watershed since 1845 include sediment-associated mercury from mercury mining, mercury losses from gold amalgamation activities in mines of the Sierra Nevada, aerial deposition of mercury from global and regional emissions to air, and the direct discharge of mercury to Bay waters associated with the urbanization and industrialization of the estuary. We assessed historical trends in mercury bioaccumulation by measuring mercury concentrations in feathers of the endangered California Ridgways rail (formerly California Clapper Rail) using museum specimens. We developed a structural equation model to attribute variation in historical mercury bioaccumulation in rails to sources of mercury, concentrations.

Purpose:

Because of the many historical and contemporary sources of mercury to the Bay, we sought to answer two questions: 1. How has mercury bioaccumulation in the Bay changed over time? And, 2. Can we attribute the observed variation in bioaccumulation over time to different anthropogenic sources of mercury?

Time Period of Content:

- Allows creation of an identifier for datasets
- Connection to DataCite
- Identifiers for USGS data only
- DOIs are used in citation and in metadata records

Science collaboration communities, data and metadata, web services, APIs.

Leveraged for repository, data release platform for USGS researchers

Per Public Access Plan, became a Trusted Digital Repository for USGS

sciencebase.gov

U.S. Department of the Interior U.S. Geological Survey

USGS ScienceBase

Base-Catalog Communities	Help +		
	Type some text to search	Search Advanced	Search
I want to: - Login - Add Data - Access Help - Report a Problem	Browse by Category - Map - Data - Physical Item - Project - Publication - Web Site	Browse by Tag Animal Behaviour Biogeochemistry Ecosystems Hazard Mitigation Hydrology All tags	Browse by Location
	View USGS data	releases in ScienceBase g a data release in Scienc	eBase
	Fea	atured Item	
Long-term effects of management in the (Greater sage-grouse (Centrocercus u) fragmentation of sagebrush ecosystem Act multiple times, and was most rece frequency and changing climate, which threat assessment within the Great Ba hypes: Clatton, Map Service, OGC WFS Lay	wildfire on greater sage-gro Great Basin rophasianus; hereinafter, sage-grouse) ma across most of its geographical ran; ntly ruled to not warrant protection as hare identified frequently as two envir sin for listing sage-grouse in 2015 und er, OGC WMS Layer, OGC WMS Service; Tage	use - integrating population ar are a sagebrush obligate species that har g. The species has been considered for of September 2015. Nevertheless, the sp onmental drivers contributing to declines or the I: Fire Suppression, Great Basin, Greater Sage-G	nd ecosystem concepts for a declined concomitantly with the loss and listing under the federal Endangered Species acies faces threats from increasing wildfire of sage-grouse populations. To help inform a rouse, Habitat, Monitoring Trends in Wildlifre Severity

U.S. Department of the Interior | U.S. Geological Survey URL: https://www.sciencebase.gov/catalog/

Privacy

FOIA

Policies and Notices

USA.gov

Data release landing page

> Note the data, metadata, citation, digital object identifier...

U.S. Department of the Interior U.S. Geological Survey

System → USGS Data Release Products → Historical methyl mercury in ...

Dates

Publication Date : 2015 Start Date 1850 End Date -2007

Citation

Schwarzbach, S.E., Ackerman, J.T., Eagles-Smith, C.A., Casazza, M.L., Yee, J.L., Heyvaert, A.C., Krabbenhoft, D.P., Bui, T.D., and Takekawa, J.Y., 2015, Historical methyl mercury in San Francisco Bay; U.S. Geological Survey data release, http://dx.doi.org/10.5066/F78P5XKP,

Data Citation

Summarv

San Francisco Bay, California is considered a mercury-impaired watershed. Elevated concentrations of mercury are found in water and sediment as well as fish and estuarine birds. Sources of mercury to the watershed since 1845 include sediment-associated mercury from mercury mining, mercury losses from gold amalgamation activities in mines of the Sierra Nevada, aerial deposition of mercury from global and regional emissions to air, and the direct discharge of mercury to Bay waters associated with the urbanization and industrialization of the estuary. We assessed historical trends in mercury bioaccumulation by measuring mercury concentrations in feathers of the endangered California Ridgway's rail (formerly California Clapper Rail) using museum specimens. We developed a structural equation model to attribute variation in historical mercury bioaccumulation in rails to sources of mercury, and estimated the toxicological consequences of extreme mercury exposure to rails from known correlations between feather and blood mercury concentrations.

Contacts

file

Historical methyl mercury in San Francisco Bay_2015.csv "Data in CSV format"		2015-07-07 08:36	wmperry@usgs.gov	5.16 KB	Process file contents
L SFBayCLRA5yrAvg_LSHg Historical methyl mercury in San Francisco Bay Metadata_09July2015.xml Original Metadata	View	2015-07-08 18:01	wmperry@usgs.gov	18.92 KB	Process file contents

Communities

- USGS Data Release Products
- USGS Western Ecological Research Center - Remove

Associated Items

% Associate an Item

Tags

Categories : Data Theme : California Ridgway's Rail, Feathers, Ice Cores, Mercury, Sediment, Structural Equation Model

Place : Lake Tahoe, San Francisco Bay, Wyoming System Types : Data Release

Provenance

Data source : Input directly

Audit History >>

Created : by wmperry@usgs.gov on Wed Jun 17 09:39:47 MDT 2015 Updated : by thorkin@usgs.gov on Fri Mar 04 15:21:25 MST 2016

Permissions *

Readable By: (INHERITED) PUBLIC

Metadata

USGS Trusted Digital Repositories

- USGS is establishing internal USGS TDRs for science data:
 - Use internationally agreed on criteria from World Data Center Data Seal of Approval-World Data System (now, Core Trust Seal)
 - Review committee approves/disapproves submissions
 - Approvals in place for 3 years recertification
 - Only an internal process at this time
 - https://www2.usgs.gov/fsp/usgs_trusted_digital_respository.asp
- Outside repositories: USGS established guidelines for use of repositories outside of USGS
 - Requires same steps for authors (review and approval, metadata with dois, connection back to SDC, etc)

		Levels of Digital Preserv	ation	
RLEMENT	LEVEL ONE	LEVEL TWO	LEVEL THREE	LEVEL FOUR
Storage and Geographic Lacation	Two names stored physically approach from each other Transfer the depaid content from temporary moles into an established storage system Managed storage system in place	Three cosise stored shysically asserate how each other A least the cary in a different page solution (off-site Regions must folder <u>high 1311 (ordering)</u>) Document the company system and storage media	Al feast one create in a gasprophic location with a different disense thread (a.g. burnane create and another and a contraction on an and	Al least three copies in geographic locations with different diselve threats Theorem a componential plan that keeps files and metadola o currently accessible systems and media
Data Integrity	Vorly checksons or logast, if provides Oratis checksons if not provides Vorla check all particles	Verify checkware on all data input Use reaching procedures when working with original mode	Verify checkwars at fixed interval of 2 years Nantach logo of diversions and supprivate thremation an domaid Mantain procedures to shock comuse other	Verify checksums of all content in response to specific events or delettes Nerrolation procedures to replace or repair comupted data Norule no see precision last write access to all opera Content, using, acti write access, of Nerrolations the all conto
Information Security	Identify who has authorization to read, write, move, and delete individual files Unit authorizations to individual files	Document access restrictions for content	Maintain logs of who performed what actions on Tios, including deletions and ampervation actions	 Perform subit of logs
Meladoka	Inventory of content and its storage location Trave backup and physical logarities of inventory information Aplices to current USOS metadate standards	 Score all meware database management information Store information describing changes to the structure or formet of the data, including time of estuments Involve access to all times of the rescalate 	Ansame canderd motivity, descriptive, and preservation metadate	Some as Level Three
File Formets	Encourage the use of a limited set of documented and open file formatis, orders, compresses achieves, and ancapeulation schemes	Inventory the file formats in use	Munitor file format obsidescence issues	Perform format migrations
Physical Media	Trentiny all physical media utilized including hand date.	Drivelop a plan to utilize trade studies to evaluate medias suitable for USOS purposes provides any from all mode utilized that are 10 years or score in Spc.	All non-recommended multis have been properly discoved of following transition activities	Base all media choises on trade studies will infermation is magnitulifrom an older media to a newer media every 3 to 5 years, including hand dialo.

Scientific Code Release Repository

 Allows code release for USGS researchers

science for a changing world

- Per Public Access Plan, software code is required to undergo review and approval under USGS FSP
- Documentation and release of code advances science reproducibility

	Projects 🗸	Groups More ~		+ ~	Search	٩	0)	n	С	•
Γ			≮ ¹ Read the]	Terms of Service						
F	Projects									
			Welcom	e to GitLab						
			Code, test, and	d deploy together						
		Create a project Projects are where you st issues, wiki and other fea	tore your code, access tures of GitLab.		Create a grou Groups are the bo members.	up est way to r	nanage	e proje	ects ar	ıd
	٢	Explore public pro There are 295 public pro Public projects are an eas to have read-only access		Learn more a Take a look at the of GitLab's capab	about Git e document bilities.	t Lab ation to	o disc	over a	I	

USGS Science Data Catalog

- Catalog of all metadata records describing data released by USGS
- Reporting capability to OMB
- Serves metadata to Department of Interior

U.S. Department of the Interior U.S. Geological Survey

data.usgs.gov

Index to official, USGS-authored publications of the USGS

Search

- Per Public Access Plan, monitors for embargo of publications
- Makes • connections to data releases (where applicable)

USGS Publications Warehouse

Study Area

CHORUS

- Clearinghouse for Open Research of the United States
- Per Public Access Plan, USGS uses CHORUS to track public access to publications

boundrivers, lateral recirculation occurs in rivers where banks exhibit strong curvature. In canyonboundrivers, lateral recirculation zones are the principal storage of fine-sediment deposits. A parallelized, three-dimensional, turbulence-resolving model was developed to study the flow structures along lateralseparation zones located in two pools along the Colorado River in Marble Canyon. The model employs thedetached eddy simulation (DES) technique, which resolves turbulence structures larger than the grid spacing in the interior of the flow. The DES-3D model is validated using Acoustic Doppler Current Profiler flowmeasurements taken during the 2008 controlled flood release from Glen Canyon Dam. A point-to-pointvalidation using a number of skill metrics, often employed in hydrological research, is proposed here forfluvial modeling. The validation results show predictive capabilities of the DES model. Th

Leveraging and Connecting USGS Tools

≊USGS

Communication Meetings Trainings

Public Access Plan: Communication

- We had 6 months from time of publication of the Public Access Plan to implement
- USGS joined together in a team effort
- Tools application leaders and science programs gathered together in March 2016 to create a 'road map' for data release with scholarly publications
- People across the Bureau who had never known each other before began to work together on solutions
- After an initial in-person meeting of ~40 people, we held monthly "accountability" meetings to ensure our work was getting done; coordinators met weekly

U.S. Department of the Interior U.S. Geological Survey

We are still working...

Communication and Training

USGS communicated <u>policy</u> and <u>implementation</u> "how-to" together

- Big push to inform Centers/Programs about both policy and data release tools
- Team approach:
 - Office of Science Quality and Integrity (OSQI) Bureau Approving Officials
 - Core Science Systems (CSS) –Science Analytics Synthesis (SAS) Science Data Management Team (SDM)

Total presentations between April and June, 2016

Total # Attendees across all presentations

ZUSGS Making Progress: Access to Data & Publications

science for a changing world

U.S. Goog

≈USGS

Prepared in cooperation with the U.S. Fish and Wildlife Service Environmental Contaminants Division

Mercury in Birds of San Francisco Bay-Delta, California-Trophic Pathways, Bioaccumulation, and Ecotoxicological Risk to Avian Reproduction

Mercury in Birds of San Francisco Bay-Delta, California: Trophic Pathways, Bioaccumulation, and Ecotoxicological **Risk to Avian Reproduction**

By Joshua T. Ackermani, Colin A. Eagles-Smith^{1,1}, Gary Heinzi, Susan De La Cruzi, John Y. Takekawai, A. Keith Milesi, Terry L. Adelsbachi, Mark P. Herzogi, Jill D. Bluso-Demens^{1,3}, Scott A. Demers^{1,3}, Garth Herringi, David J. Hoffman1, C. Alex Hartman1, James J. Wilacker1, Thomas H. Suchanek1, Steve E. Schwarzbach1, and Thomas C

Executive Summary

San Francisco Bay Estuary in northern California has a legacy of mercury or could reduce the health and reproductive success of waterbirds in the estuary. The goal of this study was to use an integrated field and laboratory approach to evaluate the risks of mercury exposure to birds in the estuary. We examined mercury bioaccumulation, and other contaminants of concern, in five waterbird species that depend heavily on San Francisco Bay Estuary for foraging and breeding habitat American avocets (Recurvirostra americana), black-necked stills (Hinantopan mexicanas), Forster's tems (Sterna forsteri), Caspian tems (Hydroprope caspia), and sarf scoters (Melanima perspicillata) These species have different foraging habitats and dises that represent three distinct foraging guids within the estuary's food web. In this report, we provide an integrated synthesis of the primary findin from this study and results are synthesized from 54 peer-reviewed publications generated to date with other unpublished results

We found that migratory waterbirds arriving in San Francisco Bay Estuary were exposed to elevated mercury concentrations, relative to other staging areas used along the Pacific Flyway. Upon arrival in the estuary, mercury concentrations in birds increased rapidly and continued to increase broughout their stay in the estuary. In particular, mercury concentrations in Forster's terms nearly trieled during the short. 2-month period from the time they arrived in the estuary to when they began reeding. Similarly, mercury concentrations in over-wintering surf scoters more than tripled prior to their departure to their northern breeding grounds. Our captive rearing studies on mallards (Anar platy/kywokos) showed that mercury obtained from exposure to a new diot source was transferred to eees in less than a week. Thus, micratory birds artiving to breed in San Francisco Bay Estuary are sposed to elevated mercury concentrations in their environment precisely at the critical time period for breeding.

¹U.S. Geological Survey. ²U.S. Fish and Wildlife Service, Environmental Contaminants Division lumboldt State University, Wildlife Department

USGS Public Access Plan allows USGS to continue to move forward...

We strive to:

- implement FAIR: Data is findable, accessible, interoperable, and reusable for future research;
- Offer more credit for data release in RGE process
- Formalize data management positions
- Increase data citation rates ٠
- Resolve large data release ٠ and preservation challenges

Science for a changing world Making Progress: Data and Publications

Over 1,651,718 visitors to USGS Publications Warehouse in FY18

Over 722,332 publication downloads

Over 1,500 USGS data releases are now publicly available in ScienceBase, representing every USGS Mission Area, and 82 different USGS Science Centers and Programs.

Thank you!

Questions? Comments?

Viv Hutchison USGS

vhutchison@usgs.gov

