



# Greater Yellowstone Network Amphibian Monitoring Protocol Science Review

## *A Summary of Reviewers' Responses*

Natural Resource Report NPS/GRYN/NRR—2022/2389



ON THE COVER

Clockwise from the top left: western tiger salamander (*Ambystoma mavortium*), boreal chorus frog (*Pseudacris maculata*), Columbia spotted frog (*Rana luteiventris*), and western toad (*Anaxyrus boreas*)

Photography by NPS

---

# Greater Yellowstone Network Amphibian Monitoring Protocol Science Review

## *A Summary of Reviewers' Responses*

Natural Resource Report NPS/GRYN/NRR—2022/2389

Michael Bozek<sup>1</sup> and Tani Hubbard<sup>2</sup>

<sup>1</sup>National Park Service  
Inventory and Monitoring Division  
Lakewood, Colorado

<sup>2</sup>Northern Rockies Conservation Cooperative and  
National Park Service/Greater Yellowstone Inventory and Monitoring Network  
Tucson, AZ

May 2022

U.S. Department of the Interior  
National Park Service  
Natural Resource Stewardship and Science  
Fort Collins, Colorado

The National Park Service, Natural Resource Stewardship and Science office in Fort Collins, Colorado, publishes a range of reports that address natural resource topics. These reports are of interest and applicability to a broad audience in the National Park Service and others in natural resource management, including scientists, conservation and environmental constituencies, and the public.

The Natural Resource Report Series is used to disseminate comprehensive information and analysis about natural resources and related topics concerning lands managed by the National Park Service. The series supports the advancement of science, informed decision-making, and the achievement of the National Park Service mission. The series also provides a forum for presenting more lengthy results that may not be accepted by publications with page limitations.

All manuscripts in the series receive the appropriate level of peer review to ensure that the information is scientifically credible and technically accurate.

Views, statements, findings, conclusions, recommendations, and data in this report do not necessarily reflect views and policies of the National Park Service, U.S. Department of the Interior. Mention of trade names or commercial products does not constitute endorsement or recommendation for use by the U.S. Government.

This report is available from the [Greater Yellowstone Inventory and Monitoring Network website](#), and the [Natural Resource Publications Management website](#). If you have difficulty accessing information in this publication, particularly if using assistive technology, please email [irma@nps.gov](mailto:irma@nps.gov).

Please cite this publication as:

Bozek, M., and T. Hubbard. 2022. Greater Yellowstone Network amphibian monitoring protocol science review: A summary of reviewers' responses. Natural Resource Report NPS/GRYN/NRR—2022/2389. National Park Service, Fort Collins, Colorado. <https://doi.org/10.36967/nrr-2293614>.

# Contents

Page

|                                                                                                                                                                                                                                                                                                             |      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| Tables . . . . .                                                                                                                                                                                                                                                                                            | v    |
| Executive Summary . . . . .                                                                                                                                                                                                                                                                                 | vi   |
| Acknowledgments . . . . .                                                                                                                                                                                                                                                                                   | vii  |
| Audience . . . . .                                                                                                                                                                                                                                                                                          | vii  |
| Terms and Acronyms . . . . .                                                                                                                                                                                                                                                                                | viii |
| Background . . . . .                                                                                                                                                                                                                                                                                        | 1    |
| Amphibian Monitoring . . . . .                                                                                                                                                                                                                                                                              | 1    |
| Protocol Review . . . . .                                                                                                                                                                                                                                                                                   | 1    |
| Reviewers . . . . .                                                                                                                                                                                                                                                                                         | 4    |
| Summarized Review Responses . . . . .                                                                                                                                                                                                                                                                       | 5    |
| Scientific Integrity and Rigor . . . . .                                                                                                                                                                                                                                                                    | 5    |
| Question 1: Is GRYN meeting objectives stated in the current amphibian protocol? If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these. . . . .             | 5    |
| Question 3: Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program? To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of out of their sampling visits? . . . . . | 8    |
| Question 4: Are current analytical approaches described in the protocol and in Gould et al. (2019) capable of detecting statistically significant trends for all species involved? For detecting principal threats to amphibians? For understanding the drivers of wetland change? . . . . .                | 9    |
| Collaboration, Science Impact, and Integrating Science into Parks . . . . .                                                                                                                                                                                                                                 | 11   |
| Question 2: Is GRYN monitoring at the appropriate/most ecologically relevant scale? Are there ways to expand outside Grand Teton and Yellowstone national parks to the GYE? To the Western U.S.? . . . . .                                                                                                  | 11   |
| Question 5: How could GRYN increase the impact of their work? How can they work with parks and partners to transform knowledge into conservation action? . . . . .                                                                                                                                          | 12   |
| Sustainability . . . . .                                                                                                                                                                                                                                                                                    | 13   |
| Question 6: How would you prioritize monitoring activities if budget/time/safety/personnel constraints restricted GRYN's ability to complete current activities? . . . . .                                                                                                                                  | 13   |
| Conclusions . . . . .                                                                                                                                                                                                                                                                                       | 15   |
| Literature Cited . . . . .                                                                                                                                                                                                                                                                                  | 16   |
| Appendix A. Questions and Instructions to Reviewers . . . . .                                                                                                                                                                                                                                               | 17   |
| Appendix B. Full Reviewer Responses to Questions . . . . .                                                                                                                                                                                                                                                  | 19   |
| Jane Carlson . . . . .                                                                                                                                                                                                                                                                                      | 20   |
| Michelle Christman . . . . .                                                                                                                                                                                                                                                                                | 24   |

# Contents (continued)

|                             | Page |
|-----------------------------|------|
| Brian J. Halstead . . . . . | 28   |
| Paige Howell . . . . .      | 41   |
| Dusty Perkins . . . . .     | 45   |
| Stephen Spear . . . . .     | 49   |
| Mary Kay Watry . . . . .    | 58   |

# Tables

Page

|                                                                                                        |   |
|--------------------------------------------------------------------------------------------------------|---|
| <b>Table 1.</b> Objectives from the Greater Yellowstone Network amphibian monitoring protocol. . . . . | 2 |
|--------------------------------------------------------------------------------------------------------|---|

# Executive Summary

Science reviews are an essential cornerstone of all excellent science programs and are a requirement of monitoring programs within the Inventory and Monitoring Division of the National Park Service (NPS). Science reviews provide necessary professional critique of objectives, study design, data collection, analysis, scientific interpretation, and how effectively information is transferred to target audiences. Additionally, reviews can help identify opportunities to cooperate more effectively with interested and vested partners to expand the impacts of collective findings across larger landscapes.

In December 2020, seven biologists from USGS, USFWS, and NPS provided a critical review of the Greater Yellowstone Network Amphibian Monitoring Protocol for monitoring Columbia spotted frogs (*Rana luteiventris*), boreal chorus frogs (*Pseudacris maculata*), western toads (*Anaxyrus boreas*), western tiger salamanders (*Ambystoma mavortium*), and environmental conditions at wetland sites clustered within watershed units in Yellowstone and Grand Teton national parks. This review followed sixteen years of GRYN amphibian and wetland monitoring, allowing us to evaluate the impact of the work thus far and to discuss potential improvements to the protocol.

Reviewers were asked to assess the following amphibian monitoring objectives per Bennetts et al. (2013, Cooperative amphibian monitoring protocol for the Greater Yellowstone Network: Narrative, version 1.0, <https://irma.nps.gov/DataStore/Reference/Profile/2194571>) and to assess the degree to which GRYN is meeting the objectives based on the current sampling, analyses, and reporting:

- Objective 1: Estimate the proportion of catchments and wetland sites used for breeding by each of the four common, native amphibian species annually, and estimate the rate at which their use is changing over time.
- Objective 2: Determine the total number of wetlands within sampled catchments that are suitable for amphibian breeding (i.e., have standing water during the breeding season) annually.
- Objective 3: For western toads, estimate the proportion of previously identified breeding areas that are used annually, and estimate the rate at which their use may be changing over time.

Generally, reviewers commended the GRYN Amphibian Monitoring Program, including the design, the statistical rigor of current analytical approaches, the large number of monitoring reports and publications, and the audiences reached. Reviewers unanimously felt that the first two objectives of this protocol are being met for two species (Columbia spotted frogs and boreal chorus frogs) in medium- and high-quality catchments, and all but one reviewer also felt these objectives are being met for western tiger salamanders. It was universally recognized that objective 3 for western toads is not being met, and reviewers attributed this to its rarity and life history differences compared to the other species rather than inherent design flaws. To address objective 3, reviewers suggested additional funding and capacity. Reviewers felt the current design provides an adequate base for parlaying additional work and offered suggestions focused on increasing efficiencies, maximizing information that can be collected in the field, strengthening analyses, and improving scientific outreach. In this document, we summarize reviewers' comments and include their full written reviews in Appendix B.

# Acknowledgments

We thank the GRYN staff for their work in hosting this workshop during the challenging Covid-19 pandemic, which helped buoy the spirits of participating scientists and staff. In particular, we thank Moses Chun, who designed the meeting layout and worked to keep everything running smoothly during the meeting, Jana Cram and Rob Daley who assisted Moses, Ben LaFrance for taking notes and filling in where necessary, and Kristin Legg and Andy Ray for organizing and hosting the workshop and their extensive review of this summary of reviewers' comments. We thank Blake Hossack for his assistance identifying reviewers and support throughout the review process. Very importantly, a special thank you to the protocol reviewers who volunteered their time and expertise and all the participants and presenters who made this review a success.

## Audience

The primary audiences for this report are current and future Greater Yellowstone Network and national park scientists and program managers and other scientists working on amphibian monitoring. Reviewers' comments and suggestions are summarized in this report. This document does not provide a response to these comments or a description of any changes that will be made to the protocol. We will determine what changes will be made by the end of 2022 and release a new version of the protocol based on this review, but at this time we anticipate limited changes will be necessary to address concerns.

# Terms and Acronyms

|                      |                                                                                                    |
|----------------------|----------------------------------------------------------------------------------------------------|
| ARMI:                | Amphibian Research and Monitoring Initiative; a USGS program                                       |
| <i>Bd</i> (chytrid): | A fungus called <i>Batrachochytrium dendrobatidis</i> that kills amphibians, also known as chytrid |
| Catchments:          | Relatively small (~200 hectare) surface water drainage areas                                       |
| eDNA:                | Environmental DNA or DNA that is released from an organism into the environment                    |
| GRTE:                | Grand Teton National Park                                                                          |
| GRYN:                | Greater Yellowstone Inventory and Monitoring Network                                               |
| GYE:                 | Greater Yellowstone Ecosystem                                                                      |
| Inventory 2.0:       | The second generation of NPS inventories of natural resources in parks                             |
| NPS:                 | National Park Service                                                                              |
| NCPN:                | Northern Colorado Plateau Network                                                                  |
| NWI:                 | National Wetland Inventory                                                                         |
| PARC:                | Partners in Amphibian and Reptile Conservation                                                     |
| Ranavirus            | A viral disease in amphibians                                                                      |
| RMAP:                | Rocky Mountain Amphibian Project                                                                   |
| SDM:                 | Structured decision making                                                                         |
| TTD:                 | Time-to-detection surveys                                                                          |
| USGS:                | U.S. Geological Survey                                                                             |
| VES:                 | Visual encounter surveys                                                                           |
| Wetland Sites:       | Specific area within a catchment where the land is covered in water or saturated with water        |
| YELL:                | Yellowstone National Park                                                                          |

# Background

## Amphibian Monitoring

The National Park Service Vital Signs Monitoring Program (Fancy et al. 2009) was designed to assess the status and trends of key park resources. In Yellowstone and Grand Teton national parks (YELL and GRTE, respectively), amphibians were identified as one of a select group of vital signs (Jean et al. 2005) to be monitored by the Greater Yellowstone Network (GRYN). Using monitoring techniques previously developed by Idaho State University and the U.S. Geological Survey's Amphibian Research and Monitoring Initiative (ARMI), GRYN piloted monitoring of native amphibians in 2005. The sampling design and draft protocol were prepared in 2006 and finalized in 2013 (Bennetts et al. 2013).

The overall goal of the program was to detect long-term trends for four widespread native amphibian species and document patterns of wetland flooding and drying (the dominant breeding habitat for amphibians) across the landscape. The four amphibian species are Columbia spotted frogs (*Rana luteiventris*), boreal chorus frogs (*Pseudacris maculata*), western toads (*Anaxyrus boreas*), and western tiger salamanders (*Ambystoma mavortium*).

Amphibian habitat quality was assigned to one of three categories based on the areal extent, within each catchment, of the following water regime classification values from the 1980s National Wetlands Inventory (Cowardin et al. 1979).

- Seasonally Flooded (Water Regime Code 'C')
- Semi-permanently Flooded (Water Regime Code 'F')
- Intermittently Exposed (Water Regime Code 'G')
- Permanently Flooded (Water Regime Code 'H')

Catchments considered high-quality habitat contain four or more hectares of NWI-classified Semi-permanently Flooded, Intermittently Exposed, or Permanently Flooded water regime features and two or more hectares of Seasonally Flooded water regime features.

Catchments not meeting the criteria for high-quality habitat are considered medium-quality habitat if they

contain one or more hectares of Seasonally Flooded water regime features and any features, regardless of size, classified in NWI as Semi-permanently Flooded, Intermittently Exposed, or Permanently Flooded water regimes.

Catchments not meeting the criteria for high- or medium-quality habitat are considered low-quality habitat if they contain any features, regardless of size, classified in NWI as Seasonally Flooded, Semi-permanently Flooded, Intermittently Exposed, or Permanently Flooded.

The ongoing, combined wetlands and amphibian monitoring in YELL and GRTE represents one of the longest running (16 years at the time of review) multi-species amphibian monitoring programs in the U.S. When combined with partner-led monitoring and restoration efforts at the U.S. Forest Service Blackrock Ranger Station Research Site, Yellowstone's Lodge Creek, the National Elk Refuge in Jackson Hole, Wyoming, and the University of Wyoming's Rocky Mountain Amphibian Project, knowledge on amphibians in the Greater Yellowstone Ecosystem (GYE) is considerable. However, little effort has been made to formally coordinate ongoing regional efforts, rigorously evaluate their individual and collective impact, and discuss to what extent these efforts are meeting the needs of land managers.

## Protocol Review

Regular reviews are important for maintaining enthusiasm and integrating best available science from long-term ecological monitoring with management and is also a requirement of the NPS's Vital Signs Monitoring Program (Mitchell et al. 2018). This review of the GRYN amphibian monitoring protocol was designed as a big-picture assessment of the study design, methods, data analysis and interpretation, transfer of information, and insights gained from the monitoring program. To date, the information generated from this program has been regularly used in GRTE and YELL vital signs reporting, annually incorporated in Yellowstone's Interpretation Handbook, and reviewed as needed by the U.S. Fish and Wildlife Service for federal listing considerations (e.g., the Boreal Toad review in 2017).

**Table 1.** Objectives from the Greater Yellowstone Network amphibian monitoring protocol (Bennetts et al. 2013; <https://irma.nps.gov/DataStore/Reference/Profile/2194571>).

| Objective   | Description                                                                                                                                                                                              |
|-------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Objective 1 | Estimate the proportion of catchments and wetland sites used for breeding by each of the four common, native amphibian species annually, and estimate the rate at which their use is changing over time. |
| Objective 2 | Determine the total number of wetlands within sampled catchments that are suitable for amphibian breeding (i.e., have standing water during the breeding season) annually.                               |
| Objective 3 | For Western Toads, estimate the proportion of previously identified breeding areas that are used annually, and estimate the rate at which their use may be changing over time.                           |

This science review critically examined the relevance of the work and the progress made towards each objective listed in the original protocol (Table 1). Furthermore, the review assessed constraints on the current program and recommended ways to improve amphibian monitoring effectiveness. The reviewers were asked to address the following three dimensions of the current program:

- **Scientific Integrity:** Provide a comprehensive science critique of GRYN’s current amphibian and wetland monitoring program, including a review of the strengths and weaknesses of the program and suggestions for alternative approaches where appropriate.
- **Effective Science Information Transfer and Use:** Carefully review how the current work and resulting reporting are meeting information needs of parks and recommend strategies that best meet the needs of parks and other resource management partners.
- **Collaboration to Increase Impact:** Explore opportunities to engage in collaborative work that can increase the impact of this amphibian monitoring program both within a park (intra-ecological impact) and across the Greater Yellowstone Ecosystem (GYE; inter-ecological impact).

Reviewers completed the following activities prior to writing the reviews that are summarized in this report:

- Reviewers first assessed the current version of the protocol and methods therein, a subset of examples of published reports and peer-reviewed papers emanating from this monitoring work by GRYN and cooperators, and a previous comprehensive analysis of the monitoring results by Gould et al. (2019).

- The reviewers next participated in an Amphibian Science Symposium and Protocol Review from December 8–10, 2020.
  - The symposium provided a comprehensive view of the program with speakers presenting the study design, methods, results, and uses of the monitoring data. Also included in the symposium were presentations on the biology, conservation, natural history, and threats to amphibians in the GYE that could inform improvements to the current monitoring program. Presentations were interspersed with topical breakout groups that allowed reviewers to ask each speaker questions and interact with GRYN and their collaborators as well as other presenters.
  - After the symposium, reviewers met with GRYN staff and key collaborators for discussions and a question-and-answer session with the intent of clarifying questions and addressing additional concerns before preparing their final assessments of the protocol and recommendations for improvement. Reviewers were also allowed to communicate with GRYN staff or other reviewers during the preparation of their individual assessments for additional clarification, if needed.

Reviewers then individually responded to the following questions (further instructions to reviewers are in Appendix A and reviewer responses are provided in Appendix B).

1. Is GRYN meeting objectives (Table 1) stated in the current amphibian protocol? If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these.

2. Is GRYN monitoring at the appropriate/most ecologically relevant scale? Are there ways to expand outside Grand Teton and Yellowstone national parks to the GYE? To the Western U.S.?
3. Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program? To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of their sampling visits?
4. Are current analytical approaches described in the protocol and in Gould et al. (2019) capable of detecting statistically significant trends for all species involved? For detecting principal threats to amphibians? For understanding the drivers of wetland change?
5. How could GRYN increase the impact of their work? How can they work with parks and partners to transform knowledge into conservation action?
6. How would you prioritize monitoring activities if budget/time/safety/personnel constraints restricted GRYN's ability to complete current activities?

# Reviewers

The following people reviewed the GRYN amphibian monitoring program. In the next chapter, we summarize the reviews using reviewers' initials.

**JC: Jane Carlson, Ph.D.** (*Gulf Coast Network, Inventory and Monitoring Division, NPS, Lafayette, LA*)

Dr. Carlson is engaged with long-term monitoring of amphibians in three national parks. She has a theoretical background in population ecology, plant functional morphology, pollination biology, entomology and invertebrate ecology, natural and sexual selection, and multivariate statistical analyses.

**MC: Michelle Christman, M.S.** (*Natural Resource Program Center, USFWS, Fort Collins, CO*)

Michelle Christman is a biologist with the U.S. Fish & Wildlife Service and currently serves as the Federal Coordinator for Partners in Amphibian and Reptile Conservation. She works on the conservation, recovery, and management of amphibians and reptiles and is particularly interested in collaborative partnerships for greater conservation impact.

**BH: Brian Halstead, Ph.D.** (*Western Ecology Research Center, USGS, Dixon, CA*)

Dr. Brian Halstead's research focuses on applications of population ecology on the conservation of reptiles and amphibians. Specifically, he combines field methods like capture-mark-recapture and radio telemetry with flexible Bayesian hierarchical models to learn about demographic parameters and the individual, biotic, and environmental variables that affect them. Central themes of his research include the influence of land use, climate, and biotic interactions on species distributions and demographic rates; the behavioral response of individuals to habitat change; and providing relevant information about wildlife conservation to resource managers and private landowners.

**PH: Paige Howell, Ph.D.** (*National Raptor Program, Division of Migratory Bird Management, USFWS, Laurel, MD*)

Dr. Howell is an applied population ecologist who has worked with amphibians, birds, small mammals, and fish. She is interested in the fundamental question of how and why populations change over

space and time. A central theme in Dr. Howell's research is developing statistical models and monitoring protocols to provide scientific support for wildlife conservation and management.

**DP: Dusty Perkins, Ph.D.** (*Northern Colorado Plateau Network, Inventory and Monitoring Division, NPS, Crested Butte, CO*)

Dr. Perkins serves as the Program Manager for the NCPN. He is responsible for the planning, management, implementation, and coordination of all aspects of the program and serves as protocol lead for big rivers, invasive exotic plants, and landbird monitoring protocols. Dr. Perkins' dissertation work focused on applied conservation of amphibians in landscapes with active riparian management in western Maine.

**SS: Stephen Spear, Ph.D.** (*Upper Midwest Environmental Sciences Center, USGS, LaCrosse, WI*)

Dr. Spear is a Research Biologist with the USGS Upper Midwest Environmental Sciences Center. He has worked on multiple aspects of conservation biology including conservation genetics, environmental DNA, endangered species head-starting and reintroduction, and restoration of wildlife habitat, often in coordination with federal, state, and NGO partners, and primarily with herpetofauna. A major focus of his research has been to use molecular tools to inform management.

**MW: Mary Kay Watry, M.S.**  
(*Rocky Mountain National Park, NPS, Estes Park, CO*)

Mary Kay Watry currently serves as a conservation biologist for Rocky Mountain National Park and has focused on the stabilization and recovery of species for the last 20 years. She enjoys collaborating with teams and across boundaries to explore conservation actions. Current efforts with the Boreal Toad Conservation Team evaluate potential future management strategies including reintroductions that benefit boreal toad recovery.

# Summarized Review Responses

For each question posed to reviewers, highlights of their responses occur in this section in generalized bulleted form under a respective review topic.

In many instances, reviewer responses were quite detailed. We recommend readers examine the highlights here, and then read the complete responses from each reviewer in Appendix B. Recommendations are often presented word for word from reviewers' comments and not put in quotes in this summary.

## Scientific Integrity and Rigor

***Question 1: Is GRYN meeting objectives stated in the current amphibian protocol? If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these.***

### Objective 1

"Estimate the proportion of catchments and wetland sites used for breeding by each of the four common, native amphibian species annually, and estimate the rate at which their use is changing over time."

The four native amphibian species monitored are Columbia spotted frogs (*Rana luteiventris*), boreal chorus frogs (*Pseudacris maculata*), western toads (*Anaxyrus boreas*), and western tiger salamanders (*Ambystoma mavortium*).

- Overall, reviewers felt that the first objective was being met for two species: Columbia spotted frogs and boreal chorus frogs.
- Five reviewers (JC, BH, PH, DP, MW) felt that this objective was being met for western tiger salamanders. SS made arguments that the objective was not being met for this species and suggested increasing the number of catchments sampled for western tiger salamanders even if the design became unbalanced.
- DP, SS, and MW did not feel Objective 1 was met for western toads because there was not enough information available to model the species and make that determination.
- MW felt that the Objective 1 is only being met in medium- and high-quality catchments for three species.

- SS questioned if the current sampling design adequately represented medium- and high-quality sites given that the current sampling only represents 11% of high-quality catchments and just 2% of medium-quality catchments in YELL, although others did not express that concern.
- There were varied comments regarding GRYN's decision to suspend surveys in low-quality catchments after 2010:
  - JC and DP pointed out that excluding low-quality catchments is not a problem because GRYN is still sampling low-quality wetland sites. JC said the spatial coverage of the wetland area is pretty good without low-quality catchments.
  - BH suggested the elimination of low-quality sites affects inference but doesn't significantly impact the overall protocol objectives.
  - MW objected to the elimination of low-quality catchments and said that Objective 1 is not being met for any species in low-quality catchments. MW said that it may take longer to observe a detectable change in medium- and high-quality sites as these remaining catchments will take longer to degrade to a lower quality condition before becoming unsuitable for breeding.
  - SS and MW said inference to the majority of catchments in the park and the broader study area is not possible without low-quality catchments.

### Objective 2

"Determine the total number of wetlands within sampled catchments that are suitable for amphibian breeding (i.e., have standing water during the breeding season) annually."

- Six reviewers (JC, BH, PH, DP, SS, and MW) felt that Objective 2 was being met and one reviewer did not directly comment on this; some felt coverage of this objective was concomitant with work done on Objective 1, and many of the comments and modelling suggestions apply across both Objectives.

- MW commented that Objective 2 is also only met for medium- and high-quality catchments because low-quality catchments are not monitored.

### **Objective 3**

"For western toads, estimate the proportion of previously identified breeding areas that are used annually, and estimate the rate at which their use may be changing over time."

- All reviewers, except one that did not address this in their response, thought Objective 3 was not being met.
  - BH, DP, and SS pointed out that funding and time limitations were the reason the objective was not met.
  - JC suggested dropping this objective from the protocol and DP suggested discussing that option with the parks. DP thought that this objective could be an Inventory 2.0 project and recommended leaving Objective 3 in the protocol but that GRYN indicate in the protocol that it will remain unmet unless funding increases. SS felt that dropping western toads was not ideal.
  - SS cautioned that Objective 3 may not be met even if more resources were available because of the difference in habitat use by western toads and the difficulty in determining if a site is suitable and suggested this species may need a toad-specific protocol.
  - BH commented that Objective 3 does not easily overlap with the timing of current surveys because of earlier breeding and shorter larval development periods for western toads.

### **Suggested monitoring improvements to better meet objectives**

Universally, reviewers complimented the protocol and the multi-state occupancy modelling approach to analyze amphibian breeding trends. It was clear to reviewers that despite their many suggestions for potential improvements and expansions, there are sampling limitations related to both resources and safety (backcountry work and animal encounters) in GRTE and YELL.

### ***Sampling design***

- Spatial autocorrelation issues in the design were brought up by BH, SS, and DP:
  - BH indicated it could be handled in the analysis (see Question 4 below).
  - SS suggested testing whether wetland sites were independent of catchments and questioned how the long-term experimental unit was defined. This question stemmed from the Gould et al. (2019) analysis where neither catchment nor drainage was a factor in the multi-state models. SS suggested that it was critical to know how catchment factors into the analyses for assessing breeding dynamics.
  - DP expressed concern about independence of sites and wasn't sure about the value of site-level analysis since the monitoring was set up to evaluate catchments, though DP defers to Gould on this. However, JC recommended dropping the catchment concept except when explaining initial study design because site-level analysis is the appropriate focus.
- Elimination of catchments having low-quality habitat:
  - MW recommended low-quality catchments be added back in because they represent most catchments in these parks. MC suggested this could be done by replacing medium- and high-quality *far* and *difficult to access* catchments with medium- and high-quality *near* catchments to increase logistical sampling efficiency, monitoring medium- and high-quality catchments on a rotating basis, or reducing sites such that the number of low-, medium-, and high-quality sites are proportional to their representation across the landscape.
  - However, JC, BH, and DP felt that the elimination of low-quality catchments was not impacting objectives. Additionally, BH felt that visiting the same sites every year was better because it produces more precise estimates of trends.
- SS proposed adding catchments near current catchments to increase sample size, which would not substantially increase travel time. SS also commented that reducing the number of

wetlands sampled per catchment (if it didn't affect power for most common species and if catchments are independent) could save time and would be better than reducing catchments because catchment sample size is already low and should be preserved in the design. However, SS pointed out that sampling each wetland in a catchment is appropriate for Objective 2. BH said the multi-scale approach with random selection of catchments and all sites visited in the catchment is an efficient design.

- Several reviewers (JC, BH, SS, and MW) commented that remote sensing data could improve and enhance objectives.
  - MW stated that this would be a way to include some data from low-quality catchments.
  - JC suggested that remote sensing could also help quantify spatial proximity to other wetlands (i.e., nearest neighbor analysis), better measure wetland size, and widen the sampling frame.
  - BH pointed out that remote sensing also provides hydroperiod information (i.e., serial assessment of wetland size), which can provide more information on climatic and other drivers of wetland dynamics than the current approach of using "at-the-time-of sampling" data can.
  - SS recommended expanding Objective 2 with remote sensing. If wetland state is a primary driver of amphibian trends, this expansion could allow parks to predict trends across the whole park.
- DP recommended that Objective 2 be modified to include a long-term component since GRYN is already doing this and it should be reflected in the objective. Specifically, DP thought that looking at wetland desiccation over longer-time periods (and prior to GRYN's first field observations) could be just as valuable as the amphibian presence data.
- Western toad and tiger salamander sampling suggestions:
  - PH, SS, and MW suggested targeted sampling of catchments/regions/habitats for one or both rarer species (western toad, tiger

salamander) to increase sample size (and cover the range of important covariates) even if this violates randomness of the design.

- DP indicated that toad inventories could happen in a single, good condition year and felt that adding catchments to increase tiger salamander sampling may be more beneficial than targeting toad locations, which would require a substantial increase in effort and an alternate/additional design.
  - BH proposed that sampling for western toads should occur at a different time to align specifically with their breeding.
  - MW commented that rare species can be better indicators of wetlands reaching trigger points, so including western toads is a good way to understand how amphibian breeding is changing over time. MW recommended targeting catchments with historical western toad presence, shifting sampling period to earlier or multiple site visits, a rotating panel to have more time for western toad sampling, and using environmental DNA (eDNA) to discover if breeding is being missed. In contrast, PH thought that eDNA is not worth it in this situation. Further description of reviewers' comments on eDNA can be found under Question 3 below.
  - MW also pointed out that western toads have been shown to breed sporadically (they don't breed every year) making it more difficult to confirm active breeding sites and making this behavior contrary to a single inventory style sampling approach. Therefore, MW recommended a modified methodology that is spatially balanced and uses time blocks of 3–5 years, rather than annual time blocks, to assess whether a breeding site is active.
  - SS said dropping tiger salamanders and western toads is not ideal because they are ecologically different and trends in boreal chorus frogs and Columbia spotted frogs may not apply to them.
- PH offered time-to-detection (TTD) surveys to improve the protocol. TTD is described in more detail under Question 3 along with other field work modification recommendations.

- MC asked whether the objectives relate to reference points and how meeting the objectives can be integrated into management decisions.

#### *Analysis*

- Reviewers made suggestions for improving analyses to better meet the protocol objectives; their comments are included under Question 4 below.

#### *Science collaboration*

- MW said that Dr. Larissa Bailey is developing a monitoring strategy for western toads for the Southern Rocky Mountains as part of the Boreal Toad Conservation Plan and GRYN might consider adopting it once it is complete.
- PH suggested community science efforts to help add sites and collect more data.

### ***Question 3: Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program? To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of out of their sampling visits?***

#### Current field methods

- Reviewers agreed that the current Visual Encounter Surveys (VES) methods (supplemented with dipnets) are appropriate and among the best available methods for this protocol. Suggestions below for other methods are in addition to VES.
  - BH pointed out that high detection probabilities reinforce the effectiveness of the method chosen for three species (but not for western toads).
  - BH felt that focusing on amphibians and any interacting species (e.g., competitors, predators, prey, and disease vectors) that might affect amphibian occurrence, rather than all species, made sense for efficient field monitoring.

#### Suggested field method improvements or additions

##### *Time-to-detection (TTD)*

- BH, PH, DP, and SS suggested that TTD should be added to the current VES (not as a replacement) and would add very little time to fieldwork. BH said the information provided on the detection process is likely still worth the minimal cost of collecting TTD data. PH said TTD shouldn't affect the cost/time of surveys but could be integrated into existing occupancy models to potentially yield better results than the current surveys alone.
- BH, PH, DP, and SS suggested not using single observer techniques when adding TTD because safety constraints do not allow single observer methods to be employed in place of VES.
- JC recommended using GPS units in TTD surveys to acquire the additional information more easily if GRYN chooses to try out TTD.

##### *Environmental DNA (eDNA)*

- BH, DP, SS, and MW suggested eDNA as *additional* work or as a better way to detect tiger salamanders and western toads. It could be used for quick sampling of a site and would be useful for detecting the four amphibians and interacting species (covariates that affect amphibian occupancy), for biodiversity assessments, and for detecting disease. SS mentioned that adding eDNA to the methods could provide an archived sample and facilitate relative abundance estimations, though the latter is not feasible at this time with current technologies. DP recommended putting samples in storage even if GRYN does eDNA analysis now because of the long-term benefit to science.
- SS said eDNA may be most beneficial for achieving objectives for tiger salamanders and western toads. A rapid eDNA survey targeted at park locations where suitable habitat exists for each species would provide an efficient method for identifying new sites if funding allowed. In addition to amphibians, SS said eDNA sampling for disease (or swab tissue sampling) and aquatic invertebrates (or dipnetting) may be valuable if funding allowed, though the latter would primarily reflect water quality rather than a direct interaction with amphibians.
- BH, DP, SS, and MW acknowledged there are issues regarding sampling, storage, and analysis of eDNA samples and interpretation of data (e.g., DNA degradation, identifying reproducing versus migrating populations). MW said eDNA can let you know if breeding was potentially missed in site even if it can't tell you if breeding occurred.

- JC and PH felt that it is not worth adding eDNA because maintaining continuity with past field methods is more important and there are funding and other constraints.
- BH stated that eDNA is not currently a cost-effective alternative to VES for these species, but to consider it in the future as technology advances.

#### *Acoustic loggers*

- Reviewers noted the use of acoustic loggers (BH and DP) in amphibian monitoring but were reluctant to recommend them because they can't detect non-vocalizing species and more visits would be required with this technology.

#### *Disease sampling*

- SS, BH, DP, and MW discussed disease sampling for chytrid (Bd) and ranavirus. SS and MW said disease might be the most important beneficial component missing in this work. SS and MW suggested eDNA or swab tissue sampling for disease, particularly for ranavirus since it is a serious threat in GRYN parks.

#### *Community science*

- SS, DP, and PH recommended community-sourced (i.e., citizen science) data to increase sampling, add sites, or implement additional methodologies.

#### *Wetland biodiversity*

- SS, DP, and BH suggested that sampling additional taxonomic groups to better characterize the aquatic community and other terrestrial species likely to interact with amphibians would be useful (e.g., fish, invertebrates). JC cautioned that the focus of such an effort should be on species that are highly interactive with the amphibian(s) present (e.g., mammals, birds) and did not see a need to add more fieldwork for biodiversity unless desired by parks and executed outside the purview of this protocol.
- MW suggested recording rare and federal- or state-listed plants in the wetlands and DP suggested that adding presence of the top five most dangerous invasive riparian plants (and invasive aquatic animals) would be easy to collect.

#### *Metapopulation dynamics*

- SS expressed significant interest in population genetic sampling (swab tissue) to understand metapopulation dynamics in these systems. SS said that this could provide important complementary data to estimates of occupancy and the impact of drought and climate change on populations but acknowledged that it is expensive and offered ideas for ways to reduce cost and effort. Read more on genetic sampling in the reviewers' comments in Appendix B.
- MC referred again to low-quality catchments saying they are important for metapopulations. It should be noted that MC refers to these catchments as "low-quality sites" and other reviewers pointed out that there are low-quality sites in medium- and high-quality catchments.

#### *Other suggestions*

- DP mentioned tightening up photo points that were used in presentations, which are powerful storytellers; it was unclear at times that the same sites were being presented in repeat comparison photographs.
- DP also suggested that GRYN not add too much additional data collection because unpredictable circumstances occur and it's important to have continuity in the core protocol data set for meeting long-term monitoring objectives.

#### ***Question 4: Are current analytical approaches described in the protocol and in Gould et al. (2019) capable of detecting statistically significant trends for all species involved? For detecting principal threats to amphibians? For understanding the drivers of wetland change?***

Reviewers made comments on analytical approaches in responses to multiple review questions and all responses are summarized here. See Sampling Design under Question 1 for related suggestions.

#### **Trends**

##### *Current analytical methods*

- BH, PH, and MW complimented GRYN for expanding methods beyond those described in the protocol and using multi-state occupancy models to meet Objectives 1 and 2. PH pointed out that western toad is the exception for reasons already outlined.

- JC, BH, PH, SS, and MW thought that the current methods were adequate for the common species, but inadequate for the rarer species. PH said the current analytical approaches can detect statistically significant trends for most species (not western toads), threats, and drivers of wetland change.
- SS asserted that survey limitations rather than analytical framework are the primary reason current methods are limited for tiger salamanders and western toads. SS commented that the current analytical framework is well suited to estimate trends in amphibians accounting for natural cycles of wetland drying and sets the stage to understand future climate change and amphibians. It provides enough power to model occupancy and trends for the two more common species. BH agreed with SS that the ability to detect trends is less a function of the analytical approach and more a function of the study design, which impacts the rarer species in this protocol.
- BH and MW felt current methods are suitable for making inferences across medium- and high-quality catchments. However, SS felt that trends may not represent all medium- and high-quality catchments due to sample size but acknowledged that there are funding issues.
- SS pointed out that current analyses do not incorporate the catchment and drainage (i.e., subbasin) levels in the analysis despite the monitoring being stratified by these levels.

#### *Recommendations*

The following are suggested analytical approaches in addition to the current multi-state modelling (for details beyond this summary, see reviewers' comments in Appendix B).

- JC, BH, and PH proposed exploring multi-species modeling approaches that share detectability estimates across species, which would be helpful in modeling rarer species. BH and PH felt a Bayesian approach to multi-species modeling would be best.
- DP recommended identifying hotspots where all four species are present and adding nearest neighbor-type analysis to help identify wetlands/catchments that are most important to the species for recolonization (metapopulation dynamics). JC also recommended quantifying spatial proximity to other wetlands and adding this as a covariate in the models.
- BH favored accounting for the clustered nature of data by modeling occurrence at the site-scale as conditional on occurrence at the catchment scale in a multi-scale occupancy model or using catchment as a random effect in a hierarchical occupancy model in addition to the multi-state occupancy models currently used.
- BH recommended Bayesian approaches over frequentist approaches because statistical significance can fail to detect trends in noisy systems. Bayesian approaches can produce probabilistic statements of certainty about trends for a parameter of interest using credible intervals. This approach can be useful for expressing confidence in a certain outcome and can be integrated with management reference points (e.g., the probability a 10% or greater annual decline occurred is 0.75).
- JC suggested GRYN consider analyzing the northern and southern ranges separately for some species since there are different things going on at each.
- JC emphasized that site-level occupancy was most meaningful for amphibians in GRTE and YELL and perhaps dropping the catchment concept in analyses would make sense. However, DP felt the monitoring was set up to evaluate catchments and was unsure of the value of the site-level analysis.
- SS suggested TTD as an addition to current approaches to enhance detection and modeling. See Question 3 for a summary of comments on TTD.

#### **Threats to amphibians**

- BH, PH, and SS thought the core of the monitoring program adequately addresses threats to amphibians: population declines, habitat change, climate change (changes in temperature, precipitation, evapotranspiration), die-offs, and invasive species.
- JC, BH, DP, SS, and MW recommended adding disease monitoring to the program to assess disease threats.

### Drivers of wetland change

Few comments focused on this question directly, but answers were woven into other parts of the review. Numerous references were made regarding drivers of wetland change including climate, hydroperiod, wetland quality (both water retention and water quality), invasive species, disease, and other drivers.

- BH commented that wetland change is well-addressed in the protocol, as evidenced by at least two peer-reviewed papers published by the network and should only improve over time and with increased spatial resolution of remote sensing data.
- SS stated that understanding wetland change is a strength of the current monitoring, though there is much to be done. One possibility SS suggested was adding fine-scale hydrography data to the multi-state model, particularly for the northern range subbasin.
- JC indicated that the current protocol could not address all drivers of change, but said the protocol does a good job quantifying wetland losses and gains, and this is most important.

## **Collaboration, Science Impact, and Integrating Science into Parks**

***Question 2: Is GRYN monitoring at the appropriate/most ecologically relevant scale? Are there ways to expand outside Grand Teton and Yellowstone national parks to the GYE? To the Western U.S.?***

### Current monitoring scale

- JC, BH, PH, DP, SS, and MW commented that the current monitoring scale is appropriate and relevant for the objectives and is effective for site, catchment, and park-scale inference. MC did not comment on scale.
- BH said sampling all (wet) sites in a catchment is the natural and appropriate scale to evaluate breeding amphibian occurrence, produces a higher sample size because it is more efficient, and reduces the likelihood of failing to detect shifts in breeding location. BH also said that sampling based on catchments likely captures the metapopulation structure of the target species. While there may be some spatial autocorrelation issues with this design, BH said this could be handled in the analysis.

- SS pointed out that there were no covariates for the catchment and major drainage scales and that drainage characteristics could be added without additional field effort.

### Suggestions for expanding

- MC, BH, PH, DP, SS, and MW mentioned partnering with external agencies and organizations adjacent to the park (USGS-Amphibian Research and Monitoring Initiative [ARMI], USFS, Partners in Amphibian and Reptile Conservation [PARC], Rocky Mountain Amphibian Project [RMAP], Boreal Toad Conservation Team, universities).
  - MC recommended PARC as a resource to share science, expand partnerships, improve coordination and communication, and expand the applicability of the data. (GRYN could request a special session at a NWPARC annual meeting).
  - MW explained that the Boreal Toad Conservation Team uses a metapopulation model that simultaneously predicts occupancy dynamics for both toads and *Bd* and scaled up comparisons could happen if GRYN collected comparable data.
  - BH, PH, and SS suggested using community science data to supplement current work (RMAP volunteer data; RMAP uses a similar design except GRYN uses one visit with two people and RMAP uses two visits with one to many people). PH offered Fletcher et al. (2016) as an example of integrating community science data into a model of species distributions.
- PH proposed occupancy analysis with shared data at larger spatial scales (e.g., regional) but that these frameworks are spatially explicit and assume all potential breeding habitat is known. PH said GRYN would have to use a spatially implicit version of this analysis. Examples of this type of work can be found in Howell et al. (2018, 2020).
- PH said an option to expand impact beyond the parks is to alternate sampling years and use saved time to sample sites outside the parks.
- JC suggested that remote sensing data could expand the area where wetlands are tracked.

**Question 5: How could GRYN increase the impact of their work? How can they work with parks and partners to transform knowledge into conservation action?**

It was clear from reviewers that they were impressed with the reach and impact of the work that GRYN was currently engaged in.

- BH felt GRYN is already conducting impactful work and disseminating findings to appropriate audiences including publishing in well-respected journals (beyond what is typical in most monitoring programs) and reaching broader audiences through Yellowstone Science.
- DP commended GRYN for doing an excellent job and being regional leaders. DP said the combination of a long-term data set, holding one of the first science reviews, excellent analytical techniques, publications in peer-reviewed science and managerial journals along with regular reporting in the NPS publication series, and being a leader in regional amphibian conservation is the gold standard. DP proposed that GRYN help other networks achieve similar success.

**Suggestions**

*Science communication*

Some reviewers acknowledged they did not have the full picture on GRYN science communication.

- MC asked if there was a communication strategy and conservation goals, which would help the program get more information to interested parties and partners that could facilitate transforming science into conservation action. MC felt that increasing impact is the area that could use the most attention and work in this monitoring program.
- JC advised that more emphasis should be placed on getting information to interpretation staff, which could include information about sterilizing waders to prevent disease transfer among wetlands.
- SS favored developing firsthand experiences with amphibians and their habitats that could demonstrate the effects of climate change and be an opportunity to educate the public on boot and gear disinfection protocols.

- MW suggested that data collected from die-offs could be used to promote disinfection of gear to the public.
- Annual meetings with a variety of participants were recommended by MW as a good way to disseminate information; they can be more meaningful than other communications because they spark discussion and facilitate leveraging of resources.

*Collaboration*

- BH recommended that GRYN continue to collaborate with universities and other agencies and continue to contribute to large-scale syntheses.
- PH promoted connecting with managers or researchers already sampling in other locations to share data.
- PH and MW proposed developing a community science program as a great way to expand the spatial scale, increase public engagement, and meet Objective 3.
- SS was not clear on whether there is direct integration of other GRYN vital signs with amphibian monitoring and mentioned that network collaboration could be important even if other I&M networks are not monitoring amphibians because they are monitoring trends in climate and water that could be relevant across larger regions.
- SS suggested partnering with neighboring landowners, including agencies and private landowners. There are models for private lands conservation, and this may be a way to incentivize landowners to maintain and protect wetland habitats.

*Other suggestions to increase impact:*

- JC thought getting feedback from parks about when and how they may use the data was important even if they don't specifically manage for amphibians.
- JC recommended adding reference points (thresholds for additional research or management action).
- BH pointed out that transforming knowledge into conservation action requires specific conservation objectives and GRYN amphibian monitoring data could inform several steps of the structured decision making (SDM) process.

- PH proposed shifting sampling effort in a subset of years to include additional sites to increase impact outside the current monitoring spatial scale.
- DP recommended making sure amphibian data are known and inserted into climate decisions and said David Thoma and Mike Tercek presented some ideas for this during the symposium.
- DP suggested the network think about replicating the grasslands and pinyon-juniper workshops NCPN has done in a wetland application.
- MW noted that the current work focuses on two of five potential responses to environmental stressors: occupancy and distribution. GRYN should also consider, if possible, abundance, health and malformation, and genetic change. MW said that while Chytrid (*Bd*) is not impacting the GYE to the degree it is in other areas, it may have more impact in the future as climate changes, so it could be valuable to collect data now to understand when that tipping point is reached.
- MW felt remote sensing to better understand pond persistence could provide good information on wetland drying and indicate years when survival through metamorphosis could not occur.
- MW commented that abundance is a missing link when trying to detect amphibian decline.

## Sustainability

***Question 6: How would you prioritize monitoring activities if budget/time/safety/personnel constraints restricted GRYN's ability to complete current activities?***

### Current monitoring

- DP, BH, SS stressed that the current monitoring is occurring at a bare minimum for maintaining scientific integrity in the trend analyses and that it is already constrained by budget. BH said to do everything possible to maintain annual visits to all medium- and high-quality catchments and SS said the program would suffer with further cuts.
- MC recommended prioritizing the linkage between amphibians and wetland habitats. MC felt that GRYN amphibian monitoring is likely a unique dataset of landscape dynamics and offers an opportunity to better understand amphibian

metapopulation dynamics with implications regionally, nationally, and globally given climate change impacts on amphibians, a broadly declining group of animals.

- SS stated that more funding is needed to meet objectives for tiger salamanders and western toads and SS said any reduction in funding would put the objectives for the two frog species that are currently being sampled adequately in jeopardy.

### Suggestions if future cuts to funding or other resources must be made

No one advocated for canceling the monitoring if funding became tight. Reviewers made suggestions for cuts but only if there was no other choice. See reviewers' full comments, including how they would prioritize these suggestions for cuts, in Appendix B.

#### *Reduce the number of catchments*

- JC identified dropping some catchments if necessary, and a power analysis could be used to find the lower limit of catchments required.
- BH brought up curtailing sampling at some selected catchments by dropping sites that are least accessible or randomly drop catchments if cuts were required, though BH stressed that maintaining all the catchments is obviously preferred.
- PH suggested that a prioritization of sampling at a smaller subset of sites based on the historical database and for which it is most important to capture inter-annual changes in occupancy or in wetland dynamics would be a first option if cuts had to be made. PH felt it would be straightforward to analyze the entire time series of data in a Bayesian framework despite a different number of sites in each year.
- SS recommended investigating whether there are sets of hydrographically similar catchments in drainages that could be sampled in alternate years. SS said that further reducing the number of catchments is a second choice but would likely reduce statistical power to detect trends.

#### *Reduce wetland sites*

No reviewers thought wetland sites should be reduced. Two offered "what if" scenarios:

- SS felt it may not be necessary to sample every wetland in a catchment, but it would not be much savings since getting to sites is the main time cost.

- MW suggested reducing the number of sites sampled in a catchment either randomly or in a way that specifically retains catchment-level data if significant cuts are required but preferred other options first.

#### *Eliminate entire drainages*

- SS stated that removing entire drainages from the monitoring would reduce costs and save time but would eliminate inference to significant portions of the parks.

#### *Alternating years/rotating panel*

- JC said switching to sampling in alternate years is a possibility if cuts are necessary and monitoring all sites in alternate years is better than sampling half the sites each year because annual weather differences can confound trends across sites. As a secondary option, JC put forth sampling the northern and southern ranges in alternate years or parks in alternate years as a possibility (though JC was uncomfortable with this option).
- MW proposed using a rotating panel design to monitor medium- and high-quality catchments.
- In contrast, SS recommended that years NOT be skipped entirely due to year-to-year variation in climate that strongly influences wetland availability. Rather, as mentioned above, SS suggested identifying catchments with similar hydrography and alternately sampling them so environmental-occupancy response could be tracked annually.

#### *Other suggestions*

- MW favored replacing medium- and high-quality far sites with closer medium- and high-quality sites to reduce travel time and expense and maintain current sample sizes.

- DP said this question should be considered in the context of the entire suite of monitoring programs at GRYN. The amphibian program cannot sustain much of a cut and still be effective, particularly considering the rarer species. Therefore, GRYN should consider reviewing its entire program and take the bold step of cutting entire protocols rather than a piecemeal reduction in amphibian sampling.
- BH and PH discussed single visits using TTD methods if safety allows but acknowledged the challenges of GRYN safety issues.
- MW mentioned limiting western toad surveys to one visit per year.
- PH and MW felt that developing a community science program or integrating existing community science data would provide information at a reduced cost.

#### *Suggestions if new efficiencies are realized and/or funding increases*

- BH and MW suggested meeting Objective 3 first.
- DP proposed adding more catchments to better meet objectives for tiger salamanders and western toads.
- BH and DP brought up incorporating eDNA monitoring but felt that all objectives should be met first.
- SS, BH, DP, and MW recommended adding disease sampling.
- MW felt strongly that low-quality catchments be added back in.

# Conclusions

The current design and analytical methods employed in GRYN's Amphibian Monitoring Protocol provide good, reliable occupancy trend estimates for two to three of the four target species identified in Objective 1. While most reviewers felt that the first objective was also being met for western tiger salamanders, there was concern by one reviewer that it was not being met. This needs to be further assessed. For the fourth species, western toads, there was consensus among reviewers that this species is not adequately sampled; sample size is too low if observations are restricted to the current long-term catchments. Considerable additional effort would be required to meet Objective 1 for western toads. There was also agreement among reviewers that Objective 2 is being met with the current methods. Objective 3 is not being met due to funding issues.

Current design and field methods were deemed to be the minimum required for successful monitoring of amphibians and wetlands. Reviewers understood that fiscal and safety constraints limit options for increasing sample size to meet the objectives for western toads and to increase analytical power for other species. Similarly, these fiscal and safety constraints limit the addition of data collection

methods that could improve our understanding of trends in amphibian breeding and wetland change. Reviewers offered recommendations on statistical approaches, which might benefit the analyses and interpretation of long-term data, yet all agreed these would be in addition to the current analytical approaches. They also suggested minor modifications to field work and sampling design that could improve inference, save time, and create a window for additional sampling.

GRYN was commended for the science, science transfer, and science communication they are already engaged in with this protocol. In fact, one reviewer suggested they should be held up as a model for other inventory and monitoring programs, both in terms of the science they do and the dissemination of information across a wide array of outlets and audiences. Reviewers commended GRYN for current collaborations and made suggestions for other collaborations to increase impact and efficiencies.

GRYN will determine what changes will be made by the end of 2022 and release a new version of the protocol based on this review, but at this time GRYN anticipates limited changes will be necessary to address concerns.

# Literature Cited

- Bennetts, R., P. S. Corn, R. Daley, W. R., Gould, C. Jean, D. A. Patla, C. R. Peterson, and A. Ray. 2013. Cooperative amphibian monitoring protocol for the Greater Yellowstone Network: Narrative, version 1.0. Natural Resource Report NPS/GRYN/NRR—2013/654. National Park Service, Fort Collins, Colorado.
- Cowardin, L. M., V. Carter, F. C. Golet, and E. T. LaRoe. 1979. Classification of wetlands and deepwater habitats of the United States. U.S. Department of the Interior, Office of Biological Services, U.S. Fish and Wildlife Service, Washington, D.C.
- Fancy, S. G., J. E. Gross, and S. L. Carter. 2009. Monitoring the condition of natural resources in US national parks. *Environmental Monitoring and Assessment* 151(May):161–174.
- Fletcher, R. J., R. A. McCleery, D. U. Greene, and C. A. Tye. 2016. Integrated models that unite local and regional data reveal larger-scale environmental relationships and improve predictions of species distributions. *Landscape Ecology* 31:1369–1382.
- Gould, W., A. Ray, L. Bailey, D. Thoma, R. Daley, and K. Legg. 2019. Multistate occupancy modeling improves understanding of amphibian breeding dynamics in the Greater Yellowstone Area. *Ecological Applications* 29(1):e01825. <https://irma.nps.gov/DataStore/Reference/Profile/2258537>.
- Howell, P. E., E. Muths, B. R. Hossack, B. H. Sigafus, and R. B. Chandler. 2018. Increasing connectivity between metapopulation ecology and landscape ecology. *Ecology* 99:1119–1128.
- Howell, P. E., B. R. Hossack, E. Muths, B. H. Sigafus, A. Chenevert-Steffler, and R. B. Chandler. 2020. A statistical forecasting approach to metapopulation viability analysis. *Ecological Applications* 30(2):e02038. 10.1002/eap.2038.
- Jean, C., A. M. Schrag, R. E. Bennetts, R. Daley, E. A. Crowe, and S. O'Ney. 2005. Vital signs monitoring plan for the Greater Yellowstone Network. National Park Service, greater Yellowstone Network, Bozeman, Montana.
- Mitchell, B., A. Chung-MacCoubrey, J. Comiskey, L. Garrett, M. MacCluskie, B. Moore, T. Philippi, G. Sanders and J. P. Schmit. 2018. Inventory and Monitoring Division protocol review guidance. Natural Resource Report NPS/NRSS/IMD/NRR—2018/1644. National Park Service, Fort Collins, Colorado. <https://irma.nps.gov/DataStore/Reference/Profile/2253227>.

# Appendix A. Questions and Instructions to Reviewers

List of Questions supplied to reviewers prior to the symposium to address with their review. Note that the questions (1–6) were supplemented with the more detailed list of topics to consider in their responses, which follow.

1. Is GRYN meeting objectives (Table 1) stated in the current amphibian protocol? If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these.
2. Is GRYN monitoring at the appropriate/most ecologically relevant scale? Are there ways to expand outside Grand Teton and Yellowstone national parks to the GYE? To the Western U.S.?
3. Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program? To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of their sampling visits?
4. Are current analytical approaches described in the protocol and in Gould et al. (2019) capable of detecting statistically significant trends for all species involved? For detecting principal threats to amphibians? For understanding the drivers of wetland change?
5. How could GRYN increase the impact of their work? How can they work with parks and partners to transform knowledge into conservation action?
6. How would you prioritize monitoring activities if budget/time/safety/personnel constraints restricted GRYN's ability to complete current activities?

As you prepare responses to questions above, reviewers are asked to consider the following topics (e.g., Sampling Design, Methods & Metrics) and changes the program has identified:

1. **Sampling Design:** Sampling design changes that have occurred since protocol implementation. Examples of these changes include additions or reductions in number of sites monitored, changes to frequency of sampling, and altered response design (how and what is measured at each site).
  - a. A reduction in catchment numbers (40 to 31; GRYN dropped surveys in 9 low quality catchments) occurred in 2011. This reduction was taken in coordination with a statistician in order to help reduce costs and because greater sampling effort was needed to represent the low-quality stratum.
  - b. Discuss how current sampling design limits our ability to document trends of rare species (e.g., western toads and to some extent tiger salamanders). While this understanding has been further informed through monitoring, we seek better elucidation regarding their effective inclusion in future monitoring.
2. **Methods & Metrics:** Consider the addition or removal of methods or metrics, as well as revising methods in a way that may not be compatible with old methods. Revised methods need also consider the cost of changing to a new approach when that affects the continuity of the data set.
  - a. Climate/water balance metrics have been used in multiple analyses. Model ranking indicates these metrics are useful in describing occupancy dynamics – should other climate/water balance metrics be considered?
  - b. Multi-state occupancy models have been used in recent analyses and better reflect habitat/breeding condition states relative to original protocol proposed analysis techniques and allows co-modelling of wetland and amphibian dynamics.
  - c. Survey method considerations – are visual methods the most sensitive technique for detecting larval amphibians?

3. **Monitoring Objectives:** Changes to monitoring objectives alter the intent of the monitoring and warrant scientific review. Altered monitoring objectives affect the protocol's sampling design, data collection methods, analytical approaches, and quality assurance steps.
  - a. Changes to objectives have not been made, but GRYN is not meeting all objectives equally (i.e., inability to detect trends in rare species [Western Toads]).
  - b. Are GRYN's monitoring objectives still relevant?
    - i. Check on relevance, to ensure that monitoring objectives are appropriate, achievable, and use methods that produce quality data to support sound science-based management of park resources.
4. **Linking Monitoring Activities:** Assess potential alignment and compatibility with related vital signs and park management needs, and effects of changes on other network monitoring activities.
  - a. Links to climate drivers/water balance metrics have been explored.
  - b. GRYN remains committed to exploring opportunities to collaborate with ongoing disease monitoring and wetland water-level monitoring on Yellowstone's Northern Range.
5. **Boundary-spanning Opportunities:** Opportunity to collaborate or align with other NPS Inventory & Monitoring networks or other monitoring programs that have similar protocols and are conducting similar work in the GYE.

## Appendix B. Full Reviewer Responses to Questions

The following seven documents are the full responses received from reviewers. We made only minor formatting changes and fixed three misspelled words for publication purposes.

## Jane Carlson

### *Greater Yellowstone Network Amphibian Monitoring Protocol Review*

Prepared For: Michael Bozek (michael\_bozek@nps.gov)

Prepared By: Jane Carlson, Ph.D., Ecologist for Gulf Coast I&M Network, Lafayette, Louisiana

Date: January 15, 2021

***Question 1: Is GRYN meeting objectives (Table 1) stated in the current amphibian protocol? If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these.***

#### Objective 1

Yes, GRYN appears to be meeting the objective of estimating the proportion of sites occupied by four amphibian species and the rate of change in use over time. For the rarer species, especially the western toad, there is a shortage of data, such that many potentially useful covariates can't be included in the models. It was recommended that to address this, the network could explore multi-species modeling approaches, to borrow detectability estimates across species. This isn't my area of specialization, but it seems like it could be a good idea—better than asking the network to gather a bunch of additional co-variate data, since they already seem to measure quite a few covariates. That said, perhaps some of the covariate variables could be improved, although I know this can end up as never-ending quest.

#### *Covariates*

One covariate I'd look at more closely is the one quantifying the spatial proximity to other wetlands: nearest neighbor distance. If there was opportunity to assemble spatial data on wetlands every year using remote sensing (see objective 2 paragraph below), proximity could be characterized differently, such as the percent of area within a X-meter buffer around the wetland of interest that is also wetland—that gets at both distance to nearest neighbor wetlands and size of those wetlands. Though it matters what buffer size you would choose, and a lot of computer work would be required. Sounds fun to me, though (no, I'm not available to help out, sorry!).

#### *Increased use of GPS units*

One of the presentations brought up an occupancy approach that uses time-to-detection rather than double observer. Since the network really needs to have two people travelling to each site for safety reasons, then I don't see a reason to move away from the double observer approach. But if the network wanted to try out time to detection, I might recommend that crew use GPS units to help out. Our field crews carry around small GPS units (GARMIN GPSMap64 st) for pretty much every protocol, allowing us to quickly store waypoints and tracklogs. Your crews use GPS units too, and if they're not already doing so, taking waypoints and tracklogs is one way to time-stamp the start and end of an activity that can be a little faster than pulling out a datasheet or tablet. For our tortoise monitoring, we have pre-loaded our sampling polygon boundaries on to the GPS units, which crew member use to keep themselves within the tortoise survey units. I could imagine that if your crews were trying to retrace steps or visualize wetland boundaries from years past, pre-loaded shapefiles could be useful for them as well. In our case, our shapefiles are saved as .gpx tracklogs, since these units aren't able to display GIS shapefiles. In any case, if this approach seems like it could be useful, you can contact me and we'll help out however we can.

If crew members really wanted to get fancy with the GPS units, they could walk the perimeter of the wetland each time and record the tracklog, so the wetland size is better quantified than just length and width. I understand that this could add extra time in the field, though, so I wouldn't take it on lightly.

#### *Site vs. catchment level analyses*

If the analyses are really focusing on the site-level (and this makes the most sense to me, too), is it worth retaining the catchment level tables and models into the future? Is there a way to downplay or drop the catchment concept, except perhaps when you explain the initial study design? It would also be included as an effect in the models

as a spatial block, but in agreement with what the network has leaned towards, reporting on occupancy at the catchment level isn't as meaningful or interpretable as at the site level. I'd also recommend describing catchment quality language differently so it doesn't indicate that you're not monitoring poor quality sites. You measure site quality in a continuous way in each visit. I'm not particularly worried about the exclusion of the low quality catchments, because high and medium are most logical in terms of payoff for time invested. GRYN is still sampling low-quality wetland sites, and that's an important point, especially now that the analysis focuses on the site level, rather than catchment level. Yes, most of the park has the low quality catchments, but that's because most of the park is minimum to no wetland. So the spatial coverage of the wetland area is actually pretty good. And just a detail about the wording in the protocol describing the catchment classification. Taken at face value, the classifications are overlapping and not mutually exclusive. You might just need to add at the start of each classification criterion something like: 'if not a low quality catchment, then ...'

#### *Difference between objective 1 and objective 2*

All in all, I support the current multi-state occupancy analysis approach, which includes modelling wetland status transitions simultaneously with occupancy status transitions. It seems like you can't get around that, and its useful for interpretation: you may see loss of occupancy if wetlands are drying up, or you may see loss of occupancy without wetlands drying up. If it's climate warming, then the park can't really do anything about it, but if it's local factors not related to pond-drying, then maybe they can. But the fact that wetland presence or absence is imbedded in objective 1 it does raise the question of whether there's any work left to be done under objective 2 as it is currently written. I'll share a few more thoughts on that in my objective 2 paragraph.

#### Objective 2

Yes, GRYN seems to be meeting its objective of determining the number of wetlands available in the focal catchments each year. As I mentioned above, the objective is basically met through the analytical process required for objective 1, although 1 is about modelling the probability of changing status in a given year, and objective 2 appears to require no statistics at all. It might be fine just to leave this objective alone, but as came up in discussions during the meeting, there could be an opportunity to expand this objective so it has more impact.

Specifically, I think it would be worth setting up a process to measure wetlands using remote sensing across a wider sampling frame, even if forested areas can't be included. This would have a pretty big impact for the parks, I would think. As far as things to add, it would probably not be too costly, and it wouldn't require doing every year—probably just in sync with the trend reports. I agree with Kristin that once the analytical process was researched and set up by a cooperator, then the future work could probably be completed in-house.

#### Objective 3

No, GRYN is not currently meeting the objective of checking the full suite of known western toad sites each year. I suppose it's most practical to drop this objective from the protocol, and have it be a repeat inventory or side project, depending on outside interest. We've also had to accept for our amphibian protocol that we do not have the methods to track rare things, even though we all get extra perky when we encounter them or think of them.

#### ***Question 2: Is GRYN monitoring at the appropriate/most ecologically relevant scale? Are there ways to expand outside Grand Teton and Yellowstone national parks to the GYE? To the Western U.S.?***

I'm comfortable with the scale of monitoring. As I mentioned for Objective 2 above, the use of remote sensing data could really expand the area over which wetlands are tracked, and I think that could have big benefits for science and monitoring and for the state agencies concerned about these species, if data and methods could be shared. Since I'm probably the most eastern person involved in this review, and I've never even been to Yellowstone (!), I'll refrain from trying to guess about other opportunities for expansion around the parks and in the western USA.

***Question 3: Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program? To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of out of their sampling visits?***

I'm comfortable with the field methods employed. It makes sense to focus on visual detection of larval amphibians with two crew members walking around the wetlands. From what I heard about eDNA during the meeting, it doesn't sound reliable enough, the logistics seem challenging, and it also would not mesh with all of the earlier data collected for this protocol. In fact, it seems pretty darn important to maintain continuity with past field methods—unless it could be shown that crew were really missing a lot, but I didn't hear anything to that effect.

From the perspective of this protocol review, I don't see a need to add work for the field crews so they could record more about the biodiversity of wetland flora and fauna. The network may have other reasons or interests in doing this, which is fine, but perhaps outside the purview of the protocol. I wouldn't mandate recording other species in the protocol unless they were likely to have a strong impact on amphibian presence, like fish or other predators that would eat the larvae, or plants that provide food and habitat, which the crews already record to some extent.

***Question 4: Are current analytical approaches described in the protocol and in Gould et al. (2019) capable of detecting statistically significant trends for all species involved? For detecting principal threats to amphibians? For understanding the drivers of wetland change?***

**Analytical approaches**

My comments on the analytical approach are pretty much captured under question 1. It may not be possible to really track the western toad with this protocol, but I would not recommend dropping it from among the 4 species in objective 1. Currently, the species is too rare for the models to work well, and indeed, there is outside evidence that they are declining in the area, but the monitoring protocol is not picking that up. I'd still argue to keep toads in because they could become more common in the future, and it's not really extra work to record them during routine fieldwork. If it seemed important to recognize this the challenge of the western toad in the language of objective 1, it could be stated that the protocol focuses on all amphibians that are sufficiently common to be monitored with these methods, which would be up to 4 species.

It came up a few times that different things were going on in the northern and southern range of study area. Perhaps those two areas should be analyzed separately, at least for some species?

In general, I would argue that the protocol should not be terrible prescriptive in terms of methods of analyses, leaving flexibility for improvement and new techniques. In contrast, the study sites and field methods should remain consistent over time.

**Principle threats to amphibians**

The protocol does not explicitly monitor for amphibian diseases, which could be a principal threat. Is ARMI covering the disease tracking work? It may not be possible to explicitly add an intermittent check for disease in the protocol, but it seems important that relationships are maintained with cooperators so that disease presence is available as a viable explanatory variable if the network detects a declining trend.

**Drivers of wetland change**

Wow, that's a tall order. I suspect that the protocol is not capable of pinpointing all the key drivers of wetland change, but I think it's doing as good as one could hope in the context of an amphibian monitoring study. If the protocol does a good job quantifying wetland losses and gains, that's most important. Whether there are more useful covariates out there can remain a question for future scientists.

**Question 5. How could GRYN increase the impact of their work? How can they work with parks and partners to transform knowledge into conservation action?**

It seems like there is opportunity to communicate more and better with park staff about their amphibians. It would be nice to get feedback from the parks about when and how they'd be likely use data from this protocol to make management decisions. I think it's OK that the parks might not manage their lands for amphibians, but then you still need to identify with them what other value these data might have. Getting information into the hands of interpretative staff could help, or even communicating with the public directly about your knowledge on amphibians. The public presentation you all had at the end of the meeting was great, and I was really impressed by the interest and questions from the attendees at the end. It seems like there is definitely opportunity for more public outreach and involvement in that part of the US.

Another suggestion that might help GRYN identify ways to transform knowledge into action: For our amphibian protocol peer review process, we were asked to identify thresholds for when we'd want to trigger additional research, communicate with the park, etc. That seemed like a good suggestion to think about, anyway, and I can share the little bit that we wrote. The topic comes up in our protocol narrative ([DataStore - Published Report - \(Code: 2256511\) \(nps.gov\)](#)) and in SOP AM 07 [Standard Operating Procedure AM07: Data Analysis and Reporting—Version 2.0 \(nps.gov\)](#)

**Question 6. How would you prioritize monitoring activities if budget/time/safety/personnel constraints restricted GRYN's ability to complete current activities?**

Ho hum... this would be my least favorite question to think about, and happily for me, it's not my area of specialization. My first instinct would be that GRYN drop some catchments, since I don't feel like they're already at their lower limit in terms of sample size and spatial coverage. To quantify where that lower limit would sit, power analysis could be employed, and possibly, this has already been done. It seems like GRYN has been OK with dropping and adding catchments already, so this approach has probably already been considered a fair bit.

If a real significant savings was necessary, GRYN could switch to alternate years of monitoring rather than every single year, although I know with our annual budgets, this is not especially practical. I would prefer a switch to all site in alternate years versus sampling half of the sites each year, just because for the half-per-year approach, differences in annual weather are confounding with documenting trends across all sites. Perhaps if GRYN settled on analyzing the northern range apart for the southern range, then they could sample all northern sites in one year and all southern units in the other, or Teton in one year and Yellowstone in the other. But that sits uncomfortably for me, relative to sampling all sites in the same year.

## Michelle Christman

### *Greater Yellowstone Network Amphibian Monitoring Protocol Review*

**Prepared For:** Michael Bozek (michael\_bozek@nps.gov)

**Prepared By:** Michelle R. Christman (michelle\_christman@fws.gov), Fish and Wildlife Biologist - Federal Agencies Coordinator for Partners in Amphibian and Reptile Conservation (PARC), Inventory and Monitoring Branch, Natural Resource Program Center, U.S. Fish and Wildlife Service, Fort Collins, CO.

**Date:** January 31, 2021

Thank you for the opportunity to learn about the Amphibian Monitoring Program occurring within the Greater Yellowstone Network, as part of the National Park Service Vital Signs Monitoring Program and offer suggestions or insights on the program. My understanding is that a variety of outside partners were asked to provide their perspective and expertise. Therefore, my perspective for this review will focus on the application of the science and monitoring for management and decision making and leveraging partnerships for greater conservation effectiveness.

In the Review document provided to reviewers, the following was presented:

*The steps in the review process (outlined below) that combined, will allow for reviewers to address three dimensions of the current program:*

- *a comprehensive science critique of GRYN's amphibian monitoring program as it's currently being conducted, including offering insights and ideas into the strengths and weaknesses of the program, while offering better alternatives where appropriate,*
- *a careful review of how the work is meeting current information needs and recommendations for prioritizing monitoring activities, and*
- *opportunities to engage in collaborative work that can increase the impact of this amphibian monitoring program both within the park (intra-ecological impact) and across the Greater Yellowstone Ecosystem (inter-ecological).*

I understand that the presentations of science with facilitated conversations among experts working on amphibian monitoring and reviewers was intended as a format for reviewers to learn and ask questions to complete a thorough review. This was without a doubt, an excellent format to learn about and better understand the vast amount of work and complexities of the science (and satisfies bullet one above), and touches on the second bullet. This is also a better strategy than a 200 page document that few would want to wade through for an understanding. However, as an outside reviewer (not familiar at all with the program), particularly as one focusing on the application of the science to management decisions (third bullet), this format was challenging. First, there is a lot of work happening, and it was difficult to sort the work of the program, the work of partners, and how work is coordinated and what objectives are being met where (e.g. who was doing what and why). For example, some partners are pursuing independent research questions, which support the overall understanding of the system, and some work is specifically being done to meet program objectives. However, it was difficult to understand linkages or relationships of the amphibian work across the Network (where work was coordinated and contributing to certain objectives, versus, work that is addressing certain questions, that is occurring in the same place, focusing on the same targets of amphibians or their habitat). For me, this also contributed to an unclear overall understanding of how the amphibian monitoring program and information gained from it is being used to understand the ecological significance of amphibians and amphibian habitat and applied either for conservation or management objectives or decisions. I learned much about the science and implementation of the protocols, changes and factors that were considered regarding the protocols with changing science. It appears that there are strong partnerships and communication, and broad support of conducting high quality work related to amphibian and amphibian habitat inventory and monitoring. However, how the Network and partners use the data to manage resources was not abundantly clear in this format. I got the impression (this could be wrong) that the overall learning and understanding of the monitoring data is and can be used, particularly in

certain specific management situations or decisions; however, it was not apparent if or how the program and its science is synthesized and applied to the overall strategic management of the resources managed in Network. While annual reports may exist, there is too much information to attempt to get through and synthesize in a meaningful way, especially in a short time frame. I would recommend that for future reviews of the GRYN program or other Network reviews that the leads prepare summaries of key attributes of the programs for which feedback is desired. It is with this perspective that I provide recommendations for increasing the impact of the amphibian monitoring program, and transforming the information into conservation action (limited understanding, and recognition that the application of the information may be underway, but may not have been communicated in a way to clearly understand).

Reviewers were asked to provide responses to the following questions about current monitoring activities:

***Question 1: Is GRYN meeting objectives (Table 1) stated in the current amphibian protocol? If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these.***

From conversations during the review process, I believe other reviewers will focus on providing important feedback on this specific question as it relates to the science of inventory and monitoring strategies to meet enumerated objectives and criteria. I have one comment that goes beyond the question of whether the stated objectives have been met or not, as to how the objectives relate to a reference point/s. I do not have a good understanding of how, meeting the objectives (which, from the presentations, have seemingly been met to various degrees depending on the objective), can be or are integrated into management decisions. First, what baseline are the monitoring objectives in reference to? For example, are toads rare across the Network because toads are inherently rare, or because they have experienced a decline prior to initiation of monitoring, and what remains is now uncommon? This question would apply to all species being inventoried and monitored, is what is the baseline or reference condition, and what is the management objective? Is it simply to observe change? Is it to maintain the status at the start of the amphibian monitoring program? Is it to identify what a healthy, functional ecological system looks like, and the role of wetlands and amphibians have in that system, and how to develop strategies needed to get there? There seems to be a wealth of understanding how natural ecological functions have changed within the GRYN, depending on the driver/s (e.g. fire, wolves, etc), and this altered ecological context through time was not apparent as being part of the amphibian monitoring program.

A better understanding of how meeting the stated amphibian monitoring program objectives contribute to management objectives or decisions would also help understand why all amphibians present in GRYN are not included. Given that the area is not specious for amphibians, the consideration of one (Plains spadefoot toad, at one known site) or two (if you were to include Northern leopard frogs, with at least one historical observation in the area) species into the program would provide a holistic approach and understanding for managing for all amphibians and their habitats in the area, and a better understanding of historical baseline, desired conditions for management, and could better contribute reaching management objectives, for not much, if any, additional work.

***Question 2: Is GRYN monitoring at the appropriate/most ecologically relevant scale? Are there ways to expand outside Grand Teton and Yellowstone national parks to the GYE? To the Western U.S.?***

Is there communication with other NPS networks that have included amphibian monitoring as part of their vital signs, or other agencies (e.g. National Wildlife Refuges, Forest Service, Department of Defense BLM) regarding with their Inventory and Monitoring efforts, or how some of their management needs could benefit from either sharing information or protocols, or how the information and understanding from GRYN could be applied?

PARC is a network of partners that spans state and federal agencies, researchers, land managers, and more that is most known for supporting the connections and partnerships of people, and applying science to amphibian and reptile conservation, at multiple scales. PARC could be a resource to GRYN and other networks to share science, expand partnerships, and improve coordination and communication. One strategy could be to request a special

session or symposium at a NWPARC annual meeting, not just for sharing information from the amphibian work in GRYN, but also as a tool to expand the applicability of the information, and to seek new collaborations and partnerships to leverage the ongoing work in GRYN.

***Question 3: Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program? To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of out of their sampling visits?***

One question I had here is if the option of adding additional “nearby” sites that are in relatively close proximity to existing sites that always get monitored has been considered (for adding efficiencies)? For example, if travel to and from a site is the most (or one of the most) time consuming part, are there “lower quality” habitats/potential sites nearby that have not been considered for monitoring because they are “low quality”, could certain low quality sites be easily added in some cases? While perhaps not “independent” sampling sites, could you gain important data not only regarding the wetting and drying of these “low quality sites” but perhaps also monitor the expansion and contraction of amphibians from some of the regularly monitored sites that could act as a source in contributing to metapopulation dynamics. When I think about “wet years” and “dry years”, the wet years are likely very important for maintaining connectivity and metapopulation dynamics, providing the stepping stones for expansion, genetic maintenance of populations, and ultimately, maintaining the health of those populations. This is especially important for rare amphibians, because of the direct correlations between genetic variation of populations, resistance to stressors and pathogens, and ability to respond to stochasticity (both environmental and genetic). However, if there are few (or fewer) opportunities (fewer number of wet years, fewer consecutive number of wet years, or shorter duration of intermittently wetted sites) for amphibians to disperse and maintain genetic variation across sites, this can ultimately influence the metapopulation dynamics, and ultimately site occupancy (e.g. the “extinction vortex”). There seems to potentially be a unique opportunity to better understand amphibian connectivity and metapopulations that could help inform not only how to best manage for healthy ecosystems within GRYN, but also well-outside of GRYN, given the long-term data sets, and that both rare and relatively more abundant species are monitored. For example, looking to the future, considering climate change alone, let alone other factors influencing, reducing availability, and often degrading habitat, moving beyond managing for the sensitive, rare, threatened and endangered, while still important, identifying strategies and management actions to keep “the common species common” will also be critical. The point here is that “doing nothing” or “monitoring natural processes” are management strategies, that will more than likely only lead to the observation of declines as habitat availability maintains its trajectory and decreases over time. If connectivity and metapopulations dynamics are better understood, targeted management actions could be trialed to support the long-term maintenance of habitats, and their associated amphibian populations.

***Question 4: Are current analytical approaches described in the protocol and in Gould et al. (2019) capable of detecting statistically significant trends for all species involved? For detecting principal threats to amphibians? For understanding the drivers of wetland change?***

I’ll defer to other reviewers to address these questions.

***Question 5: How could GRYN increase the impact of their work? How can they work with parks and partners to transform knowledge into conservation action?***

Science communication may be the most difficult aspect of science, and yet equally as important as the science itself. From what I’ve seen, it appears that several strategies have been employed to communicate the amphibian monitoring science of the GRYN; however, I would ask if communication strategies are being leveraged to target the impact of the work. Does a Communications Strategy exist for the program? If so, does it have targets and audiences identified? For example, specific audiences (Park management, NPS Leadership, other park/Network programs, in-reach into the parks and outreach to local, regional, national partners) that could facilitate transforming the knowledge/science/information into conservation action? Have conservation goals been identified that then leads to identification of conservation priorities and action? From what I’ve learned,

this Question 5 is the area in the amphibian monitoring program that could use the most attention and work. An adequate response to this question would require a more in-depth understanding of existing goals, objectives, communications, plans etc that link the inventory and monitoring work to conservation objectives, management objectives, and desired landscape conditions or desired outcomes. I would be happy to follow up afterward to continue conversations about potential strategies of transforming knowledge to conservation action, as to me, this is the most critical part of a program, such as the amphibian monitoring program, and may be the portion that could use the most attention for effectiveness.

***Question 6: How would you prioritize monitoring activities if budget/time/safety/personnel constraints restricted GRYN's ability to complete current activities?***

I would recommend that the linkage between amphibians and wetland habitats be prioritized for maintenance. Oftentimes, changes to habitats are not well understood, and occupancy data are limited without the understanding of habitat dynamics. Again, other partners will likely better comment on the statistics of sites and how to add, reduce, or otherwise alter sampling schemes. I will just refer back to that the ongoing inventory and monitoring occurring in GRYN is likely a unique dataset of landscape dynamics including amphibians and wetlands, and with it, likely offers a unique opportunity to better understand amphibian metapopulation dynamics. I would recommend keeping this in mind, that learning from this system could have implications not only for the GRYN, but regionally, nationally, and globally given climate change impacts on amphibian metapopulation dynamics and conservation of a broadly declining taxonomic group of animals.

## Brian J. Halstead

### *Greater Yellowstone Network Amphibian Monitoring Protocol Review*

**Prepared For:** Michael Bozek (michael\_bozek@nps.gov)

**Prepared By:** Brian J. Halstead, Research Wildlife Biologist, U.S. Geological Survey, Western Ecological Research Center, Dixon Field Station, 800 Business Park Drive, Suite D, Dixon, CA 95620-4309; [bhalstead@usgs.gov](mailto:bhalstead@usgs.gov)

***Question 1: Is GRYN meeting objectives (Table 1) stated in the current amphibian protocol? If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these.***

The Greater Yellowstone Network (GRYN) Amphibian Monitoring Protocol outlines three primary objectives related to amphibian breeding and habitat. The justification and rationale for each objective is sound, and all objectives are important. Because the protocol's effectiveness varies by objective and recommendations are often objective-specific, I treat each objective in turn below.

The first objective of the GRYN Amphibian Monitoring Protocol (hereafter, protocol) is to “estimate the proportion of catchments and wetland sites used for breeding by each of the four common, native amphibian species annually, and estimate the rate at which their use is changing over time.” This objective is largely being achieved by the current monitoring protocol. The sampling design is generally very well-conceived and implemented. Revisiting the same catchments and sites annually will provide more precise estimates of trends than going to different catchments and sites each year or split or rotating panel designs (Hankin et al. 2019). Dropping the low-quality catchments in 2011 has the effect of reducing inference to catchments with any mapped wetlands to catchments with any semi-permanent and >1 ha seasonally flooded wetlands. This effect is likely to have a larger influence at the catchment scale than at the site scale, where the quality of individual sites can vary greatly across years. Nonetheless, I don't expect this change to be consequential to the stated objectives, as most amphibian populations appear to occur in these medium- and high-quality catchments (Gould et al. 2012). The exception would be if catchment quality changes over time with changes in climate. In that case, state transitions of catchments (like those of wetland sites) also would be of interest and worth monitoring. A related concern is that if sites are chosen based on being occupied or having high-quality habitat, then inference about trends can be biased toward documenting declines (Pechmann et al. 1991, Fournier et al. 2019, Mentges et al. 2020). Because most amphibian dynamics are occurring at sites, rather than catchments, I do not expect either changes in catchment quality (in terms of extent of wetlands) or potential bias caused by selecting medium- and high-quality catchments to be problematic at management-relevant time scales.

Another effect of dropping low-quality catchments is simply the reduction in sample size and, therefore, precision. Because it appears that few of these catchments and sites were occupied by amphibians (Gould et al. 2012), this reduction in precision is likely inconsequential to inference from the program, particularly at the site level. Nonetheless, it is important to continue restrict analysis of trends to those (medium- and high-quality) catchments monitored consistently as done in Ray et al. (2016) and Gould et al. (2019) to avoid biasing occupancy upward in recent years as unoccupied catchments were dropped from the monitoring protocol.

The field methods for achieving the first objective of the protocol are appropriate and likely the best methods available for surveying the amphibian assemblage in the GRYN. Please see my response to Question 3 for more detailed commentary on current and potential field methods.

Although analytical methods are covered elsewhere in this review (see Question 4), I will spend some time here discussing the general approach and some alternatives not specifically related to detecting trends, threats, or drivers of wetland change. The occupancy approach to monitoring and analyzing (breeding) amphibian status and trends at a landscape scale is necessary and appropriate, as more detailed abundance-based metrics are infeasible at these larger spatial scales. In addition, amphibian populations often demonstrate large fluctuations,

and estimating trends from these naturally noisy data are difficult. Furthermore, in a metapopulation context and in dynamic systems such as the wetlands amphibians use for breeding, occupancy is a more appropriate metric than local population size.

The use of multi-state occupancy models that condition amphibian breeding occurrence on wetland state is an excellent use of the monitoring data. It achieves the analysis of two monitoring objectives within a single analysis and appropriately forces amphibian breeding to only occur in wetted sites. Current analyses are single-species dynamic multi-state occupancy models, carried out at both the site and catchment scales (Gould et al. 2012, 2019, Ray et al. 2016). Recent developments in occupancy modeling, including the Bayesian analysis of such models, allows for several features that could benefit the analysis. I discuss these models and how they might be used to combine information across the larger data set to potentially improve precision and performance while evaluating fewer models.

The first and, in my opinion, most important way that data analysis for Objective 1 could be improved would be to explicitly account for the clustered (also considered hierarchical, grouped, or nested) nature of the data. Sites are naturally nested within catchments, and the occurrence of amphibians at sites within catchments is more closely linked (less independent) than at sites found in different catchments. Two possible approaches to deal with the clustered nature of the data include modeling occurrence at the site scale as conditional on occurrence at the catchment scale in a multi-scale occupancy model (Nichols et al. 2008) and using catchment as a (logit-normal) random effect in a hierarchical model for occupancy. The multi-scale occupancy approach essentially fixes site occupancy at zero when the catchment is not occupied by the species (though it should be noted that the data collected at sites, including imperfect detection, are used to infer presence or absence at the catchment scale). It also focuses the detection process on the natural unit described by the data. In other words, the detection process is modeled based on independent surveys at the site scale without pooling across sites within a catchment as currently described for catchment scale analysis (Gould et al. 2012). Although intuitively appealing, inference at the catchment scale of a multi-scale occupancy model might be somewhat limited by sample size ( $n = 30$  catchments). Furthermore, sites in unoccupied catchments provide no further information on the occurrence patterns of amphibians at the site level, which might further limit statistical power to detect effects of covariates on site-level occurrence or trends in occupancy over time. Regardless of these limitations, I think it is worth further exploring multi-scale occupancy models, not as a replacement for the multistate models, but rather as a supplement to address different questions. Combining the two models is possible, but I don't know of any examples of such a model.

A simpler approach would be to treat catchment as a (logit-normal) random grouping effect with catchment-specific occupancy allowed to vary around a park-wide mean occupancy probability. Including random intercepts for catchments is straightforward in a Bayesian setting and would fit easily into the multistate framework, as it is just part of the linear predictor for site occupancy. Although this approach is not as intuitively appealing as conditioning site occupancy on catchment occupancy, it does allow site occupancy rates to vary among catchments and allows dependence structure among sites within catchments. This partial pooling is a good analytical solution in situations where non-independence among sample units might occur (Gelman et al. 2003, Gelman and Hill 2006, Hobbs and Hooten 2015).

The hierarchical model described above is easy to implement in a Bayesian analysis, and this approach would likely be beneficial for other variables. For example, observer effects are often evaluated in active searches for animals; when more than a few different observers conduct surveys it is more efficient and parsimonious to treat observer as a (logit-normal) random effect on detection probability unless interest in the performance of specific observers exists. Treating year as a random effect for detection, extirpation (or persistence), and colonization probabilities also is more efficient than treating each year as a completely independent parameter to be estimated from the data. Of course, one could also estimate a linear trend of year or annual climatic covariates (Ray et al. 2016) on these probabilities in addition to random effects.

A final and more recent development in the use of hierarchical occupancy models is multi-species models that treat individual species and their responses to variables of interest as random effects. Such multi-species

occupancy models improve the performance of occupancy models for rare species by sharing information across species, with the influence of each species on the community mean and each other implicitly weighted by the amount of data for the species (Zipkin et al. 2010, Hunt et al. 2013, Sauer et al. 2013). A primary caveat in these (and other) hierarchical models is that the units (in this case, species) are assumed to be “exchangeable,” meaning that they can be thought of as random realizations from a statistical distribution. This assumption likely holds for pond-breeding anurans, and likely even for pond-breeding amphibians because of their biphasic life histories. The other limitation to multi-species (and again, other hierarchical) models is that they perform best when there are many groups (catchments, observers, or years above; species here). Because the pool of pond-breeding amphibians in the GRN [sic] (GRYN) is small (4 species), there is little information with which to estimate hyperparameters (i.e., the variance of the distribution from which species occupancy or the effects of variables on occupancy are drawn), so these variances will likely be estimated with poor precision and the benefits of shrinkage realized to a lesser extent than in larger assemblages. These parameters are often not of interest, so the utility of multi-species occupancy models is still worth exploring.

A final note on these potential modeling approaches is that they are not intended to replace either the multistate models currently in use nor the model selection and model-averaging approaches used. Ideally, one would add the features described above to the current multistate models. Although AIC cannot be used in a Bayesian MCMC setting, a number of information criteria have been developed for Bayesian model selection (Gelman et al. 2014, Vehtari et al. 2017). Alternatively one can use variable selection techniques for model selection and averaging (O’Hara and Sillanpää 2009, Tenan et al. 2014, Hooten and Hobbs 2015); when model selection is limited to selecting influential variables, this is my preferred approach for its simplicity and because the model only needs to be run one time, rather than separate runs for each model. My final plug for Bayesian analysis of these models is that they make much more intuitive probabilistic statements about estimated parameters than frequentist methods. I expand on this idea in my response to Question 4.

The second objective of the GRYN Amphibian Monitoring Protocol is to “determine the total number of wetlands within sampled catchments that are suitable for amphibian breeding annually.” This objective is very closely tied to the first objective, and most of my comments there apply here as well. The current protocol is achieving this objective, with the same sampling design and caveats regarding the dropping of low-quality catchments. As for the first objective, field methods are relevant and adequate, as is data management and analysis. The use of multistate occupancy models to jointly achieve Objectives 1 and 2 is a strength of the design and analysis of this protocol.

The potential for remote sensing to enhance this objective is great. Recent developments that use sub-pixel satellite information to inform the extent of water in small wetlands at frequent intervals (approximately every 2 weeks if interference from cloud cover is not a problem) can provide important information relevant to pond-breeding amphibians across a much larger landscape than can feasibly be visited (Kissel et al. 2020). Remote sensing also provides hydroperiod information that isn’t available in a single monitoring visit to wetlands; this information has the potential to provide a more nuanced view of climatic and other drivers of wetland dynamics, and in turn, the influence of wetland dynamics on amphibian occurrence. A major limitation to remote sensing will be canopy cover; nonetheless, for wetlands in meadows and in the more open Northern Range, the development of remote sensing tools to better characterize wetland dynamics in space and time is exciting.

The third and final objective of the GRYN Amphibian Monitoring Protocol is to “...estimate the proportion of previously identified breeding areas that are used annually [by boreal toads], and estimate the rate at which their use may be changing over time.” As indicated in the workshop and in Bennetts et al. (2013), the supplemental sampling design for boreal toads is only implemented after work associated with Objectives 1 and 2 have been completed. Given the short field season and intensive effort required to achieve Objectives 1 and 2, Objective 3 has not been consistently met. This is largely a limitation caused by lack of funding, rather than planning or sampling design. Indeed, the dual-frame sampling design proposed is ideally suited to rare organisms that consistently breed in the same locations (Haines and Pollock 1998, Hankin et al. 2019), and I think it is a good approach to achieve the stated objective while limiting potential downward bias in trends caused by visiting

known occupied locations (Pechmann et al. 1991, Fournier et al. 2019, Mentges et al. 2020). The justification for the prioritization of sampling and split panel design based on status and access in the event of partial resources makes good intuitive sense. In short, the sampling design is well-suited to the objective.

The implementation of the sampling design, however, is problematic. Because monitoring for this objective is implemented “only when funding and survey crews are available and after work associated with Objectives 1 and 2 has been completed” (Bennetts et al. 2013, p. 25) and boreal toads tend to breed earlier than other amphibian species and have a relatively short larval period, it seems that even when the sampling design for Objective 3 can be implemented, it is likely to be at a suboptimal time of year for many sites, though persistence of recently metamorphosed juvenile toads near natal wetlands could counter this potential limitation.

The field methods for Objective 3, which will typically employ dual observers or repeat visits within a sampling window appropriate to detect boreal toad breeding, are appropriate to achieve this objective. The sampling design and field methods might limit the utility of dynamic occupancy analysis because of small sample sizes and relatively few (5-12) list sites visited annually, but the information gathered on boreal toad populations will be useful regardless of the ability to fit statistical models.

In summary, the GRYN Amphibian Monitoring Protocol is meeting Objectives 1 and 2, but despite appropriate sampling design and field methods, funding is too limited to meet Objective 3. Continued development of remote sensing tools and recent developments in occupancy models, especially those made possible by Bayesian analysis, could extract additional information from the monitoring data, but even without these enhancements the first two objectives are met.

***Question 2: Is GRYN monitoring at the appropriate/most ecologically relevant scale? Are there ways to expand outside Grand Teton and Yellowstone national parks to the GYE? To the Western U.S.?***

The GRYN is monitoring amphibians and wetlands at the appropriate and most ecologically relevant scales. The multi-scale approach, whereby catchments are selected at random and all sites within catchments are visited, and if wet, surveyed, is an efficient design within this large landscape. Catchments are a necessary scale to consider because it would be impossible to survey as many sites if a completely random or spatially balanced selection of sites throughout the landscape were selected. Furthermore, sampling based on catchments likely captures the metapopulation structure of the target amphibian species. Sampling sites within catchments likely results in non-independence among sites within catchments, but both the conditional nature of occurrence at sites within catchments and potential variation in occupancy rates among catchments can be dealt with analytically as indicated above for Question 1.

The sampling of all sites within catchments is a good choice for several reasons. First, sites are the natural and appropriate scale to evaluate breeding amphibian occurrence. Second, it makes good sense logistically to sample all sites within a catchment because the greatest cost of conducting the surveys is often reaching the catchments themselves. Third, the dynamic nature of wetlands results in substantial variation in the suitability of sites for amphibian breeding among years. Because amphibians can breed in alternative proximate sites, monitoring all sites within a catchment reduces the likelihood of failing to detect such shifts in breeding location and interpreting them as extinction events within a catchment. Finally, to detect trends in occupancy requires that the sampling unit changes occupancy status during the study. This happens routinely at the site level, allowing estimation of extirpation and colonization probabilities and the influence of climate on wetland dynamics and in the influence of wetland dynamics on amphibian extirpation and colonization. Smaller spatial scales (e.g., plots within wetlands) would be irrelevant to amphibian occurrence dynamics at the scale of the GRYN, and larger spatial scales (e.g., catchments or watersheds) would undergo changes in occurrence status at a slower rate, providing less information about occupancy dynamics of amphibians except over much longer time scales.

Expanding amphibian and wetland monitoring outside Grand Teton and Yellowstone national parks to the Greater Yellowstone Ecosystem (GYE) and western U.S. is an admirable aspiration. The only feasible way to

do so that I can think of is to partner with other organizations. For example, existing or new collaborations with nearby universities (e.g., Idaho State University and Montana State University) and national forests within the GYE could expand sampling using similar protocols outside Yellowstone and Grand Teton national parks. Similar collaborations with other national parks, universities, national forests, and other federal agencies could expand this further to the western U.S. Continued collaboration with the U.S. Geological Survey's Amphibian Monitoring and Research Initiative (ARMI) is a natural fit because ARMI has similarly designed studies elsewhere in the western U.S. (e.g., Yosemite National Park and probably others in the Pacific Northwest and Rockies, though I am not sure how many of these are ongoing). Volunteer groups are another means of expanding outside the parks of the GRYN, although the time and expertise required to train volunteers can be limiting. Nonetheless, efforts like the Rocky Mountain Amphibian Project seem very promising.

***Question 3: Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program? To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of out of their sampling visits?***

The GRYN is using visual encounter surveys (VES) aided with a dipnet to detect (mostly larval) pond-breeding amphibians. This method is effective for pond-breeding amphibians, especially where the focus is on species occurrence rather than abundance (though VES is often the basis for capture-mark-recapture studies of adult amphibian abundance as well). The VES methods used in the protocol are therefore appropriate, and the high detection probabilities for the four amphibians in the GRYN demonstrate the effectiveness of the field methods (Gould et al. 2019, Patla and Ray 2020). Importantly, VES provide the opportunity to document other components of biodiversity that can be seen or heard, and extension of VES to include other species is straightforward. The primary concern here is that focusing on a small group of organisms, such as the pond-breeding amphibians and important interacting species (like fish, beaver, and aquatic plants as in this protocol; Bennetts et al. 2013), likely produces higher detection probabilities in less survey time than a complete inventory of all species would. Thus, focusing on amphibians and the interacting species that might affect their occurrence is appropriate for efficient field monitoring, and I think the protocol makes good use of field visits.

Recent developments in several techniques warrant consideration for improving survey efficiency, however. Acoustic surveys using automated recording units (ARUs) are one technique that can be used to detect some amphibians. The advantages of ARUs include recording data over longer periods of time (thus providing detailed data on phenology) and the collection of data on other species (insects and birds) that might be missed by visual encounter surveys. Despite these potential advantages, ARUs are likely not a viable option for amphibian monitoring in the GRYN for several reasons. First, not all amphibians produce breeding vocalizations. Salamanders produce none, and although boreal toads and Columbia spotted frogs vocalize, the vocalizations are not always indicative of breeding, breeding can occur without vocalization, and the calls are often not loud enough to be reliably detected on ARUs. Second, ARUs are susceptible to destruction by large mammals, which are abundant in the GRYN. Third, because ARUs must be deployed and retrieved (or at least data downloaded and batteries changed), there likely are few gains in efficiency by using automated recorders. The data files must be examined, either with automated recognizer files (trained by the researcher) or by visual and auditory examination of output. These additional tasks further limit the efficiency of ARUs in this system. Finally, ARUs are considered an installation in wilderness, so their widespread use in the GRYN is likely infeasible. I therefore do not recommend that ARUs be used as a replacement for VES, but they might provide interesting supplemental data in some circumstances.

Another recent development that warrants consideration for improving survey efficiency is the use of environmental DNA (eDNA) to detect amphibians and other aquatic organisms (Dejean et al. 2011, Goldberg et al. 2011, Jerde et al. 2011, Pilliod et al. 2013, Rees et al. 2014). Environmental DNA has been shown effective for amphibians under diverse circumstances and guidelines regarding the effects of variables on detection of eDNA are available and continue to be updated (Dejean et al. 2011, Pilliod et al. 2014, Laramie et al. 2015, Bedwell and Goldberg 2020a, 2020b). Indeed, eDNA is a rapidly developing technique that demonstrates great promise for many organisms. The advantages of using eDNA include the ability to survey for all aquatic species

recently at a site, including pathogens and other microbes (Mosher et al. 2017), from replicate water samples. Extracted DNA also can be archived for future analysis as assays are developed and interest shifts to organisms not considered at the time of sampling (e.g., Halstead et al. 2020). Another benefit of eDNA is that samples can be collected by a single observer, although safety considerations might nullify this advantage. Finally, the use of eDNA does not necessarily need to replace VES; rather, it could supplement VES, and one could use a single VES observer as one detection method and an eDNA sample as another to constitute the multiple surveys required to estimate detection probabilities in occupancy models. Although the potential advantages of eDNA sampling are numerous, limitations to the use of eDNA as a monitoring tool exist.

Several disadvantages to using eDNA as a primary sampling method in the GRYN Amphibian Monitoring Protocol exist. Perhaps most obvious is the increased cost. Collecting eDNA samples requires additional equipment, which increases costs. The amount of time necessary to collect eDNA samples at sites is likely not appreciably less than that required for VES, although advances in technology might make collecting eDNA samples more efficient in the future. In addition to field sampling costs, laboratory costs are incurred to extract DNA from filters and conduct quantitative PCR on the extracted DNA; recent estimates of these costs for similar work was \$15 per species per sample. Additional study is likely needed to determine the relationship between detection probabilities and filter type, pore size, volume of water filtered, etc. to optimize sampling design. Environmental DNA also cannot distinguish between life stages indicative of breeding and adult amphibians, though the former are likely to shed much more DNA into the water. Thus, although eDNA sampling is very promising, at this time I do not think it is a cost-effective alternative to VES for the four amphibian species covered under the protocol. As costs come down and if objectives change, however, eDNA could be a very powerful tool for examining a larger suite of aquatic organisms.

Another potential modification to field methods uses existing VES surveys, but includes the collection of time-to-detection data to model detection probabilities (Garrard et al. 2008). Briefly, time-to-detection occupancy models use information about how long it takes to detect the first individual of a species (broken down by life stage, if desired) *within* a survey to estimate the detection rate. The detection rate is then used to estimate the detection probability, given that the species occurs at the site and the duration of the survey. Thus, using time-to-detection methods can allow the use of occupancy models with only a single survey by a single observer. The method has been successfully used for diverse organisms, including plants (Garrard et al. 2008, 2015, McCarthy et al. 2013, Bornand et al. 2014), birds (Whittington et al. 2019, Henry et al. 2020), mammals (Medina-Romero et al. 2019), and amphibians (Halstead et al. 2018, 2021). The implementation of time-to-detection methods requires only small modifications to existing protocols; namely, the use of a stopwatch to determine how much survey time has elapsed between the initiation of the survey and the detection of the first individual of each life stage for each species. We have incorporated a survey timer into our Survey123 forms (see attachments) to help field personnel seamlessly collect these data, and the collection of time-to-detection data does not appreciably increase total survey time.

An important application of time-to-detection methods is the optimization of surveys (Henry et al. 2020, Halstead et al. 2021). By providing detailed information on detection rates and the variables that affect how much search time is required to achieve a desired detection probability, time-to-detection methods can help researchers plan how long a survey at a given site should be, given site and survey characteristics, to achieve good precision in occupancy estimates. Although time-to-detection methods can improve efficiency by limiting surveys to a single visit by a single observer, multiple time-to-detection surveys also can be accommodated (Henry et al. 2020), and these also can be optimized with regard to survey duration. All occupancy model extensions (multistate, multi-scale, multi-species, etc.) can be accomplished with time-to-detection methods.

Despite these benefits of time-to-detection methods, their utility is limited in some cases. Time-to-detection methods perform poorly for rare and cryptic species whose time to detection is long or that are detected at few sites (Henry et al. 2020, Halstead et al. 2021). These species often pose difficulties in traditional repeated measures occupancy designs as well. Based on published detection probabilities for amphibians in the GRYN (Gould et al. 2019, Patla and Ray 2020), I expect time-to-detection methods to perform well in this system. Pilot

studies (by incorporating time-to-detection data into existing survey protocols) are recommended to determine whether time-to-detection sampling using a single visit is a feasible alternative to independent double-observer sampling. Indeed, this is exactly what Halstead et al. (2021) did for a very similar sampling situation in Yosemite National Park, and we have since switched to single time-to-detection surveys.

The potential increased efficiency of time-to-detection surveys relative to independent double-observer surveys must be balanced with safety, however. In Yosemite National Park, we still send crews of two technicians to each watershed (= catchments in GRYN), and they remain in radio contact or, preferably, visual and auditory contact. In the GRYN, where grizzlies, wolves, and large ungulates can be a concern, it might not be feasible for crews to split up and survey different sites. Even if the potential efficiency of single-survey time-to-detection sampling cannot be realized, the information provided on the detection process is likely still worth the minimal cost of collecting time-to-detection data.

In summary, current field methods are among the best available, despite the development of alternative methods in the past decade or so. Acoustic monitoring is likely inappropriate for amphibians in this system. Environmental DNA sampling would be a boon to efficiently sample a broader swath of biodiversity, but its cost-effectiveness is a concern. Time-to-detection occupancy methods are likely feasible in this system, and they can readily be incorporated into existing surveys to test them prior to implementation. Even if single-survey time-to-detection sampling cannot be used in this system, the information it provides on the detection process is likely still worthwhile.

***Question 4: Are current analytical approaches described in the protocol and in Gould et al. (2019) capable of detecting statistically significant trends for all species involved? For detecting principal threats to amphibians? For understanding the drivers of wetland change?***

Detecting statistically significant trends in amphibian populations is notoriously difficult (Scherer and Tracey 2011, Greenberg et al. 2018). This pattern is especially true for abundance, but also holds for occupancy.

Broadly speaking, detecting trends depends on

1. Uncertainty in estimates of interest (in this case, occupancy) at each time point, which largely depends on the number of sites monitored and detection probabilities, and
2. Variability over time in the estimates of interest, which depends on inherent variability in the system and the duration of the study.

Thus, the ability to detect trends is less a function of the analytical approach, and more a function of the study design and constraints and variability inherent in the system. Species with greater uncertainty in annual estimates (e.g., tiger salamanders (Gould et al. 2012, Ray et al. 2016)) will be difficult to estimate trends for because of 1) above, whereas species whose occurrence varies drastically across years (e.g., boreal chorus frogs (Gould et al. 2012, Ray et al. 2016)) will be difficult to estimate trends for because of 2) above. Because of the small sample size for catchments ( $n = 30$ ), the protocol is likely inadequate to detect trends at the catchment level except over very long time scales. Trends at the site level are more appropriate to amphibian breeding and offer a larger sample of sites with which to work, increasing statistical power, but might still require a long study duration to detect a statistically significant trend. Despite these potential difficulties, there are approaches that might yield more meaningful results than seeking statistical significance.

A statistically significant trend is based on both the effect size (slope of the trend line) and variability over time. Thus, given the same amount of annual variation, a large trend can be detected more easily than a small trend and given the same amount of change, a trend is more easily detected when there is less annual variation. Statistical significance based on P-values can therefore fail to detect even large, biologically significant trends in noisy systems, because they focus on minimizing Type I errors (the probability of rejecting the null hypothesis [of no change] when it is true; Wade 2000). Instead, when concern is about detecting trends, one should also be concerned about Type II errors (declaring there is no trend when in fact one exists). I therefore encourage

the use of estimates of the magnitude of trends and probabilistic statements about trends rather than statistical significance, which focuses primarily on reducing Type I errors.

When it comes to estimating trends, Bayesian approaches have advantages over frequentist approaches (Wade 2000). First, Bayesian posterior distributions are probabilistic statements of certainty about the parameter of interest after accounting for prior information, the data collected, and the model used for estimation. Thus a Bayesian 95% credible interval of (-15% – 5%) means that we are 95% certain that the trend is between a 15% decrease and a 5% increase, whereas the same frequentist 95% confidence interval means that if we repeated the experiment 100 times, we would expect the true value of the trend to be within the interval about 95 times. The Bayesian intervals are much more intuitive to interpret and relate directly to the parameter of interest, rather than the experiment or observations themselves.

Another advantage of using a Bayesian approach to trend estimation is that one can express their confidence in a certain outcome. For example if one were interested in a threshold of a 10% annual decline to trigger some management action, the posterior distribution of the trend could be used to make a probabilistic statement about that trigger being exceeded (e.g., “The probability a 10% or greater annual decline occurred is 0.75”). These types of statements about thresholds are very useful for management because they can be formally included in a decision analysis (Wade 2000). Decision analysis, in turn, relates directly to my response to Question 5 regarding structured decision making for conservation.

The GRYN Amphibian Monitoring Protocol is designed with threat detection in mind. The monitoring of wetland status documents habitat change (including that caused by climate change; Ray et al. 2016, 2019), and reporting die-offs is an important component of disease monitoring. Auxiliary studies on *Batrachochytrium dendrobatidis* (Bd) and ranavirus are particularly relevant for detecting the threat to amphibian populations posed by disease. Because monitoring includes all amphibian species and also indicates whether fish are detected at each site, the protocol allows the detection of invasive animals relevant to amphibians. Invasive plants also are likely to be captured by the protocol, if not through early detection, then at least when their presence begins to dominate a site. Other threats, like anthropogenic habitat loss (exclusive of climate change), overharvest, and pollution are less likely in the GRYN. Overall, the protocol does a good job of detecting threats to amphibian populations.

As mentioned above, wetland change is well-addressed in the protocol. Indeed, the protocol has produced at least two peer-reviewed papers examining drivers of wetland change in the GRYN (Ray et al. 2016, 2019). The protocol is therefore performing very well in this regard and should only improve over time and with increased spatial resolution of remote sensing data.

**Question 5: How could GRYN increase the impact of their work? How can they work with parks and partners to transform knowledge into conservation action?**

The GRYN is already conducting impactful work and disseminating findings to appropriate audiences. In addition to annual reports, the publication of results of wetland and amphibian monitoring in well-respected peer-reviewed journals (e.g., Gould et al. 2012, 2019, Hossack et al. 2015, Ray et al. 2016, 2019) increases the reach and importance of the protocol and extends it beyond what is typical in most monitoring programs. Other outlets geared toward broader audiences, like the Vital Signs special issue of *Yellowstone Science*, make the monitoring of the health of the GYE accessible to the public. Continuing to collaborate with universities and other agencies is the most obvious way to leverage the GRYN Amphibian Monitoring Protocol to have a larger effect on knowledge, research, and management of wetlands and amphibians in the GRYN and beyond. In particular, continuing to contribute to large-scale syntheses (e.g., Adams et al. 2013, Grant et al. 2016, Muths et al. 2017) whenever possible is an excellent way to increase the impact of the work.

Of course, the impact of the work not only applies to publishing it for others to learn from, but also to improving management of the resource. Transforming knowledge into conservation action requires defining specific conservation objectives, and it is difficult to make specific recommendations in this regard. In general, however,

structured decision making (SDM) provides a framework for transforming knowledge into action, and is particularly useful when multiple objectives and uncertainty exist (Martin et al. 2009, Gregory et al. 2012). Structured decision making is a process that explicitly describes the context for the decision, the objectives used to identify and evaluate alternatives, the alternative actions under consideration, expected consequences of these actions, important uncertainties, key trade-offs, and adaptive learning to revise actions over time (Gregory et al. 2012). This framework has demonstrated utility in terms of water management for wildlife occupancy (Martin et al. 2009), including multistate occupancy approaches involving dynamic water availability (Martin et al. 2010) and optimizing management decisions for multiple species (Sauer et al. 2013). An example of amphibian conservation decisions in U.S. national parks based on SDM can be found in Grant et al. (2013), and SDM and SDM workshops seem to be gaining traction within the U.S. Department of the Interior. Knowledge gained from the GRYN Amphibian Monitoring Protocol would contribute to several steps in SDM and is essential for making informed conservation decisions.

***Question 6: How would you prioritize monitoring activities if budget/time/safety/personnel constraints restricted GRYN's ability to complete current activities?***

The implementation of the GRYN Amphibian Monitoring Protocol is already constrained by budget (which relates to time and personnel constraints). For example, the sampling design for Objective 3 has not been fully implemented, and low-priority watersheds were dropped from the sampling design for Objectives 1 and 2. If further constraints to budget, time, and personnel were encountered, my recommendation would be to do everything possible to maintain annual visits to all high- and medium-quality catchments currently monitored. If safety protocols allow, this could include single visits to sites using time-to-detection methods. Indeed, we have made this change to our similar survey protocol in Yosemite National Park for this reason. Depending on the time to reach catchments and sites, this change can result in substantial cost savings (with more savings incurred as travel time is reduced and makes up a smaller portion of the total time). Further reductions in funding would likely entail curtailing sampling at some selected catchments. In this case, I recommend dropping those for which access is least reliable (e.g., closures because of safety or animal concerns) or dropping a randomly selected catchments or catchments. The first approach, maintaining sampling of all long-term monitoring catchments, is much preferred.

If budgets remain constant (i.e., in line with general increases in salaries and operating expenses) and time-to-detection surveys are implemented or if budgets increase enough to allow additional work, then pursuing surveys of the list frame sites for Objective 3 would be a useful course of action. The dual-frame study design for this objective is well thought out, but suffers from lack of funding for implementation. Another potential avenue for expansion under an increased funding scenario would be to incorporate eDNA monitoring for its biodiversity benefit, though my recommendation would be to pursue the successful accomplishment of all three monitoring objectives first.

***Literature Cited***

- Adams, M. J., D. A. W. Miller, E. Muths, P. S. Corn, E. H. C. Grant, L. L. Bailey, G. M. Fellers, R. N. Fisher, W. J. Sadinski, H. Waddle, and S. C. Walls. 2013. Trends in Amphibian Occupancy in the United States. *PLoS ONE* 8:6–10.
- Bedwell, M. E., and C. S. Goldberg. 2020a. Spatial and temporal patterns of environmental DNA detection to inform sampling protocols in lentic and lotic systems. *Ecology and Evolution* 10:1602–1612.
- Bedwell, M. E., and C. S. Goldberg. 2020b. Spatial and temporal patterns of environmental DNA detection to inform sampling protocols in lentic and lotic systems. *Molecular Ecology Resources* 00:1–11.
- Bennetts, R., S. Corn, R. Daley, W. R. Gould, C. Jean, D. A. Patla, C. R. Peterson, and A. Ray. 2013. Cooperative amphibian monitoring protocol for the Greater Yellowstone Network: Narrative, version 1.0. Fort Collins, Colorado.

- Bornand, C. N., M. Kéry, L. Bueche, and M. Fischer. 2014. Hide-and-seek in vegetation: Time-to-detection is an efficient design for estimating detectability and occurrence. *Methods in Ecology and Evolution* 5:433–442.
- Dejean, T., A. Valentini, A. Duparc, S. Pellier-Cuit, F. Pompanon, P. Taberlet, and C. Miaud. 2011. Persistence of environmental DNA in freshwater ecosystems. *PLoS ONE* 6:8–11.
- Fournier, A. M. V., E. R. White, and S. B. Heard. 2019. Site-selection bias and apparent population declines in long-term studies. *Conservation Biology* 33:1370–1379.
- Garrard, G. E., S. A. Bekessy, M. A. McCarthy, and B. A. Wintle. 2008. When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. *Austral Ecology* 33:986–998.
- Garrard, G. E., S. A. Bekessy, M. A. McCarthy, and B. A. Wintle. 2015. Incorporating detectability of threatened species into environmental impact assessment. *Conservation Biology* 29:216–225.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2003. *Bayesian Data Analysis*. Second edition. CRC Press, Boca Raton, Florida, USA.
- Gelman, A., and J. Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. First edition. Cambridge University Press, Cambridge, UK.
- Gelman, A., J. Hwang, and A. Vehtari. 2014. Understanding predictive information criteria for Bayesian models. *Statistics and Computing* 24:997–1016.
- Goldberg, C. S., D. S. Pilliod, R. S. Arkle, and L. P. Waits. 2011. Molecular detection of vertebrates in stream water: A demonstration using rocky mountain tailed frogs and Idaho giant salamanders. *PLoS ONE* 6:e22746.
- Gould, W. R., D. A. Patla, R. Daley, P. S. Corn, B. R. Hossack, R. Bennetts, and C. R. Peterson. 2012. Estimating occupancy in large landscapes: Evaluation of amphibian monitoring in the greater yellowstone ecosystem. *Wetlands* 32:379–389.
- Gould, W. R., A. M. Ray, L. L. Bailey, D. Thoma, R. Daley, and K. Legg. 2019. Multistate occupancy modeling improves understanding of amphibian breeding dynamics in the Greater Yellowstone Area. *Ecological Applications* 29:1–13.
- Grant, E. H. C., D. A. W. Miller, B. R. Schmidt, M. J. Adams, S. M. Amburgey, T. Chambert, S. S. Cruickshank, R. N. Fisher, D. M. Green, B. R. Hossack, P. T. J. Johnson, M. B. Joseph, T. A. G. Rittenhouse, M. E. Ryan, J. H. Waddle, S. C. Walls, L. L. Bailey, G. M. Fellers, T. A. Gorman, A. M. Ray, D. S. Pilliod, S. J. Price, D. Saenz, W. Sadinski, and E. Muths. 2016. Quantitative evidence for the effects of multiple drivers on continental-scale amphibian declines. *Scientific Reports* 6:1–9.
- Grant, E. H. C., E. F. Zipkin, J. D. Nichols, and J. P. Campbell. 2013. A Strategy for Monitoring and Managing Declines in an Amphibian Community. *Conservation Biology* 27:1245–1253.
- Greenberg, C. H., S. J. Zarnoch, and J. D. Austin. 2018. Long term amphibian monitoring at wetlands lacks power to detect population trends. *Biological Conservation* 228:120–131.
- Gregory, R., L. Failing, M. Harstone, G. Long, T. McDaniels, and D. Ohlson. 2012. *Structured Decision Making: A Practical Guide to Environmental Management Choices*. Wiley-Blackwell, West Sussex, UK.
- Haines, D. E., and K. H. Pollock. 1998. Estimating the number of active and successful bald eagle nests: an application of the dual frame method. *Environmental and Ecological Statistics* 5:245–256.

- Halstead, B. J., C. S. Goldberg, R. B. Douglas, P. M. Kleeman, and D. W. Ulrich. 2020. Occurrence of a suite of stream-obligate amphibians in timberlands of Mendocino County, California, examined using environmental DNA. *Northwestern Naturalist* 101:194–209.
- Halstead, B. J., P. M. Kleeman, and J. P. Rose. 2018. Time-to-detection occupancy modeling: An efficient method for analyzing the occurrence of amphibians and reptiles. *Journal of Herpetology* 52:415–424.
- Halstead, B. J., J. P. Rose, and P. M. Kleeman. 2021. Time-to-detection occupancy methods: performance and utility for improving efficiency of surveys. *Ecological Applications* In Press.
- Hankin, D., M. S. Mohr, and K. B. Newman. 2019. *Sampling Theory: for the Ecological and Natural Resource Sciences*. Oxford University Press, Oxford, UK.
- Henry, D. A. W., A. T. K. Lee, and R. Altwegg. 2020. Can time-to-detection models with fewer survey replicates provide a robust alternative to traditional site-occupancy models? *Methods in Ecology and Evolution* 11:643–655.
- Hobbs, N. T., and M. B. Hooten. 2015. *Bayesian Models: A Statistical Primer for Ecologists*. Princeton University Press, Princeton, New Jersey, USA.
- Hooten, M. B., and N. T. Hobbs. 2015. A guide to Bayesian model selection for ecologists. *Ecological Monographs* 85:3–28.
- Hossack, B. R., W. R. Gould, D. A. Patla, E. Muths, R. Daley, K. Legg, and P. S. Corn. 2015. Trends in Rocky Mountain amphibians and the role of beaver as a keystone species. *Biological Conservation* 187:260–269.
- Hunt, S. D., J. C. Guzy, S. J. Price, B. J. Halstead, E. A. Eskew, and M. E. Dorcas. 2013. Responses of riparian reptile communities to damming and urbanization. *Biological Conservation* 157:277–284.
- Jerde, C. L., A. R. Mahon, W. L. Chadderton, and D. M. Lodge. 2011. “Sight-unseen” detection of rare aquatic species using environmental DNA. *Conservation Letters* 4:150–157.
- Kissel, A. M., M. Halabisky, R. D. Scherer, M. E. Ryan, and E. C. Hansen. 2020. Expanding wetland hydroperiod data via satellite imagery for ecological applications. *Frontiers in Ecology and the Environment*:1–7.
- Laramie, M. B., D. S. Pilliod, C. S. Goldberg, and K. M. Strickler. 2015. Environmental DNA Sampling Protocol—Filtering Water to Capture DNA from Aquatic Organisms: U.S. Geological Survey Techniques and Methods. Book 2 Collection of Environmental Data:1–15.
- Martin, J., S. Chamaillé-Jammes, J. D. Nichols, H. Fritz, J. E. Hines, C. J. Fonnesebeck, D. I. Mackenzie, and L. L. Bailey. 2010. Simultaneous modeling of habitat suitability, occupancy, and relative abundance: African elephants in Zimbabwe. *Ecological Applications* 20:1173–1182.
- Martin, J., M. C. Runge, J. D. Nichols, B. C. Lubow, and W. L. Kendall. 2009. Structured decision making as a conceptual framework to identify thresholds for conservation and management. *Ecological Applications* 19:1079–1090.
- McCarthy, M. A., J. L. Moore, W. K. Morris, K. M. Parris, G. E. Garrard, P. A. Vesk, L. Rumpff, K. M. Giljohann, J. S. Camac, S. S. Bau, T. Friend, B. Harrison, and B. Yue. 2013. The influence of abundance on detectability. *Oikos* 122:717–726.
- Medina-Romero, M., A. O’Reilly-Nugent, A. Davidson, J. Bray, E. Wandrag, B. Gruber, A. Lopez-Aldana, R. Palit, T. Reid, A. Adamack, R. Pietsch, C. Allen, R. Mac Nally, and R. P. Duncan. 2019. Effect of detection heterogeneity in occupancy-detection models: an experimental test of time-to-first-detection methods. *Ecography* 42:1514–1522.

- Mentges, A., S. A. Blowes, D. Hodapp, H. Hillebrand, and J. M. Chase. 2020. Effects of site-selection bias on estimates of biodiversity change. *Conservation Biology* 00:1–10.
- Mosher, B. A., K. P. Huyvaert, T. Chestnut, J. L. Kerby, J. D. Madison, and L. L. Bailey. 2017. Design- and model-based recommendations for detecting and quantifying an amphibian pathogen in environmental samples. *Ecology and Evolution* 7:10952–10962.
- Muths, E., T. Chambert, B. R. Schmidt, D. A. W. Miller, B. R. Hossack, P. Joly, O. Grolet, D. M. Green, D. S. Pilliod, M. Cheylan, R. N. Fisher, R. M. McCaffery, M. J. Adams, W. J. Palen, J. W. Arntzen, J. Garwood, G. Fellers, J. M. Thirion, A. Besnard, and E. H. C. Grant. 2017. Heterogeneous responses of temperate-zone amphibian populations to climate change complicates conservation planning. *Scientific Reports* 7:1–10.
- Nichols, J. D., L. L. Bailey, A. F. J. O’Connell, N. W. Talancy, E. H. C. Grant, A. T. Gilbert, E. M. Annand, T. P. Husband, and J. E. Hines. 2008. Multi-scale occupancy estimation and modelling using multiple detection methods. *Journal of Applied Ecology* 45:1321–1329.
- O’Hara, R. B., and M. J. Sillanpää. 2009. A review of Bayesian variable selection methods: What, how and which. *Bayesian Analysis* 4:85–118.
- Patla, D., and A. Ray. 2020. Greater Yellowstone Network amphibian and wetland monitoring: Status report for 2017 and 2018. Fort Collins, Colorado.
- Pechmann, J. H. K., D. E. Scott, R. D. Semlitsch, J. P. Caldwell, L. J. Vitt, and J. W. Gibbons. 1991. Declining amphibian populations: The problem of separating human impacts from natural fluctuations. *Science* 253:892–895.
- Pilliod, D. S., C. S. Goldberg, R. S. Arkle, and L. P. Waits. 2013. Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. *Canadian Journal of Fisheries and Aquatic Sciences* 70:1123–1130.
- Pilliod, D. S., C. S. Goldberg, R. S. Arkle, and L. P. Waits. 2014. Factors influencing detection of eDNA from a stream-dwelling amphibian. *Molecular Ecology Resources* 14:109–116.
- Ray, A. M., W. R. Gould, B. R. Hossack, A. J. Sepulveda, D. P. Thoma, D. A. Patla, R. Daley, and R. Al-Chokhachy. 2016. Influence of climate drivers on colonization and extinction dynamics of wetland-dependent species. *Ecosphere* 7:1–21.
- Ray, A. M., A. J. Sepulveda, K. M. Irvine, S. K. C. Wilmoth, D. P. Thoma, and D. A. Patla. 2019. Wetland drying linked to variations in snowmelt runoff across Grand Teton and Yellowstone national parks. *Science of the Total Environment* 666:1188–1197.
- Rees, H. C., B. C. Maddison, D. J. Middleditch, J. R. M. Patmore, and K. C. Gough. 2014. The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology* 51:n/a-n/a.
- Sauer, J. R., P. J. Blank, E. F. Zipkin, J. E. Fallon, and F. W. Fallon. 2013. Using multi-species occupancy models in structured decision making on managed lands. *Journal of Wildlife Management* 77:117–127.
- Scherer, R. D., and J. A. Tracey. 2011. A power analysis for the use of counts of egg masses to monitor Wood Frog (*Lithobates sylvaticus*) populations. *Herpetological Conservation and Biology* 6:81–90.
- Tenan, S., R. B. O’Hara, I. Hendriks, and G. Tavecchia. 2014. Bayesian model selection: The steepest mountain to climb. *Ecological Modelling* 283:62–69.
- Vehtari, A., A. Gelman, and J. Gabry. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* 27:1413–1432.

- Wade, P. R. 2000. Bayesian methods in conservation biology. *Conservation Biology* 14:1308–1316.
- Whittington, J., B. Shepherd, A. Forshner, J. St-Amand, J. L. Greenwood, C. S. Gillies, B. Johnston, R. Owchar, D. Petersen, and J. Kimo Rogala. 2019. Landbird trends in protected areas using time-to-event occupancy models. *Ecosphere* 10:e02946.
- Zipkin, E. F., J. Andrew Royle, D. K. Dawson, and S. Bates. 2010. Multi-species occurrence models to evaluate the effects of conservation and management actions. *Biological Conservation* 143:479–484.

## Paige Howell

### *Greater Yellowstone Network Amphibian Monitoring Protocol Review*

Prepared For: Michael Bozek (michael\_bozek@nps.gov)

Prepared By: Paige Howell, USGS Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel, MD 20708

Date: January 15, 2021

***Question 1: Is GRYN meeting objectives (Table 1) stated in the current amphibian protocol? If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these.***

Overall, the current GRYN amphibian protocol is sound and allows for estimating both the number of wetlands that are suitable for amphibian breeding and the proportion of wetland sites that are used for breeding by the more common amphibian species. The sampling protocol allows for dealing with detectability (Bennetts et al. 2013) and the use of a multistate models (Gould et al. 2019) is a clever way to estimate both wetland dynamics and occupancy dynamics. Below I have a few suggestions for modifications.

- A. Multi-species occupancy models in a Bayesian framework – Without making any changes to the sampling protocol, I would suggest analyzing data for all four species together in a multi-species occupancy model. In this way, we can share information across species and potentially improve precision around parameter estimates. Here I provide an example that outlines this type of modeling framework (Zipkin et al. 2009).
- B. Community-science data / non-random site selection – The stratified random site selection that was originally used to select sites for sampling is a good idea, particularly if design-based statistical methods are used to analyze the data. However, given that we can make use of model-based statistical methods, it is less imperative that we have random site selection. Rather, it is important that we a) span the entire covariate space (e.g., wetland hydroperiod) and b) achieve high enough sample sizes (i.e., for a dynamic occupancy model, extinction and colonization events). So, for species like the western toad where it has been difficult to obtain precise estimates, I would suggest adding in additional sites that have a higher probability of being occupied by that species. So long as the sampling protocol still allows for making inference at a larger scale (i.e., the GYE), it is reasonable to have some sites that are not randomly selected. To avoid having to drop some sites from the survey, a community-science effort would be one way to add sites. If existing community-science databases have decent location information available there may already be some data that can be integrated into existing analyses. Otherwise, engaging members of the public to take part in community-science efforts would be another potential avenue for data collection. Not only would this effort add to the data available for analyses, but it would provide a platform for education and outreach. Integrating community-science with structured survey data is an active area of research, and so there is great potential to use these data with the structured survey data in an occupancy modeling framework (Fletcher et al. 2016).
- C. Comment on time-to-detection and eDNA – Single time-to-detection (TTD) surveys would potentially cut down on the time spent in the field. Although for safety reasons (i.e., needing two observers to be within eyesight) I'm a little skeptical that it would substantially cut down on cost/time. Additionally, evidence suggests that while occupancy models fit to single survey TTD data perform well for widespread / conspicuous species, they do not perform well for rare / inconspicuous species (Henry et al. 2020, Halstead et al. 2018 – but see their commentary on informative priors). I would however suggest collecting the TTD data during each replicate detection – non-detection (DND) survey for two reasons. First, there is some evidence that repeat TTD surveys yield better results than replicate DND surveys (Bornand et al. 2014). Second, in the event that monitoring is disrupted in a given year (e.g., budget shortfall) and only a single survey could be conducted at each site, a single TTD survey is better than a single DND survey (Henry et al. 2020), and could be used in an integrated model with the historical DND data.

With eDNA, I do not see this approach as being less costly than the current visual encounter surveys, given the cost of analyzing samples in the lab. It would require a long-term agreement with a lab, and I'm not sure how easy it is to guarantee a price and long-standing, stable cooperation between the lab and GRYN. Second, eDNA isn't really measuring the state of the system at the same temporal scale as the visual encounter surveys. Depending on how quickly the DNA degrades at a site, you could be measuring a single transient individual that was present during the breeding season but didn't actually colonize and breed in a site. Given the objectives in Table 1 (i.e., estimate proportion of sites with breeding), I think the visual surveys should be preferred. Third, my understanding is that in general you need more samples than you would think to achieve a higher detection probability with eDNA, relative to the visual surveys.

***Question 2. Is GRYN monitoring at the appropriate/most ecologically relevant scale? Are there ways to expand outside Grand Teton and Yellowstone national parks to the GYE? To the Western U.S.?***

The monitoring protocol is sound and currently allows for making inference at the site and catchment scale. In terms of expanding beyond Grand Teton and Yellowstone, my recommendation would be to either a) shift sampling effort in a subset of years to include additional sites outside the current monitoring program b) connect with managers or researchers already sampling in other areas to share data or c) develop a community-science program to incentivize members of the public to help collect these data at additional sites within the GYE.

For an example of an occupancy analysis of a larger study area, than what is sampled each year (Howell et al. 2018, Howell et al. 2020). Note, these frameworks are spatially-explicit and assume all potential breeding habitat is known. So, the exact model formulation for the GRYN would be a spatially-implicit version of this analysis, unless all potential breeding habitat is known OR could be inferred from additional remote-sensing data.

For an example of integrating community-science data into a model of species distributions (Fletcher et al. 2016).

***Question 3. Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program? To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of out of their sampling visits?***

Overall, I think this is a very sound protocol and allows for detecting most amphibian species as well as modeling wetland dynamics. I would suggest collecting TTD data during the repeated visual encounter surveys. It shouldn't affect the cost/time of surveys but could be integrating into existing occupancy models to potentially yield better results than the current DND surveys alone (Bornand et al. 2014). As suggested in questions 1 and 2, community-science data could expand the scope of inference for the GRYN and serve a dual purpose of outreach and education.

***Question 4. Are current analytical approaches described in the protocol and in Gould et al. (2019) capable of detecting statistically significant trends for all species involved? For detecting principal threats to amphibians? For understanding the drivers of wetland change?***

The current analytical approaches are capable of detecting statistically significant trends for most species, including detecting the principal threats and understanding drivers of wetland change. However, for species like the western toad, the current sampling / statistical design does not provide precise estimates of occupancy dynamics. As noted in my responses to the preceding questions, I think a combination of a) multi-species occupancy modeling in a Bayesian framework, and either b) site selection targeting likely western toad breeding habitat or c) community-science data will potentially allow for estimating trends for this and other rare species.

***Question 5. How could GRYN increase the impact of their work? How can they work with parks and partners to transform knowledge into conservation action?***

To expand the scope of inference regarding amphibian occupancy dynamics and wetland breeding habitat, my recommendation would be to either a) shift sampling effort in a subset of years to include additional sites outside

the current monitoring program's spatial scale b) connect with managers or researchers already sampling in other areas to share data or c) develop a community-science program to incentivize members of the public to help collect these data at additional sites within the GYE.

Developing a community-science monitoring program, or plugging in with existing community-science efforts, would be a great way to a) expand the spatial scale over which data are collected and b) provide a mechanism to increase public engagement with the GRYN. I am not very familiar with applied social science research, but I could see this avenue of monitoring as potentially very fruitful for improving our understanding of how engaging the public in research changes public perception of the value of natural resources.

***Question 6. How would you prioritize monitoring activities if budget/time/safety/personnel constraints restricted GRYN's ability to complete current activities?***

In the event of a shortfall in budget/time or some other constraint to monitoring, I would suggest two potential options.

First, is prioritizing sampling (using the current visual encounter survey protocol) at a smaller subset of sites that are a) part of the historical database and b) for which it is most important to capture inter-annual changes in occupancy or in wetland dynamics. It would be straightforward, particularly in a Bayesian framework, to analyze the entire time series of data, despite year-to-year changes in the number of sites surveyed (e.g., Howell et al. 2018).

An alternative would be to conduct single TTD surveys at all or a larger subset of sites than would be possible with the repeat visual encounter surveys. This scenario may be less feasible, since safety concerns necessitates two observers being within eyesight of one another.

Finally, another benefit to developing a community-science survey, or integrating existing community-science data, is that it would help to provide at least some amount of survey data in situations where ALL normal monitoring activities are disrupted (e.g., budget shortfalls, pandemics, etc.).

***Literature Cited***

- Bennetts, R., S. Corn, R. Daley, W. R. Gould, C. Jean, D. A. Patla, C. R. Peterson, and A. Ray. 2013. Cooperative amphibian monitoring protocol for the Greater Yellowstone Network: Narrative, version 1.0. Natural Resource Report NPS/GRYN/NRR—2013/654. National Park Service, Fort Collins, Colorado.
- Bornand, Christophe N., Marc Kéry, Lena Bueche, and Markus Fischer. "Hide-and-seek in vegetation: time-to-detection is an efficient design for estimating detectability and occurrence." *Methods in Ecology and Evolution* 5, no. 5 (2014): 433-442.
- Fletcher, Robert J., Robert A. McCleery, Daniel U. Greene, and Courtney A. Tye. "Integrated models that unite local and regional data reveal larger-scale environmental relationships and improve predictions of species distributions." *Landscape Ecology* 31, no. 6 (2016): 1369-1382.
- Gould, W., A. Ray, L. Bailey, D. Thoma, R. Daley, and K. Legg. 2019. Multistate occupancy modeling improves understanding of amphibian breeding dynamics in the Greater Yellowstone Area. *Ecological Applications* 29(1):e01825.
- Halstead, B. J., Kleeman, P. M., & Rose, J. P. (2018). Time-to-detection occupancy modeling: An efficient method for analyzing the occurrence of amphibians and reptiles. *Journal of Herpetology*, 52(4), 416–425.
- Henry, Dominic AW, Alan TK Lee, and Res Altwegg. "Can time-to-detection models with fewer survey replicates provide a robust alternative to traditional site-occupancy models?." *Methods in Ecology and Evolution* 11, no. 5 (2020): 643-655.

- Howell, Paige E., Erin Muths, Blake R. Hossack, Brent H. Sigafus, and Richard B. Chandler. "Increasing connectivity between metapopulation ecology and landscape ecology." *Ecology* 99, no. 5 (2018): 1119-1128.
- Howell, Paige E., Blake R. Hossack, Erin Muths, Brent H. Sigafus, Ann Chenevert-Steffler, and Richard B. Chandler. "A statistical forecasting approach to metapopulation viability analysis." *Ecological Applications* 30, no. 2 (2020): e02038.
- Zipkin, Elise F., Amielle DeWan, and J. Andrew Royle. "Impacts of forest fragmentation on species richness: a hierarchical approach to community modelling." *Journal of Applied Ecology* 46.4 (2009): 815-822.

## Dusty Perkins

### *Greater Yellowstone Network Amphibian Monitoring Protocol Review*

**Prepared For:** Michael Bozek (michael\_bozek@nps.gov)

**Prepared By:** Dusty Perkins, Northern Colorado Plateau Network, Inventory and Monitoring Program, Crested Butte, Colorado

**Date:** December 22, 2020

### **Objectives**

1. Estimate the proportion of catchments and wetland sites used for breeding by each of the four common, native amphibian species annually, and estimate the rate at which their use is changing over time.
2. Determine the total number of wetlands within sampled catchments that are suitable for amphibian breeding (i.e., have standing water during the breeding season) annually
3. For Western Toads, estimate the proportion of previously identified breeding areas that are used annually, and estimate the rate at which their use may be changing over time.

Reviewers are asked to provide responses to the following questions about current monitoring activities:

***Question 1: Is GRYN meeting objectives (Table 1) stated in the current amphibian protocol? If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these.***

GRYN is meeting objective 1 for three of the four species it monitors. It is not meeting objective 1 for western toads due to low occurrence rates, but this information is still valuable to collect. I am very impressed with their analyses and wide range of publications from reports to peer-reviewed journals to more managerial and public facing journals such as Park Science and Yellowstone Science.

GRYN is meeting objective 2. However, I find the objective is an odd one for a long-term monitoring program. I think looking at wetland desiccation over long-time periods could be just as valuable as the amphibian presence data. I recommend revising the objective 2 to incorporate a long-term component. The network is already doing this, so the objective should reflect that.

It is my understanding that GRYN has made a conscious decision to not address objective 3 due to financial considerations. I would recommend they pursue Inventories 2.0 money to try and address this objective over a couple of years. In an ideal world this sampling would include at least one wet year. Sampling in dry years for western toads is probably not worth the investment. I do think you should keep this objective in the protocol but continuing to acknowledge that more funding is needed if this is to be accomplished. I also think you should “save up” when possible and do a lot of toad inventories at a time in a single year when conditions are good. Past GRYN documents have inferred they would do this “when they could”, I think it should be a bit more strategic than that. I think you would be better served to adding more catchments when you could to potentially increase your tiger salamanders sampled then targeting toad locations.

I found this statement in your protocol: “Recognizing these limitations, boreal toad monitoring efforts described here will only be implemented once all work needed to address Objectives 1 and 2 have been completed for all amphibians within a given year.” Does this mean you should eliminate objective 3 now if you still agree with this statement? You currently are very borderline as to whether you have enough samples for objective 1 for tiger salamanders, so this would imply you need more sites added before you go after objective 3. I am not advocating one way or the other, but thought the quote was worth adding for discussion with park management.

***Question 2: Is GRYN monitoring at the appropriate/most ecologically relevant scale? Are there ways to expand outside Grand Teton and Yellowstone national parks to the GYE? To the Western U.S.?***

GRYN is monitoring at the appropriate scale. There was much discussion about the exclusion of low-quality catchments from the monitoring program several years ago. I also was concerned about this because low quality areas can be some of the more sensitive wetlands and the first ones to blink out should a species start to decline. On one extreme, doing amphibian surveys for presence in only high quality wetlands could lead you to a situation where you do not detect a decline until they start to decline in the best wetlands, long-after they have gone extinct in low and moderate quality wetlands. However, upon further discussion, I think most of this issue is a naming issue. Due to the large number of absences in these low quality catchments in a wide range of hydrologic years, I think these wetlands might be better thought of as non-habitat for amphibians or as very transient wetlands that are only occupied in very wet years when populations are high and likely do not contribute very much to the overall population. It sounds like using medium and high-quality catchments is still capturing most of the amphibian population – one could argue you still have inference to the GRYN amphibian population. In addition, due to the high number of absences in most years, it would be very problematic statistically to include these low quality catchments. There is a large gradient of wetlands from ones that have fish in them that are not habitat but might occasionally have a frog hop in them to the other extreme of an artificial kiddie pool (yes, I have seen amphibians in them). Obviously, you don't include all of these in your sample. Lastly, GRYN is still monitoring low quality wetlands in medium and high-quality catchments. So, should this low-quality data ever be needed they still have some of these data.

***Question 3: Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program? To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of out of their sampling visits?***

GRYN is using the best available methods. Acoustic loggers would miss tiger salamanders and would require two annual visits to all sampled wetlands. More substantial efforts that tried to get at abundance would require more time and more precise timing to hit amphibians at their peak. Visual encounter surveys are broadly used and the use of two independent observers is excellent for methods and for safety in grizzly country.

One potential addition would be to do more eDNA work if money could be found for lab analysis. This could potentially increase detections, including amphibian disease, and is well suited for a broader time period to sample since the DNA can last 3-4 weeks after a species DNA has been present. However, even if money would not be expected to be found for several years for eDNA lab work, GRYN should still collect water samples and put them in storage so that they could be examined sometime in the future as the technology gets better, maybe semi-quantitative, and perhaps cheaper. Imagine how much we wished someone in the 1950s had taken some water samples when western toads were abundant and we could analyze them now. Storage of samples would also have the additional benefit to expand the monitoring to other species like birds in the future. If storage was still an issue and a cost, maybe you prioritize taking samples in good wet years, or on some periodic 3-5-year time frame. If you did have money to incorporate eDNA lab work now, I would still recommend that you take an additional sample that can be stored for all the advances in technology that are likely. In watching Dr. Murphy's talk and a subsequent conversation, it sounded like there may even have been advances since you attempted this with so-so results a few years ago.

***Question 4: Are current analytical approaches described in the protocol and in Gould et al. (2019) capable of detecting statistically significant trends for all species involved? For detecting principal threats to amphibians? For understanding the drivers of wetland change?***

Biostatistics is not my strong point, but it is for Dr. Gould and GRYN has used him extensively. GRYN should be applauded for using Dr. Gould to advance their analytical techniques, provide more powerful conclusions and using him almost every step of the way when developing the protocol and changing methods (i.e. dropping of low-quality catchments).

I like the identification of amphibian hotspots where all 4 species are present. I think this could be extended to wetland and catchment “hotspots”. Are there wetlands and catchments over time that are always wet or almost always wet? These could be particularly important to management as climate change continues. These persistent wetlands could be key to re-colonization of less persistent wetlands in wet years. Additional nearest neighbor type analysis could be done to identify wetlands that are really important to the species. For example, one persistent wetland that is surrounded by moderate wetlands and does not have any other persistent wetlands within y km (some number where colonization by these species is likely) could be much more important to amphibian metapopulations than persistent wetlands that have several other persistent wetlands in close proximity.

***Question 5: How could GRYN increase the impact of their work? How can they work with parks and partners to transform knowledge into conservation action?***

GRYN has done an excellent job in this department. This work has put them in a spot where they are looked to as leaders in regional amphibian conservation. It was clear in the 2-day symposium that GRYN staff and partners were looked up to by other amphibian people in the region. This will only grow as their long-term study grows and more information is gleaned from these data. So, my suggestions should be taken with a grain of salt, they are already fulfilling this goal.

GRYN amphibians should be held up by WASO leaders as an example to follow. GRYN can play a role in this by holding a webinar on the review and their achievements. They can share lessons learned, but they should also not be modest in their achievements – it will inspire other networks to try and catch them. The combination of a long-term data set, holding one of the first science reviews, excellent analytical techniques, publications in peer-reviewed science and managerial journals and NRR regular reporting, and being a leader in regional amphibian conservation is the gold standard. So, the conservation action would be to help other networks achieve similar success. If all networks were operating like this nationwide, we would be legendary.

As with other suggestions elsewhere about the value of certain wetlands and amphibian hotspots, I think this monitoring protocol has some real value in thinking about conservation planning and climate change decisions these parks are facing. I would push on with that and make sure your data is known and inserted in these climate decisions. Mike Tercek and David Thoma had some good ideas in their presentation. So perhaps this is just a reiteration of what you are doing, but I would encourage you to be bold in this arena. It is unlikely that there are other sources of data that are good as yours in helping to tackle climate change issues. Perhaps the network could think about replicating the grasslands and PJ workshops NCPN has done (David knows all about them) in a wetland application?

***Question 6: How would you prioritize monitoring activities if budget/time/safety/personnel constraints restricted GRYN's ability to complete current activities?***

I think this can only be considered with other monitoring objectives of GRYN. I would not cut very much at all from the amphibian monitoring program. I think it is close to the bare minimum given the rarity of these species and the variability of occurrence rates dependent on water availability. If cuts were to be made at GRYN, I would strongly recommend that GRYN think about cutting whole monitoring programs, rather than a death by a thousand cuts across all programs. So perhaps they keep amphibians whole while cutting another protocol entirely. On the flip side if your other protocols are equally outstanding, maybe it is amphibians that should be cut. In decisions like these, I think it is worth asking questions like “If you knew you could not do toads and knew the occurrence rate you have for tiger salamanders, would you still have picked amphibians?”. In general, I think IMD will do far better by doing a good job at a few things as opposed to a mediocre job at several things.

If you had more money, I would add more catchments first in the hopes that it would help with tiger salamanders and maybe western toads. Adding eDNA would be a close second, maybe a tie.

## Other thoughts

- GRYN's publication record is impressive. They have excellent judgement in knowing when to publish data that is important to wider audiences while also continuing NRR reports that are so important to park managers and to show what a monitoring program is doing on a regional and national level.
- Other data to collect that would be fast. Any thought to adding the top 5 most dangerous invasive riparian plants? I don't know the system well, but are other invasive critters that could become a problem, like mollusks, crayfish? Species that could be total game changers to the ecosystem would be worth noting when seen, even if the absence information might not be believable.
- Photos are great, very powerful. However, I think the SOPs could be tightened up a little. A small change in direction or location can make skeptics think you are trying to manipulate the photo. There were several examples of repeat photos in the presentations, I assume you picked some of the best ones. Even so, there were some where it was not entirely clear that it was the same location. Anything you can do to tighten that up while not spending too much time doing it could really help. There are some people that specialize in this – taking all day to try and get the light just right for a repeat photo, I am not suggesting that, it still needs to be quick. Some new cameras now have direction built into the lenses so you can be sure to take the photo is taken in the same compass direction each year. Even just emphasizing in training how important photos are, that they are not just something to run and do quickly before you leave the site can improve the quality. In our work with water rights on big rivers, a WRD hydrologist told me that while they needed all the details about data on plants, hydrology, and geomorphology, what won the Black Canyon of the Gunnison NP water right in court was repeat photos that showed what a river did in high flows to remove colluvial matter that was building up in the riverbed.
- Adding time to detection to GRYN amphibian presence seems to be a very minor but worthwhile tweak to the SOP and could be useful in expanding this work in GRYN or elsewhere.
- I liked the talk by Wendy Estes-Zumpf on RMAP with their volunteer program. Is this a possibility in GYE to expand the monitoring both within the parks and outside the parks in the GYE?
- In the discussion, Chuck Peterson suggested adding a developed site. I disagree. GYE is one of the largest protected intact ecosystems in North America. You should focus on that and make sure we remember how rare that is. I can understand if one is always in the GYE and one sees other projects in other parts of the country where development has affected amphibians and how that might be intriguing. But until you can get objectives 1-3, stay the course. Let others in more damaged ecosystems study the effects of development on amphibians.
- I am a little unsure of the value of the site level analysis. It would seem that the monitoring program was set up to evaluate by catchments and that is your unit of measure. So, by analyzing by sites, the sites are not independent – but I would also defer to the likes of Dr. Gould on such analyses and their appropriateness.
- There is mention of the value of GRYN's monitoring being part of the ARMI program. However, it seemed like GRYN is the only one really doing the ARMI protocol in the GYE? Maybe the RAMP protocol uses the same methods? If it is, have you all investigated combining your data with theirs for both statistical and ecological reasons?
- Be careful of too much interp/education. We were put here to do science. We want to contribute to interp/education – but we should not be leading these efforts. We have a world class interpretation division for that. We are not yet a world class science organization within NPS. If early IMD leaders had wanted a heavy interpretation component they would have made the program that way. They did not.
- I know that I and others have suggested things to add. But remember, you are doing a very important piece – long-term monitoring. This is one of the hardest pieces to do and hardest to fund. Don't add too much. Imagine years like the sequester, imagine a year with a horrible crew, what can you still do to continue with minimal effort to make sure this data set continues. If you do nothing else, even if you don't publish, if this program is still collecting data in its present form in 10 years at the next review, that alone is a success.

## Stephen Spear

### *Greater Yellowstone Network Amphibian Monitoring Protocol Review*

Prepared For: Michael Bozek (michael\_bozek@nps.gov)

Prepared By: Stephen F. Spear, PhD, Research Biologist, U.S. Geological Survey, Upper Midwest Environmental Sciences Center, 2630 Fanta Reed Road, La Crosse, Wisconsin 54603 (Email: sfspear@usgs.gov)

**Question 1: Is GRYN meeting objectives (Table 1) stated in the current amphibian protocol? If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these.**

Objectives from Bennetts et al. (2013) Cooperative amphibian monitoring protocol for the Greater Yellowstone Network; <https://irma.nps.gov/DataStore/Reference/Profile/2194571>:

1. Estimate the proportion of catchments and wetland sites used for breeding by each of the four common, native amphibian species annually, and estimate the rate at which their use is changing over time.
2. Determine the total number of wetlands within sampled catchments that are suitable for amphibian breeding (i.e., have standing water during the breeding season) annually
3. For Western Toads, estimate the proportion of previously identified breeding areas that are used annually, and estimate the rate at which their use may be changing over time.

The current monitoring protocol is aimed at complete wetland sampling in a subset of catchments that represent high and medium quality amphibian habitat based on the amount of suitable wetland habitat. The same catchments are sampled each consecutive year and assessed for larval amphibians using visual and dipnet surveying. Detection is determined using a two observer protocol with single visits to each site. If a wetland is not suitable for amphibians (i.e. dry), this is recorded as well. The survey protocol is well suited to meet Objectives 1 and 2 for each catchment surveyed as there is complete census of all wetlands within each sampled catchment, and each wetland is assessed for suitability each year. It is less clear if the protocol can provide reliable inference to all GRYN catchments. First, catchments that are considered low quality (based on number of NWI wetlands) have been excluded, and so results cannot be applied to the majority of catchments in the parks – these are implicitly assumed to not provide consistent amphibian habitat. However, beyond the exclusion of the low quality sites, I think it is also unclear if the monitoring protocol can be assumed to represent all high and medium quality catchments. There are a total of 135 high quality catchments and 990 medium catchments, so the current protocol represents ~11% of high quality and ~2% of medium quality catchments. Surveying a higher percentage of catchments would take a significant increase in effort and is likely not feasible given current funding levels. The investigators have attempted to maximize representation of sampled catchments as much as possible by spreading out units among different drainages of the parks that differ in elevation and landscape features, and selecting catchments within those regions randomly. While stratification by drainage is important and a strength of the protocol, a disadvantage is that it dilutes sample size even further as there are only 2-3 catchments in each category per region.

Another question I have is what the long-term experimental unit for monitoring is going to be. The protocol is set up for both catchments and wetlands, but the most recent analysis (Gould et al. 2019) does not appear to have included either catchment or major drainage as a factor in the multi-state models, unless I missed it in the methods/results. If that is the case, are individual wetland sites independent of catchment or region? Has this been tested? A better understanding of whether amphibian breeding dynamics are structured by catchment would be really useful going forward as to whether the current design could be flexible to addition of catchments or reduction of wetland sites within a catchment.

Another major concern is that the current protocol is likely at best effective for monitoring only two of the four species, specifically chorus frogs and spotted frogs. This is best illustrated by Gould et al. (2019) which was able

to model detection and occupancy for the two species and demonstrated a strong relationship between climate and amphibian breeding. Beyond yearly variation in wetland availability, there was not any apparent increase or decrease in breeding in these two species. Therefore, while it appears that Objective 1 is met for chorus frogs and spotted frogs, it has not been met for tiger salamanders or western toads, as number of observations is not sufficient to model either of these species. If the goal is to meet the objective for these two species, additional catchments will need to be added that are already known to contain a sufficient proportion of tiger salamander or western toad sites, which may require both compilation of previous survey data and additional targeted surveys to locate new sites. Both species are nonrandomly associated with park regions, especially tiger salamanders which are located throughout the northern range, Yellowstone, and southern Grand Teton drainages. Although preliminary surveys would be needed, a straightforward way to estimate trends in tiger salamander would be to add more catchments or wetlands in these regions, even if it led to unbalanced regional designs. This could be accomplished by prioritizing more resources to additional surveys. If that is not possible, then one strategy to reallocate sites would be to run multi-state models on subsamples of the existing data set to better understand how sensitive the results would be to loss of sites. Losing catchments would be difficult given the already low sample size but reducing the number of wetlands sampled per catchment could be a possibility, as well as adding catchments adjacent to current catchments to increase sample size while limiting the need for increased travel or personnel. I believe this approach could only work if reducing sites per catchment didn't influence statistical power for trends for chorus frogs and spotted frogs, and if catchments within a geographical region can be considered independent such that surveying geographically adjacent catchments wouldn't violate independence assumptions. Assessing whether these conditions are met would take further analysis and pilot surveys. Alternatively, the objective could be modified to solely focus on chorus frogs and spotted frogs. I do not think this is ideal because each of the four amphibian species have different ecological relationships to a certain extent, and I don't think we can assume that trends in one species would necessarily apply to all four.

Objective 2 is the most clear-cut of the three and the current approach is very successful for this objective. Each wetland in a catchment is surveyed, so the state of suitable wetlands for amphibian breeding is easily assessed, and very important for modeling amphibian trends. Given some of the research presented at the conference regarding modeling hydroperiod and wetland inundation using remote sensing, I think this objective could be expanded even further to include catchments not regularly sampled if they can be modeled reliably using remote sensing. Furthermore, if additional analyses demonstrate that wetland state is the primary driver behind amphibian year-to-year trends, the hydrological research could allow the park to predict trends parkwide beyond the sampled catchments. I would certainly recommend the monitoring program continue to invest in the hydrological research across the park.

As acknowledged by the monitoring team, objective 3 has not been met due to funding limitations, and it is certainly the most difficult of the three objectives due to the sampling effort and personnel required. If a pair of surveyors were dedicated to visiting known toad sites, then the survey portion of this objective could be accomplished, although it is not clear to me if trends could be reliably estimated using the same statistical tools as used for the chorus frog and spotted frog. The sample size would be much smaller, but a larger consideration is the complexity of toad breeding sites. One of the advantages of the other three species is that the majority of their breeding sites within the park are isolated, ephemeral wetlands that are easily monitored. In the case of western toads, breeding sites are often riverine backwaters, oxbows, or disturbed sites. These sites often don't have easily discrete boundaries and by their nature are often quite dynamic over time. In the monitoring plan (Bennetts et al. 2013), toad sites are defined as complexes and that is likely the best way to define as presence or absence. However, it would be much more difficult to determine if site is suitable or not since there would likely be some water present at many of these sites in most years. If sufficient resources were available to address this objective, I would recommend that a detailed and customized protocol be developed by toad experts to allow surveys to determine extent of available toad habitat and/or determine if a multistate analysis can be applied to western toad, as well as power analysis to determine what minimum sample size would likely be.

***Question 2: Is GRYN monitoring at the appropriate/most ecologically relevant scale? Are there ways to expand outside Grand Teton and Yellowstone national parks to the GYE? To the Western U.S.?***

The grain of the monitoring is ecologically relevant, as the base survey unit (wetland) is the unit at which amphibians breed, and therefore should be the finest scale at which inferences can be made. This is also the scale at which covariates were estimated for the multi-state occupancy models described in Gould et al. (2019). There are two additional scales at which monitoring is structured (catchment and major drainage). These two scales have an ecological basis being both based on drainage basins. While catchment and major drainage were critical in site selection, the most recent analysis does not really address these scales beyond a summary view of occupied catchments and drainages, as there are not covariates associated with the two broader scales. For instance, Wright et al. (2020) examined park-level covariates with amphibian species richness for monitoring in the national capital region, and found an association with these metrics related to park area and isolation. This analysis is not an exact template for looking at multi-scale effects within GRYN as the amphibian species richness is lower within GRYN and there does not exist the fragmentation present in the urbanized national capital region. However, ecological characteristics associated with different drainages such as elevation, wetland density, percentage of wetland with water, etc. will likely provide additional insight into the factors associated with species presence and species trends throughout GRYN. Such analyses would not require any additional field efforts, only additional analytical time and consideration as to the covariates that would best describe differences at each of scale.

The Rocky Mountain amphibian project (RMAP) is an intriguing approach to expand the scope of amphibian monitoring to the larger GYE. Resources and funding are already insufficient for the current monitoring plan (evidenced by the inability to meet Objective 3) so any additional funds are needed toward surveys/monitoring within the already established sampling areas, and not outside the national parks. RMAP includes the Bridger-Teton National Forest, and data already collected by community scientists in this region could serve as an excellent test case as to whether the geographic scope of the project can be expanded. The survey design for RMAP already is similar to the ongoing GRYN protocol: defined by catchments in which surveyors are asked to sample all wetlands within the catchment, documentation of amphibian breeding, and site data collected (although I assume this may not be identical to site characteristics collected in GRYN monitoring). One key difference is that surveyors visit RMAP sites on two separate occasions in contrast to the two independent observers with one site visit for GRYN monitoring. Given the generally high detectability of breeding amphibians in this system, that may not be a big discrepancy for estimating detection probability. Larger concerns relate to differences in observer skill and training, data quality related to species identification, and possibility of not having each catchment be sampled every year since it is reliant on volunteer interest. These issues are common to any community science program, and I expect that the monitoring team should be able to account for these concerns in integrating survey results into the larger GRYN program. If RMAP data were able to be integrated into the GRYN monitoring analyses, there are great opportunities for additional survey regions. For instance, the Targhee National Forest has had an amphibian monitoring effort at 10 year intervals, and as a result there already exist detailed information on amphibian breeding sites that could inform bringing this national forest into the RMAP program.

***Question 3: Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program? To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of out of their sampling visits?***

The current survey methodology is a visual encounter method supplemented by dipnetting. This is a standard method for surveying lentic amphibians and based on the estimates of detection, is likely effective for detecting the species present in the GRYN. The advantage of this method is that it is very easy to implement, surveyors can be quickly trained, and it is low cost relative to other survey methods. The disadvantage is that it does not provide much biodiversity or community data beyond the focal amphibians and other incidental species recorded such as

snakes or the presence of any fish species. It also has minimal ability to monitor biotic amphibian stressors such as disease.

There are several approaches that could be used along with visual encounter surveys that could provide additional biodiversity and monitoring information. Each would add additional survey time and cost but considering the time and effort it takes to reach each wetland, they may be justified at some level for monitoring.

Environmental DNA (eDNA) would likely have the broadest benefit to better understanding biodiversity at the sampled wetlands. The potential of eDNA for GRYN monitoring would be to: 1) improve detection of the focal amphibian species, particularly in more complex wetlands, 2) provide faster field sampling 3) estimate relative abundance of each amphibian species, 4) identify multiple levels of the biotic community, and 5) provide an archived sample that can be vouchered and investigated into the future.

The most direct benefit of eDNA to the monitoring program as currently constructed would be to increase detection rate with a faster visitation time at each wetland (assuming a visual survey was not conducted at the same time). Probability of detection is high with visual methods for chorus frogs and spotted frogs (0.7 – 0.9; Gould et al. 2019), however Melanie Murphy’s presentation at the symposium suggested that detection probability might be inflated when using visual surveys. I believe that this research will be published soon, and at that point, a critical examination would be warranted to evaluate the magnitude of detection increase with eDNA. It would need to be a significant increase to establish eDNA as a new monitoring protocol because despite shorter sampling times at a wetland, costs would be significantly higher for lab analysis compared to no additional processing costs for the current visual protocol. Furthermore, there would need to be an efficient workflow developed for either filtering water samples in the field or transport of water samples back from the field for either lab filtering or centrifugation. My expectation given my experience with visual and dipnet surveys for these species and typical wetland size in Yellowstone and Grand Teton, and the results from the monitoring is that it would not be necessary to drastically overall the current monitoring protocol to include eDNA at each site purely for the purposes of amphibian detection.

Instead, I anticipate that eDNA may be most beneficial for achieving the monitoring objectives for tiger salamander and western toad. Monitoring has not been effective for either of these species primarily due to the low number of positive sites within the sampled catchments. A rapid eDNA survey targeted in portions of the parks where suitable habitat exists for each species would provide an efficient method for identifying new sites and also perhaps to monitor existing sites. Water collection for eDNA can be done by community scientists. One could imagine an eDNA “blitz” where park visitors are recruited to help collect water, and portable eDNA filtration stations could be set up at trailheads and samples sent to collaborating laboratories. This would require minimal additional investment in staff. This objective could also be accomplished using the RAMP model using visual surveys, but RAMP results thus far have found low detection probabilities for tiger salamander, suggesting community scientist visual surveys may be ineffective for this species. After an initial eDNA sampling effort, sites with salamander or toad positives could then be incorporated into regular visual monitoring if funding allowed.

There has been much excitement about the idea of eDNA as a tool to estimate relative abundance of species. However, there are many factors that can affect the concentration of eDNA for each species. Site factors that can impact eDNA isolation such as acidity and turbidity (Goldberg et al. 2018) may be expected to impact all species equally, but eDNA in lentic systems can be localized (Li et al 2019) and so there may be a bias toward species detection based on the exact location of the sample. This latter issue could be compensated for by collecting subsamples across a wetland and combining the water samples before processing. More problematic are potential differences in shedding rates among species due to factors such as biomass and diet (Klymus et al. 2015) and would be expected to vary among the amphibian species present in the GYE. Currently the monitoring protocol is focused on trends in occupancy, not abundance, although some data relevant to relative abundance is collected by the surveyors. With the current state of the eDNA field, it would be unlikely that eDNA could add trends in abundance as a feasible objective, although archived samples could be used if future methodologies are developed.

If an eDNA protocol were to be incorporated into the current protocol, the largest gap it would fill would be for monitoring of the larger biotic community within wetlands that is missing from the current effort. Community level eDNA sequencing would not only add species across a variety of taxonomic groups to vital signs monitoring, but more specific to amphibians would provide a wider base of hypotheses to test for any change in occupancy trends. A previous pilot study in the park using eDNA metabarcoding was able to detect DNA sequences from several species (including terrestrial species such as grizzlies and wolves) yet had low number of sequence reads for the targets. However, an optimized field protocol that increases sampling volume or extraction efficiency should allow for reliable metabarcoding in the GYE wetlands. In my opinion, the two most value-added targets would be aquatic invertebrates and disease. Both insects and wildlife disease are park vital signs that lack a current long-term monitoring effort such as currently exists for amphibians. eDNA collection from amphibian field sites has the capability to contribute to monitoring both these vital signs. Both target groups are not without their challenges of using eDNA. A recent study compared eDNA detection of macroinvertebrates to bulk metabarcoding of samples from kicknets and found much higher detected diversity through the bulk sampling compared to eDNA, likely due to a high proportion of non-specific amplification (largely algae and bacteria) in the eDNA samples (Gleason et al. in press). It is possible that testing and optimization of more specific PCR primers would increase the effectiveness of eDNA for invertebrates. Similarly, multiple participants in the symposium indicated that eDNA disease testing for chytrid had limited success in the systems where it had been tried. eDNA can be effective for disease monitoring (for instance, Seiber et al. 2020 had 80% detection in a mesocosm study with *Bd*) but can also be inconsistent among sample replicates. A successful protocol would likely require sampling greater volumes and increased site replication. Recent eDNA work for ranavirus (likely the greater disease threat for amphibians in the GYE) has been relatively successful, with multiple instances of detection at sites with die-offs (Hall et al. 2018, Miaud et al. 2019, Vilaça et al. 2019). Detection of ranavirus does appear to have a seasonal pulse that is not necessarily synchronous with actual die-offs and may not correspond to the timing of die-offs.

Of course, a more direct way to sample each of these would be to kicknet sample for macroinvertebrates and swab/tissue sample for chytrid and ranavirus. Each of these could then be sequenced for species detection as an eDNA sample would. Bulk sampling for macroinvertebrates would be more time-consuming as it would require each surveyor to sort through each kick sample and sort invertebrates into sample jars and would require training of field crews. Disease sampling through swabs and tissue sampling would be more straightforward but would also require handling of each individual amphibian, which is not part of the current protocol. Furthermore, tadpoles/larva are the life stage most often encountered during the amphibian monitoring. Tail clips can be effective for ranavirus (but is invasive and would not be appropriate for small larvae). Chytrid sampling of tadpoles is even more difficult and would require either swabbing of mouthparts or euthanizing individuals.

Of these two broader targets, disease almost certainly has the biggest potential impact on amphibians. While the potential impact of chytrid fungus in the park should not be ignored, especially given the rarity of the *Bd*-susceptible western toad, ranavirus appears to be a much greater threat to Yellowstone amphibians (Patla et al. 2016). This is consistent with my experience sampling tiger salamanders in 2003 when I found two sites with die-offs of tiger salamanders from which a unique ranavirus strain was isolated (Ridenhour and Storfer 2008). Therefore, I do think that it would be useful for ranavirus to be included in the rigorous monitoring design going forward, particularly for spotted frogs and tiger salamanders. Certainly specimens could be collected when die-offs are found (as they have been already), but either direct animal sampling or eDNA would provide additional insight into ranavirus distribution or threat. Ideally, a pilot study would examine both approaches to determine which one has higher detection and when samples should be collected. I would suggest that a ranavirus survey effort be targeted toward known spotted frog and tiger salamander sites, as they are the susceptible hosts and because PCR tests for ranavirus at every sample site would likely be too cost-prohibitive. However, an ideal monitoring protocol would involve testing the same sites over multiple years to better understand the persistence of ranavirus and the interaction with climate impacts on ephemeral wetlands (i.e. does wetland drying reduce ranavirus persistence?).

With respect to aquatic insect monitoring, consultation with the biologists working on the insect vital signs would be useful to understand how high a priority that would be for wetland sampling. I would expect that macroinvertebrate sampling would primarily reflect wetland water quality, rather than a direct interaction with the amphibian population. There do not seem to be concerns regarding water quality at wetlands, so this added objective would probably be a lower priority given that there are already concerns about funding.

There is one other potential for further sampling at field sites that could benefit the overall monitoring program, which would be population genetic monitoring of amphibians in the park. This would be accomplished by tissue and swab testing of a sample of individuals during routine monitoring. There has already been population genetic work published on tiger salamanders in the northern range (Spear et al. 2005) and western toads across all of YNP (Murphy et al. 2010), but there is not currently analyzed population genetic data for chorus frogs and spotted frogs within the GYE. There are three main advantages to population genetic data and analysis. First, for chorus frogs and spotted frogs, baseline genetic data would inform how many management units or even evolutionary significant units occur within GRYN, which would provide important context for interpreting trends from the monitoring data. Second, genetic diversity and estimates of effective population size would provide important complementary data to estimates of occupancy and the impact of drought and climate change on populations. Furthermore, the existence of previous genotypic data for tiger salamanders and western toads would be especially useful for estimating any trends in genetic diversity over the past 15-20 years, although it would require using the same microsatellite loci as was used in the previous studies. Finally, if genome-wide markers are used, the investigators can examine if there is evidence of adaptation to climate change factors observed during the monitoring, and better predict if each species will be able to adapt to changing environmental conditions. While DNA extraction and genotyping would require significant additional funds, samples could be collected in the field and archived for minimal additional cost, although it would require handling and processing of individuals similar to swabbing for disease. Because generation time for each species spans several years, sampling would not need to occur each year. In addition, because individual wetlands within a catchment likely are genetically similar due to their proximity, only 1-2 wetlands per catchment would need to be sampled. The use of genetic monitoring might be especially valuable for tiger salamanders and western toads not only because of previous genetic studies, but also the sample size for each species is too low to work well for the current multistate modeling approach. Genetic connectivity occurs at a broader scale than an individual wetland, and a lower number of sampling sites can provide meaningful data and inference relative to traditional amphibian monitoring.

***Question 4: Are current analytical approaches described in the protocol and in Gould et al. (2019) capable of detecting statistically significant trends for all species involved? For detecting principal threats to amphibians? For understanding the drivers of wetland change?***

I should start by disclaiming that I am not an expert in occupancy models, although I am familiar with their use in monitoring and some of their advantages and disadvantages. The multistate models being implemented in this monitoring make a lot of sense to me in the context of GYE amphibians, especially since many of the wetlands are ephemeral. The current analytical framework is well suited to estimate trends in amphibians accounting for natural cycles of wetland drying and sets the stage to understand future climate change and amphibians. For the two species with sufficient number of detections, it appears to provide enough power to model occupancy and trends. During the symposium, Brian Halstead presented on a time to detection method that can increase efficiency in some cases. One of the advantages of time to detection is that it only requires one observer one time, but that is not really relevant to GRYN monitoring because there is already one visit and two observers are necessary for safety precautions. Although the method can decrease time per survey as well (since it is possible that a single dipnet could be the entire survey if you detect the species), this is unlikely for the Yellowstone surveys as there are four targeted species, and it is very rare for all four species to be found at a single site. In that case, even in time to detection surveys, the entire wetland would probably need to be surveyed regardless. However, it would not add any effort for surveyors to record time to first detection, which would provide data for analysts to compare estimates from a hypothetical time to detection design versus the current methodology

(assuming a time to detection occupancy model can fit within a multistate framework, which I don't know if that is a realistic assumption).

I have already discussed the limitations of the current methodology and analytical framework for tiger salamanders and western toads, but that is tied more to survey limitations rather than the analytical framework. I also mentioned in a previous answer that the current analysis does not seem to incorporate the catchment and drainage levels in the analysis despite the monitoring being stratified by these levels. Finally, the current analysis is only relevant to the high and medium quality catchments, as the monitoring team is already aware.

The primary threats to amphibians in the GYE in my view are climate change and disease, with landscape change (i.e. vegetative succession) and development minor threats that may impact some wetlands (i.e. road construction that disrupted Lodge Creek spotted frogs). I believe the protocol is well suited to assessing response to climate change. Disease is not currently part of the statistical trend framework but including it would strengthen any inference of impacts of threats on amphibian trends.

The current analytical framework is addressing climatic drivers of wetland change by incorporating temperature, precipitation, evapotranspiration, and runoff into the multi-state models, but at a relatively coarse scale (1 kilometer). The current research regarding hydrography that we heard about during the symposium on the northern range is using finer scale data (< 30 m) and has great potential to predict drivers of wetland dynamics at an individual wetland scale and be incorporated into the multi-state model. There is still much work to be done to better understand wetland change, but in my opinion this is a strength of the current monitoring effort and I anticipate that the current program will continue to provide important insights on the topic.

***Question 5: How could GRYN increase the impact of their work? How can they work with parks and partners to transform knowledge into conservation action?***

I do think this is one of the big challenges in translating monitoring to conservation action. This is especially true because the biggest threats (climate change and disease) are not localized to the GYE. However, there are a number of strategies that could leverage the long-term monitoring effort to push toward a broader impact across the park. Some of these efforts are certainly underway such as highlighting the monitoring efforts on the park website, Yellowstone Forever herpetology courses, and the recent initiative to promote community science through iNaturalist in the park. Yellowstone is of course more famous for other wildlife other than amphibians and increased public awareness of amphibians can increase support for the amphibians as high priority wildlife in the GYE.

I'm not aware of all the current outreach, and so some of these recommendations may already be implemented in some form. I think having visitors experience firsthand amphibians and their habitat could be really beneficial. For instance, you could have a ranger-led tour that visited one of the monitored catchments that was relatively accessible where the visitors could see amphibians (without entering a wetland) as well as visit wetland sites that are dry that particular year to demonstrate the threats of climate change. This would also allow the ranger to discuss the most up to date findings from the monitoring. Another public education/outreach opportunity would be to reduce spread of disease through boot and gear disinfection protocols presented to the public. This would especially apply to anglers but is relevant to anyone exploring the GYE trails and backcountry. On the fishing regulations on the website, I only saw reference to cleaning boats and gear for aquatic invasives, without direct mention of disease, or description of a disinfection protocol for boots and nets.

Currently it is unclear to me the extent to which there is direct integration of the different GRYN vital signs programs with each other. For vital signs monitoring that have metrics that overlap, or are relevant to multiple taxa or processes, addressing these metrics for management action could have a wider impact that includes amphibians. I don't know enough about the other vital signs monitoring to make any specific recommendations, but something to keep in mind. There are also opportunities to interface with other NPS inventory and monitoring networks – for instance there already is a regular communication with the Rocky Mountain network based on the participants in the symposium. To my knowledge, GRYN is the only network within the larger

western region that has amphibians as a vital sign, but most of the surrounding networks have water quality and climate as a vital sign in their monitoring, and trends in climate and water monitoring could be relevant across the different networks.

I do think there is great opportunity to use the current monitoring to improve amphibian conservation overall with partners. While the GRYN monitoring program only occurs on NPS land, the trends likely extend to other land ownerships within the GYE, and therefore there could be regular contact with surrounding landowners, especially if there are management or conservation implications that could be addressed. This should be easy with other federal landowners, such as the Forest Service, which manages a large proportion of amphibian habitat within the GYE. Outreach to private landowners, especially ranchers, in the GYE could further increase the impact of the program. There are models for private lands conservation and amphibians (such as David Pilliod presented during the symposium) and shared concern of water for ranchers and amphibians might incentivize landowners to maintain and protect wetland habitat for amphibians on their land.

***Question 6: How would you prioritize monitoring activities if budget/time/safety/personnel constraints restricted GRYN's ability to complete current activities?***

I see the current monitoring effort as really a minimum – I think the program would suffer if were to be reduced in some form. However, if there did need to be cuts that reduced survey effort, I would investigate if there are sets of catchments within each larger drainage that are similar to each other with respect to hydrography. In other words, are the proportion of wet and dry sites in the same year similar among the catchment subsets? If so, then one way to attempt to maintain the program with less effort would be to alternate years of sampling catchments with similar hydrologic dynamics in the same region. This would save time but would still involve park-wide travel and yearly survey effort. I would not advocate skipping years entirely due to the known year to year variation in climate that strongly influences wetland availability. It may not be necessary to sample every wetland in each focal catchment, but reducing the number of wetlands per catchment surveyed is unlikely to result in much time or cost savings, as a bulk of the time and cost is getting to the catchment, and catchments are generally small enough that all wetlands seem to be able to be sampled without too much difficulty. Eliminating entire drainages from sampling would certainly reduce costs and time but would eliminate any inference on amphibian trends to a significant portion of the GRYN footprint. A final option would be to further eliminate catchments completely from the program. While this is possible, it would reduce an already small sample size of catchments and I would expect reduce statistical power to detect trends. However, this would be my second choice option for reducing costs after the first option of alternating years at certain catchments (while still maintaining yearly surveys in each drainage).

I hope there will not be a need to reduce the monitoring. Instead, increased funding is needed to meet objectives regarding tiger salamanders and western toads. Reducing the current monitoring would put meeting the objectives for chorus frog and spotted frog in jeopardy.

***Literature Cited:***

- Bennetts R, Corn PS, Daley R, Gould WR, Jean C, Patla DA, Peterson CR, Ray A. 2013. Cooperative amphibian monitoring protocol for the Greater Yellowstone Network: Narrative, version 1.0. Natural Resource Report NPS/GRYN/NRR—2013/654. National Park Service, Fort Collins, Colorado.
- Gleason, J.E., V. Elbrecht, T.W.A. Braukmann, R.H. Hanner, and K. Cottenie. In press. Assessment of stream macroinvertebrate communities with eDNA is not congruent with tissue-based metabarcoding. *Molecular Ecology*, <https://doi.org/10.1111/mec.15597>
- Gould, W., A. Ray, L. Bailey, D. Thoma, R. Daley, and K. Legg. 2019. Multistate occupancy modeling improves understanding of amphibian breeding dynamics in the Greater Yellowstone Area. *Ecological Applications* 29(1):e01825.
- Hall, E.M., C.S. Goldberg, J.L. Brunner, and E.J. Crespi. 2018. Seasonal dynamics and potential drivers of ranavirus epidemics in wood frog populations. *Oecologia* 188:1253-1262.

- Miaud, C., V. Arnal, M. Poulain, A. Valentini, and T. Dejean. 2019. eDNA increases the detectability of ranavirus infection in an alpine amphibian population. *Viruses* 11:526.
- Murphy, M.A., J.S. Evans, and A. Storfer. 2010. Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. *Ecology* 91: 252-261.
- Patla, D., S. St-Hilaire, A. Ray, B.R. Hossack, and C.R. Peterson. 2016. Amphibian mortality events and ranavirus outbreaks in the Greater Yellowstone Ecosystem. *Herpetological Review* 47:50-54.
- Ridenhour, B.J. and A.T. Storfer. 2008. Geographically variable selection in *Ambystoma tigrinum* virus (Iridoviridae) throughout the western USA. *Journal of Evolutionary Biology* 21: 1151-1159.
- Seiber, N., H. Hartikainen, and C. Vorburger. 2020. Validation of an eDNA-based method for the detection of wildlife pathogens in water. *Diseases of Aquatic Organisms* 141: 171-184.
- Spear, S.F., C.R. Peterson, M.D. Matocq, and A. Storfer. 2005. Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Molecular Ecology* 14: 2553-2564.
- Vilaça, S.T., S.A. Grant, L. Beaty, C.R. Brunetti, M. Congram, D.L. Murray, C.C. Wilson, and C.J. Kyle. 2019. Detection of spatiotemporal variation in ranavirus distribution using eDNA. *Environmental DNA* 2: 210-220.
- Wright, A.D., E.H. Campbell Grant, and E.F. Zipkin. 2020. A hierarchical analysis of habitat area, connectivity, and quality on amphibian diversity across spatial scales. *Landscape Ecology* 35:529-544.

## Mary Kay Watry

### *Greater Yellowstone Network Amphibian Monitoring Protocol Review*

Prepared For: Michael Bozek (michael\_bozek@nps.gov)

Prepared By: Mary Kay Watry, Conservation Biologist, Rocky Mountain National Park

Date: January 28, 2021

**Question 1: Is GRYN meeting objectives (Table 1) stated in the current amphibian protocol? If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these.**

**Q1a: Is GRYN meeting objectives (Table 1) stated in the current amphibian protocol?**

Obj. 1: Estimate the proportion of catchments and wetland sites used for breeding by each of the four common, native amphibian species annually, and estimate the rate at which their use is changing over time.

It appears that Objective 1 is **being met** in high and medium quality catchments for Boreal Chorus Frog, Columbia Spotted Frog and Western Tiger Salamander. GRYN is not meeting this objective for the low-quality catchments which represent the majority of the GRYN area monitored or for the Western Toad. See the following section for potential adjustments to address these deficiencies.

Obj. 2: Determine the total number of wetlands within sampled catchments that are suitable for amphibian breeding (i.e., have standing water during the breeding season) annually.

Objective 2 is **being met**. As mentioned above for objective 1, only medium and high-quality catchments are sampled with no evaluation of most wetland catchments and this is of concern. This objective may be expanded through use of remote sensing techniques and is addressed in the following section.

Obj. 3: For Western Toads, estimate the proportion of previously identified breeding areas that are used annually, and estimate the rate at which their use may be changing over time.

Objective 3 is **not being met**. See the following section for potential adjustments to address this deficiency.

**Q1b: If not, what needs to be adjusted to do so? Should GRYN consider other strategies for analyzing monitoring data that would allow them to better address stated objectives? If so, please define these.**

*Potential adjustments for Objective 1:*

In 2011, there was a change to protocols that dropped all low-quality catchments. This is of concern for several reasons. Low quality catchments represent nearly 70% of all identified catchments. The current design is not representative and therefore, broader inferences cannot be made to the condition of wetlands in the study area. Marginal sites are less resilient to change. Low-quality sites may have more consecutive years of unsuitable conditions. Amphibians tend to be boom or bust breeders and therefore, are evolved to have some years with no recruitment. Eventually a tipping point is reached and there are too many years without breeding to maintain a population. Understanding what is happening in low quality catchments is key to understanding what is happening across the landscape.

Low-quality catchments were eliminated to address budget constraints. Eliminating these sites lowered the number of catchments monitored annually from 40 to 31 catchments. A few other options can be considered that would include low-quality catchments and remain budget neutral.

1. Consider replacing high and medium “far” sites, with high and medium “close” sites. The first 10-15 years of monitoring suggest no difference in amphibian detection based on remoteness criteria, yet the time and money invested is significantly greater for the “far” sites. Replacing catchments is a difficult decision and needs to be considered carefully and implemented over time. The shift away from far sites will allow monitoring of more sites in all catchment conditions.

2. Consider monitoring high- and medium-quality catchments on a rotating basis. Prior to dropping low-quality catchments, a rotating panel design was considered. A simulation by the USGS staff at Patuxent Wildlife Research Center indicated a loss of precision for most parameters. There was no discussion of the high- and medium-quality catchments considered for the rotating panel design. High- and medium-quality catchments have more wetland area and therefore, more breeding sites than low-quality catchments. Would this greater number of sites reduce the confounding effects of among and within year variation and loss of precision for parameters at the catchment level? Confounding effects and loss of precision will be greater at the site level and a trade-off for this change. If a rotating panel design was not considered in 2011 for the catchments in medium- and high-quality, it would be worth considering it now. The advantage of moving to a rotating panel design (if feasible) is that it would preserve monitoring existing medium- and high-quality catchments while adding back some low-quality catchments.
3. Consider reducing sites proportionally high-, medium-, and low-quality catchments to better represent overall changes across the spectrum of habitat available. It seems counterintuitive that the most prevalent wetland condition has no survey effort. The low-quality catchments will likely be more responsive to interannual and incremental changes to climate precisely because they are already marginal sites. By eliminating these sites, it may take longer to observe a detectable change as the remaining catchments will need to degrade to a lower quality before becoming unsuitable for breeding.

Objective 1 has not been met for the Western Toad. One reason is simply the assumption at the start of protocol development that there were four common amphibians. The last 15 years of monitoring have shown that only two species are truly common, the Boreal Chorus Frog and the Columbian Spotted Frog, and two are not, the Western Tiger Salamander and the Western Toad. Data show that western toads historically occurred in ~10-12 monitored sites and the maximum number of catchments with observed breeding in a single year was six.

One option is to acknowledge that a broad sampling strategy cannot address a species that is rare, rather than common, and focus efforts to monitor western toads wholly through objective 3. A concern with this approach is that common species typically are more adaptable in habitat used, breeding strategies, etc. and therefore less likely to respond to changes. Rare species may breed with tighter constraints and be better indicators of wetlands reaching trigger points that preclude breeding on a consistent basis. Objective 1 seeks to estimate the rate at which amphibian breeding is changing over time and this may be more quickly observed by including western toads.

The second option is to modify objective 1 design and methods to increase the ability to detect western toad breeding. There are several options that can be considered.

1. If the number of sampled catchments changes, consider historic presence of western toads within a catchment or within a neighboring catchment as one of the criteria for retaining current or including new catchments. Given the relative rarity of the toad, high breeding site fidelity and limited dispersal distances this is the best way to increase catchments with toads or potential for toads in the future. Putting such a strong emphasis on western toads would certainly bias the number of catchments that could be considered in the future. It is unclear what impact this would have on detecting other species, specifically western tiger salamanders which are also much less common than chorus frogs or Columbia spotted frogs.
2. Consider shifting the sampling period to earlier in the season. The current sampling period is approximately eight weeks between mid-June to mid-August. Western toads breed very early and monitoring to assess breeding can begin as early as mid-May. There is a possibility that GRYN is not detecting evidence of western toad breeding at sites even though it is occurring based on this timing mismatch. There are several reasons this could be occurring.
  - a. Occasionally males do not fertilize all or most of egg masses or water mold causes eggs to die. In both cases, breeding will have occurred but will not be detected either because there are no tadpoles or numbers are so low the probability of detecting the species is low.
  - b. There is evidence from reintroduced western toad tadpole monitoring that weekly tadpole survival probabilities range from 0.54 – 0.73, with greater mortality in smaller, less developed tadpoles (Muths et

al. 2014). Western toads that breed earlier will have a greater chance of not surviving to the survey date and therefore will not be detected. At historic western toad breeding sites in Rocky Mountain National Park there are many examples of egg masses and tadpoles detected during earlier site visits but no tadpoles or metamorphs detected later in the season.

- c. Detection of breeding is based on a single individual at any life stage <1 year old. It is reasonable that there will be a difference in the ability to detect one tadpole of a common species that has multiple adults depositing eggs and resulting in larger number of tadpoles versus detecting one tadpole of a rare species which may only have one egg mass deposited and fewer tadpoles available for observation. The differences between species are further exacerbated by the above factors. Is there a way to correct for this difference in the analysis? This was not discussed during the workshop.
3. If shifting the single site visit earlier is not possible, consider multiple site visits and specifically adding an earlier season site visit. To keep this change budget neutral, these site visits could be conducted in lieu of time spent on objective 3 for a pilot trial of three years. If there is no evidence, after this pilot that any detections are missed by later surveys, regular objective 3 monitoring can resume. If detection of western toad breeding increases, additional discussion is needed to determine the number and timing of site visits and the tradeoffs for the suite of species monitored. If a rotating panel design (as discussed in the above catchment section) is a viable option for some catchment types, then two visits to each site in one year and alternating sites each year would be budget neutral.
4. An alternative way to explore whether a mismatch may be occurring between western toad breeding and site visit timing is to record the developmental stage of tadpoles (Gosner stage) in sites where multiple amphibians breed. If there are general trends by which we can understand how much earlier western toads exit the pond, we could make an informed assessment of whether boreal toads were not detected because there was no breeding versus boreal toads were not detected because even if they did breed they would already have exited the pond and were unavailable for detection.
5. Consider incorporating eDNA at pond sites or streamside visual encounter surveys (VES) surveys adjacent to sampled wetlands as an additional check for species presence. Both these methods would not differentiate between breeding and non-breeding; however, they could provide insight as to whether breeding might be missed. It is unlikely that toads would be detected consistently in an area without a breeding site nearby. In Colorado, western toads were confirmed in 13 drainages and three new breeding sites when 55 previously unsampled or minimally sampled drainages were surveyed using a combination of these methods (Bailey et al. 2019). During this study, detecting toads with eDNA varied by year and sample date with detection decreasing as the season progressed. Streamside surveys had little detection variation among years or dates within years. Based on these data, streamside surveys are preferred if only one additional method is added. Additionally, the method used for filtering water was different in Colorado than described by Dr. Melanie Murphy's workshop presentation. Improving the sampling technique may improve detection.
6. Consider consistently starting surveys in the same location and prioritizing the wetland areas most likely to have evidence of breeding. Western toads do not use sites equally and warmer shallows are more likely to be used for egg deposition and tadpole development. To hedge against a partial survey missing optimal habitat, previous surveys can inform a survey starting location and direction to ensure the most productive sections of habitat are sampled, even when time is limited. If this information has not been collected in the past, it can easily be gathered by adding egg deposition, tadpole and metamorph congregation areas to the site sketch.

#### *Potential adjustments for Objective 2:*

Objective 2 as stated is being met. There are opportunities to expand this objective and use emerging technology in addition to field visits. Remote sensing could be used to look at wetland habitat condition that is not constrained by the time required for site visits. This may be a way to incorporate some data from low-quality catchments that are currently not monitored.

#### *Potential adjustments for Objective 3*

Objective 3 has only been addressed through opportunistic and informal monitoring. A formalized sampling strategy with a level of dedicated funding is needed to achieve this objective.

Dr. Larissa Bailey is currently developing a monitoring strategy for the Southern Rocky Mountains as part of the Boreal Toad Conservation Plan update. Once this is complete it may be appropriate to adopt it for GRYN. Alternatively, a possible survey strategy is outlined below that seems achievable based on survey efforts in recent years, would visit all known breeding sites within a 5-year period and that is grounded in western toad research (primarily from the Southern Rocky Mountains). There is no discussion of modeling and statistical analysis, which should be explored but is outside the scope of this brief review process. The strategy would represent a more formalized approach to addressing Objective 3 but would seek to monitor all breeding sites over a longer timeframe.

More than 45 western toad breeding sites have been confirmed from 1999 to present. Confirming a breeding site is “active” can be complicated especially in small populations due to skipped breeding. In the Southern Rocky Mountains, the female western toad does not breed in consecutive years. Muths and collaborators found that only 27% of females breed in the year following a non-breeding year, which means nearly  $\frac{3}{4}$  of females skip breeding for two or more years (Muths et al. 2010). Currently Objective 3 reads, “For Western Toads, estimate the proportion of previously identified breeding areas that are used *annually*, and estimate the rate at which their use may be changing over time (emphasis added).” The combination of limited funding, many sites, remoteness of sites and obligatory skipped breeding argues for a longer timeframe to determine breeding in all sites. One possible strategy is determining whether a western toad breeding site is active within each 5-year timeframe (2020-2024, 2025-2030, etc.) and use each time block to look for trends.

Based on the most recent monitoring report from 2017 and 2018, monitoring 10-15 sites per year is possible, if only one visit is conducted at each site annually. A single site visit appears justified based on the high probability of detection during VES reported for western toad when observed in the catchment surveys, the lack of additional breeding detections when interannual repeat site visits are added during the current western toad opportunistic surveys, and recent analysis comparing survey methods for western toads in the Southern Rocky Mountains (Bailey et al. 2019).

Incorporating consecutive years of surveys is supported by data from the Southern Rocky Mountain Boreal Toad Conservation Team and Rocky Mountain National Park. In Colorado and Wyoming, only two of 70 breeding sites have transitioned from non-breeding to breeding after >4 years without breeding. In these cases, it is unclear if the breeding site remained “active” or was recolonized. In Rocky Mountain National Park, the only two breeding sites that were non-breeding for  $\geq 3$  years and eventually transitioned to breeding, remained non-breeding for over a decade. For all other active sites breeding occurs every year or every other year and only rarely have two consecutive years without breeding been observed. For these reasons, three consecutive years seems to be a reasonable timeframe to determine whether a breeding site is “active” or “inactive”.

A spatially balanced design also appears important due to the differences observed annually in major drainages within YELL and GRTE. Therefore, sites selected for survey in any given year should be distributed across the landscape. Additionally, a metric like objective 2 to differentiate whether a site was available for breeding (standing water) or not is important to record. Over a 5-year period a survey strategy for 45-55 sites may look like the following based on the ~70% observed breeding occupancy from 2017 and 2018:

- Year 1: 15 sites sampled with 11 occupied, 4 unoccupied and carried over for next year’s sampling. Monitoring complete on 11 sites for 5-year period (11 occupied).
- Year 2: 4 sites revisited (1 confirmed occupied, 3 carried over), 11 new sites visited (8 occupied, 3 carried over). Monitoring complete for 20 sites for 5-year period (20 occupied).
- Year 3: 3rd survey at 3 sites (1 occupied, 2 considered unoccupied for 5-year block), 2nd year survey at 3 sites (1 occupied, 2 carried over), 9 new sites visited (6 occupied, 3 carried over). Monitoring complete on 30 sites (28 occupied, 2 unoccupied).
- Year 4: 3rd survey at 2 sites (2 considered unoccupied for 5-year block), 2nd survey at 3 sites (1 occupied, 2 carried over), new survey at 10 sites (7 occupied, 3 carried over). Monitoring complete for 40 sites (36 occupied, 4 unoccupied).

- Year 5: 3rd survey for 2 sites (1 occupied, 1 unoccupied), 2nd survey at 3 sites (1 occupied, 1 unknown due to only 2 years of survey), first survey at 10 sites (7 occupied, 3 unknown due to only 1 year of survey).
- Results for 5-year period. Monitoring complete for 50 sites with 45 occupied, 5 unoccupied and 4 sites with unknown status. Sites with unknown status would be prioritized in the next 5-year period so that no site goes for 2 sampling periods with unknown results.

The number of sites monitored in the above scenario will be driven by how many sites can be completed after one or two visits when occupancy is confirmed. An alternative sampling strategy could include putting a large investment in year 1 to sample all breeding sites, with reduced effort in years 2 and 3 for repeat surveys to non-breeding sites from previous year/s and no effort in years 4 and 5. Another option is repeating breeding site surveys on a 3 rather than 5-year basis but this would be more costly.

Additional methods could be used in combination with VES. As mentioned under the discussion for Objective 1, consider incorporating eDNA at pond and adjacent stream sites or streamside VES surveys adjacent to sampled wetlands. This will provide an additional check for species presence/absence during breeding site visits when toads are undetected in the pond survey. Both these methods will not differentiate between breeding and non-breeding; however, they could indicate that breeding is being missed. It is unlikely that breeding age toads will be detected consistently during the breeding season without a breeding site nearby. In Colorado, western toads were initially detected in 13 new drainages and three new breeding sites when 55 previously unsampled or minimally sampled drainages were surveyed using a combination of eDNA and streamside surveys (Bailey et al. 2019).

***Question 2: Is GRYN monitoring at the appropriate/most ecologically relevant scale? Are there ways to expand outside Grand Teton and Yellowstone national parks to the GYE? To the Western U.S.?***

The current protocol is at the appropriate scale for the goals of the NPS I&M program. Comparing existing methods to those of the Rocky Mountain Amphibian Project and incorporating similar methods or data sharing as appropriate, will expand comparisons with all of Wyoming. Similar comparisons can be conducted with the Boreal Toad Conservation Team for western toad and to other state agencies, conservation organizations or interagency teams that are also studying the amphibians in the GRYN protocol.

*Bd* is critical in the Southern Rocky Mountains and the Boreal Toad Conservation Team uses a metapopulation model that simultaneously predicts occupancy dynamics for both toads and *Bd*. Scaled up comparisons will be facilitated by collecting comparable data. This is especially important for the western (boreal) toad that has been frequently petitioned for listing over the last two decades and whose “listable entity” boundaries are still disputed.

***Question 3: Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program? To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of out of their sampling visits?***

**Q3a: Is GRYN using the best available field survey methods for detecting all breeding amphibian species covered in this long-term monitoring program?**

Under the question 1 response, eDNA and streamside VES surveys were addressed for western toad. The recent publication by Bailey and collaborators (2019) is an excellent summary of field methods and performance for this species with the caveat that eDNA sample collection methods and detection have improved as presented by Dr. Melanie Murphy.

Detection of western tiger salamanders (although not specifically breeding) would improve with use of eDNA. It may be useful to incorporate eDNA sample collection for sites where no tiger salamanders are detected as this species is more likely to be missed when present than the other amphibians studied.

Breeding from the previous year that may have been missed can be confirmed retrospectively if 1-year-olds can be positively identified for any of the four amphibians studied. One-year-old western toads tend to stay close to their breeding site given their small size and have conspicuous orange toe pads. If the other amphibian species are also distinct and unlikely to travel to adjacent sites, then identifying 1-year-olds provides an additional opportunity to detect evidence of breeding. If these data are not being specifically collected and utilized currently, then they should be incorporated with little to no additional cost.

It was unclear during the presentations whether the YELL wildlife health program is sampling for pathogens at the same sites that I&M monitors, and if the projects coordinate and share data. If not, there is a missed opportunity to collect samples for disease testing. It sounds like crews are trained and equipped to collect samples for analysis when mortality is observed; however, proactive pathogen testing is not a part of the current protocol. Pathogen data appear important, particularly with the large number of observed mortality events discovered opportunistically (48 in 15 years, Patla et al. 2016). The cost of swabs, vials, and EtOH are low and amphibian swab sampling is straightforward and efficient. Pooled samples can be relatively inexpensive to analyze and may be tested for multiple pathogens. Handling amphibians is required for swabs and detection can vary by season. Alternatively, water can be filtered and tested using eDNA, however, the efficacy of this technique as compared to amphibian swabs for *Bd* is low (Mosher et al. 2018). The known presence of ranavirus and *Bd* in GRTE and YELL, with past die-offs from ranavirus, chytridiomycosis, and the emerging threat of Bsal, indicate that pathogen presence is a useful covariate that is relatively inexpensive and easy to collect.

**Q3b: To best capture the biodiversity benefits of wetlands? Stated differently, are they making the most of out of their sampling visits?**

From the workshop discussion it does not sound like field crews have extra time to add surveys and address overall biodiversity. One addition that would require little effort, is to cross-train in identification of rare and listed plants that are likely to occur in wetland areas. Park staff may be able to provide a specific photographic field guide for these plants that crews can take with them. Current site visits match the flowering timing of most plants which is often the only time rare plants can be positively identified.

***Question 4: Are current analytical approaches described in the protocol and in Gould et al. (2019) capable of detecting statistically significant trends for all species involved? For detecting principal threats to amphibians? For understanding the drivers of wetland change?***

There are several reviewers who have greater expertise in this area and are better suited to provide substantive comments. I defer to their expertise.

The recently incorporated multi-state occupancy model is a positive change to gain understanding of a basic driver of amphibian breeding: Was the pond present for breeding in the year studied? or was it dry?

***Question 5: How could GRYN increase the impact of their work? How can they work with parks and partners to transform knowledge into conservation action?***

**Q5a: How could GRYN increase the impact of their work?**

See the response for question 2 regarding collecting comparable work with other potential collaborators.

As stated in the response for question 3a, I am concerned that there is no clear pathogen sampling and included in the I&M protocol. There are three reasons this may be of concern. First, although *Bd* does not seem to impact amphibians and specifically western toads in the GYE to the same degree that it does in the Southern Rocky Mountains, this may be due to underlying environmental conditions (temperature and moisture). As climate changes, these conditions will change, and *Bd* interactions with GYE amphibians may change. Collecting and recording the *Bd* status now and in the future may lend itself to further understanding of when the tipping point for *Bd*, environmental conditions and amphibian decline occurs. Second, because *Bd* is critical in the Southern Rocky Mountains, the Boreal Toad Conservation Team uses a metapopulation model that simultaneously predicts occupancy dynamics for both toads and *Bd*. Scaled up comparisons will be facilitated by collecting

comparable data. This is especially important for the western (boreal) toad that has been frequently petitioned for listing over the last two decades and whose “listable entity” boundaries are still disputed. Finally, the amphibian monitoring protocol focuses on two of five potential responses to environmental stressors, occupancy, and distribution. The three responses not addressed are abundance (collected but currently unanalyzed), health and malformation, and genetic change. These responses are interrelated. Negative changes to both the health and genetics can result in lower abundance. In turn, lower abundance can result in lower probability of detection. Disease has clearly been a factor in previous amphibian die offs in YELL and GRTE. Tracking the frequency and presence of disease-causing pathogens is straightforward and can inform potential differences in occurrence and distribution.

Documenting amphibian breeding is important, but recruitment is necessary for population persistence. Recruitment is difficult to measure but a critical first step is for breeding ponds to hold water until metamorphosis occurs and amphibians’ transition from their aquatic to terrestrial state. If there is no survival through metamorphosis, there will be no breeding adults from that year class at that site. Additional site visits late in the season are cost prohibitive. Remote sensing can be used to verify that wetlands are holding water through the anticipated timing of metamorphosis. Confirming ponds persist is not a definitive measure of metamorphosis; however, wetland drying can prevent successful metamorphosis. An estimated metamorphosis date based on survey date and Gosner stage could be derived from previous research and field observations. Pond persistence to these dates can be identified through remote sensing and provide good evidence that survival through metamorphosis could not occur in some years.

Abundance categories are recorded during site visits, but it is unclear how these are communicated or used. Abundance is a missing link when thinking about amphibian decline and their role in the ecosystem. When Lost Lake, Rocky Mountain National Park, experienced a decline where the estimated number of boreal toad males decreased from 700 to 2 in a matter of years, the ecosystem impacts and even recreational experience were changed to a degree that is not captured by only recording breeding occupancy (Muths et al. 2003). Changes to abundance are often precursors to loss of breeding and detectable long before breeding ceases. Explore ways to report these data and trends in abundance is important and will be informative to management.

#### **Q5b. How can they work with parks and partners to transform knowledge into conservation action?**

Interactions with parks and partners appears to be focused on distributing reports and partnering on specific needs as they arise. There are technical committee meetings every other year, but it is unclear how inclusive these meetings are of a variety of participants. An annual collaboration meeting (1/2 to full day) with I&M and a range of collaborators discussing through formal or informal presentations highlights observed in the previous year could be useful. Often by getting key folks together for a discussion there emerge topics and needs that were previously unidentified. Something that was thought to be an anomaly becomes more meaningful when several people observe the same phenomenon. Participants could include present and past wetland and amphibian researchers, park biologists (wetland, aquatic and wildlife) and conservation partners. At Rocky Mountain National Park, we have found these meetings valuable to leverage resources across years where one entity has more funding but another less in any given year. We have also partnered on funding proposals that only some entities (non-profits, states) are eligible to apply.

One conservation action that is clearly needed is to engage more Citizen Scientists to address Objective 3. Could work of the existing Biodiversity Institute and Rocky Mountain Amphibian Project be leveraged to include sites in GRTE and YELL? Not all western toad breeding sites are safe to assign to volunteers but if only a handful can be monitored by others that will spread the existing limited funding of GRYN.

If amphibian die offs are observed at sites, photo document and quantify the loss. Collect samples to assist in identifying the causes of decline. If declines could have been caused by factors that are linked to visitors (transfer of pathogens on fishing/hiking footwear), the information collected can be used to promote individual actions such as cleaning and disinfecting waders. This is a very simple action especially in its most basic form of removing

mud and other debris and drying in direct sunlight. These actions can make a difference and allow visitors to actively work to protect aquatic environments.

**Question 6. How would you prioritize monitoring activities if budget/time/safety/personnel constraints restricted GRYN's ability to complete current activities?**

Several options to potentially reduce time and cost were covered in the response to question 1b. In that context, it was to make room for increased monitoring of low-quality catchments, but these same considerations could apply under a reduced funding scenario. Briefly, three of these ideas were to consider replacing high- and medium-quality “far” sites, with high and medium “close” sites, to consider monitoring high- and medium-quality catchments on a rotating basis using a rotating panel design and to limit western toad breeding site visits to one visit, per site, per year. Under the response to 5b, increasing collaboration with partners such as the Biodiversity Institute and Rocky Mountain Amphibian Project may be a way to continue to collect the same amount of data but at a reduced cost. These ideas are further discussed earlier in this review.

If additional reductions are needed, staff time could be reduced by only sampling a portion of sites within a catchment. This could either be done randomly or designed to specifically retain catchment level data by selecting survey sites that are most likely to encounter the four target species and curtailing additional surveys to sites within a catchment once all species or likely sites for species have been visited. Choosing between these two options may depend on the anticipated length of the funding interruption. Is it one-year or a long-term reduction in funds?

If more funds become available, adding low-quality catchments is a priority. This can be done by restoring the original sample design or can include some of the other ideas proposed during this workshop and review. The second priority is to fund Objective 3 and address it using a formalized design with data analysis. The third priority is adding disease surveillance for both the Objective 1 and 3 field protocols. Once these items are addressed, consider incorporating additional methods to improve data collection for existing objectives that add cost. These could include using eDNA or streamside VES surveys to increase detection, using remote sensing to evaluate more wetland sites or improve tracking of wetland suitability for amphibian breeding throughout the full season. If coordinating an annual collaborator meeting is limited by funding, this should also be added if funds increase.

**Literature Cited**

- Bailey, L.L., P. Jones, K.G. Thompson, H.P. Foutz, J.M. Logan, F.B. Wright, and H.J. Crockett. 2019. Determining presence of rare amphibian species: testing and combining novel survey methods. *Journal of Herpetology* 53(2):115-124.
- Mosher, B.A., K.P. Huyvaert, and L.L. Bailey. 2018. Beyond the swab: ecosystem sampling to understand the persistence of an amphibian pathogen. *Oecologia* 188: 319-330.
- Muths, E., P.S. Corn, A.P. Pessier, and D.E. Green. 2003. Evidence for a disease-related amphibian decline in Colorado. *Biological Conservation* 110: 357-365.
- Muths, E., R.D. Shearer, and B.A. Lambert. 2010. Unbiased survival estimates and evidence for skipped breeding opportunities in females. *Methods in Ecology and Evolution* 1: 123-130.
- Muths, E., L.L. Bailey, and M.K. Watry. 2014. Animal reintroductions: an innovative assessment of survival. *Biological Conservation* 172: 200-208.
- Patla, D., S. St-Hilaire, A. Ray, B.R. Hossack, and C.R. Peterson. 2016. Amphibian mortality events and ranavirus outbreaks in the Greater Yellowstone Ecosystem. *Herpetological Review* 47:50-54.



The Department of the Interior protects and manages the nation's natural resources and cultural heritage; provides scientific and other information about those resources; and honors its special responsibilities to American Indians, Alaska Natives, and affiliated Island Communities.

**National Park Service**  
**U.S. Department of the Interior**



---

**Natural Resource Stewardship and Science**

1201 Oak Ridge Drive, Suite 150  
Fort Collins, Colorado 80525