3 May 2024

TO: Washington Fish & Wildlife Commission and Washington Department of Fish & Wildlife (WDFW)

SUBJECT: S comment on WDFW's 2024 periodic status review and proposal to down-list Washington's gray wolves.

FROM:
Francisco Santiago-Ávila, PhD
Adrian Treves, PhD
Bridgett von Holdt, PhD

Our comment focuses on the science used to justify down-listing. We identify two major areas of concern.

1. **One unreplicated study with many shortcomings should not be the basis for recommending a major policy decision such as down-listing. Also, Petracca et al. (2024) did not take into account scientific concerns and criticisms provided on the 2023 pre-prints and provided as public comments to WDFW. The preprint process is intended to learn from good faith peer review to improve the final product. The opposite seems to have happened resulting in a deeply flawed study (Petracca et al. (2024) informing a WDFW policy proposal.**
2. **The modeling effort in Petracca et al. (2024) did not follow best practices for predictive modeling in ecology (using past data to inform the model that predicts future conditions) nor best practices in its specialized area of modeling animal colonization. Furthermore, we show how Petracca et al. (2024) disregarded a voluminous scientific literature on cumulative threats to wolves and on the interactions between wolf-killing and colonization.**

Throughout, we refer to the two studies listed below as Petracca et al. (2024) and (2023b):

Petracca LS, Gardner B, Maletzke BT, Converse SJ. (2024) Merging integrated population models and individual-based models to project population dynamics of recolonizing species. Biological Conservation 289:110340.

Petracca L. S., B. Gardner, B. T. Maletzke, and S. J. Converse. (2023b). Forecasting dynamics of a recolonizing wolf population under different management strategies. bioRxiv doi: 10.1101/2023.03.23.534018.

We described our concerns in meticulous detail in a letter (Appendix 1) on 19 April 2024 sent to the lead author, L. Petracca, and asked her to share it with her co-authors, as she felt appropriate. They acknowledged receipt but have not yet answered any of our concerns or queries as of 2 May 2024. They directed us to WDFW for a few of our questions below about sharing missing data. **Please consider this a formal request for data therefore.**

Our letter in Appendix 1 is reproduced in its entirety below for the commissioners to see our good faith attempt to understand why some steps were taken and our good faith attempt to enter into scientific debate over many questionable steps in Petracca et al. (2024). Our questions and concerns remain unanswered, so we plan to submit a scientific commentary and rebuttal to the same journal in which Petracca et al. (2024) published.

In Appendix 2, we raise several additional concerns about Petracca et al. (2023b) which informed Petracca et al. (2024) and we fear may inform future unwise policy proposals. In summary these were: (a) unjustified assumptions in the development of alternative scenarios that are not supported in the scientific literature without appropriate sensitivity analyses, and (b) the reliance on value-based decisions focused on increased killing of wolves, as opposed to scenarios for maximizing the likelihood of reaching recovery goals.

Thank you for your time and consideration.

Adrian Treves, PhD
Professor, Carnivore Coexistence Laboratory, University of Wisconsin-Madison
Francisco J. Santiago-Ávila, PhD
Project Coyote
Bridgett M. vonHoldt, PhD
Associate Professor of Ecology and Evolutionary Biology, Princeton University

Appendix 1: Letter to Dr. I. Petracca sharing concerns with Petraca et al. (2024)
Appendix 2: Concerns with Petracca et al. (2023b)

Appendix 1
18 April 2024

Dear Dr. Petracca and co-authors,
We hope this finds you healthy and happy. We read with interest your 2024 article in Biological Conservation on integrating models of wolf movement and population change, its 2023a pre-print, and the 2023b pre-print forecasting the wolf population while modeling different scenarios.

As we prepare scientific commentary on the subject, we would like to ask a few questions about methods and share a few concerns. We hope you will see this as constructive. We're also hoping you'll engage with us in a constructive scientific discussion to advance the field and inform policy.

Before noting our concerns, we reiterate a statement posted in the pre-publication review [1], that "None of [our] comments or criticisms below is meant to undermine the hard work put in, but rather they are meant to improve the final product, improve outcomes for wolves, and improve the policy that may result from applied research."

Our first concern (A) is that a single, unreplicated study with many scientific uncertainties should not be touted as fact and suggested to be a foundation of the policy process. Our second concern (B) is with parametrization that consistently leans towards earlier attainment of the state delisting goal while not discussing other parameterization scenarios that would delay delisting.

(A) Incautious decisions using one study as a justification
Our primary concern is that the types of models presented in your 2024 article will tempt agencies to change policy in anticipation of meeting biological goals on the ground. Relatedly, we worry when past deliberative processes are overturned for new policy (down-listing in this case) that has not undergone deliberation but hinges instead on a single (or a pair of) study(ies). We find that such a deliberation is lacking, especially when based on a single paper that has been questioned during pre-publication review (2023a,b) as we describe below. We are concerned that models of this sort should not be used to make "anticipatory" policy decisions and moving policy "goalposts" (referencing statements by Dr. Carlos Carroll[2]). This problem is exemplified in the recent proposal by the Washington Department of Fish & Wildlife (WDFW) to down-list wolves in the state to 'sensitive' following your published predictions. The WDFW's proposal points to your publication as justification despite the wolf population not having achieved the geographic distribution thresholds of the Wolf Plan for the Southern Cascades and Northwest Coast recovery region, which seems to be a requirement for such down-listing proposals.

A projection may be a useful exercise but scientifically, by definition, is only one of many potential realities, and one in which warnings should abound given the use of past data to forecast future scenarios. Therefore, we suggest that such exercises should not affect protective policies and statutes that have clear precautionary guidelines for removing protections, especially when doing so will harm those policy goals. Such use of predictive models seems opposed to the spirit of the precautionary principle intended to prevent or mitigate harms to biodiversity, climate resilience, and public health. Currently, WDFW is using your modeling exercise for exactly the opposite of

---

[1] 1. Treves A. Pre-publication review of "forecasting dynamics of a recolonizing wolf population under different management strategies" by Petracca et al. . Biorxiv. 2023; https://www.biorxiv.org/content/10.1101/2023.03.23.534018v1#comments.

[2] Carroll, C. Scientific peer review of WDFW Draft Periodic Status Review for the Gray Wolf in Washington (Smith et al., 2023) and supporting documents (Petracca et al., 2023a, 2023b). July 2023.

precautionary protections, and we urge you to clearly convey to them the limits of your exercise and the perils of using it in such a way.

Similarly, we urge you to engage in clearer communication about uncertainty when addressing the public and particularly, decision-makers, lest undue confidence be applied to predictions that have influences on organisms in real ecosystems. For example, your 2024 article reported higher confidence that wolves will recolonize all suitable habitat in WA to 100%, whereas the 2023a pre-print reported only 99% confidence. Even the latter seemed high to us. While seemingly a trivial change in conclusions, the increase in confidence was not explained and moreover represents a mathematical impossibility (a 100% probability implies no uncertainty). Indeed, your own results suggest a finite risk of quasi-extinction by your definition (prediction interval nearing 33%, right?) And even extinction has a prediction interval that exceeds zero, right? Therefore, 100% confidence seems impossible. Moreover, the questions raised about 2023a,b (more below) should lower confidence in the predictions, not the opposite.

Additionally, we are concerned that the science is being used as justification for changing a policy despite the science being an unreplicated forecasting effort with many question marks that we summarize in (B) below. In short, (conveying) caution seems lacking in the 2024 article. In particular, the pre-print process did not seem to promote caution as we had hoped. For example, in 2023, Treves commented on your pre-print (Petracca et al. 2023b):

> "*I acknowledge the risk posed by preprints, such as policy-makers or the public running with results or inferences before they have been approved by qualified peer scientists. I think two aspects of the preprint process guard against such undesirable outcomes: (a) peer reviews attached to the preprint as a comment should serve to caution against such precipitous use of preprints, and (b) the authors can reinforce the need for caution in subsequent revisions to the preprint, even citing their pre-reviewers.*"[3]

It seems these two concerns were both well-founded. The WDFW acted on your preprint and Petracca et al. (2024) did not acknowledge, or more importantly, repair the shortcomings that triggered the concerns of pre-publication reviewers. Although we acknowledge that authors are not beholden to qualified pre-publication reviewers' thoughts and comments on pre-prints, it would have been encouraging to see expert advice taken to heart, discussed, and integrated or if not, justified why the pre-publication reviews were not heeded.

(B) Concern with incomplete accounting for possible parameter values.

We find the scientific pre-publication reviews by Drs. Treves and Carroll remain relevant to your 2024 article. We share concerns similar to those expressed by Dr. Carroll in a letter sent to the WDFW and its commission in 2023, in which he stated an overarching concern about the reliance of policy-makers on Petracca et al. (2023a pre-print),

> "…*[dispersal] is the component most central to the WDFW down-listing proposal, which is based on the model's prediction that the Washington Southern Cascades will be recolonized successfully even if more wolves are killed in the existing source population in eastern Washington*" (p.8) [4].

We agree that dispersal is the most critical, albeit the most uncertain, component of the models in Petracca et al. (2024). We add that colonization dynamics were modeled without considering

[3] https://www.biorxiv.org/content/10.1101/2023.03.23.534018v1#comments

[4] Carroll, C. Scientific peer review of WDFW Draft Periodic Status Review for the Gray Wolf in Washington (Smith et al., 2023) and supporting documents (Petracca et al., 2023a, 2023b). July 2023.

numerous peer-reviewed publications and the known interactions between wolf movement, mortality, habitat quality, and pack stability. While we understand models simplify reality, the 2024 model assumes zero interaction between agencies killing wolves ("removal") or allowing/promoting it ("harvest") and any other demographic rates.

Likewise, we find numerous parameterizations and specifications of the model openly oppose published work on wolves without citing or justifying such contradictions. Below, we call your attention to replicated evidence from several sites with wolves, which undermine stated and unstated assumptions in Petracca et al. (2024) (see citations at the end of our letter). We would appreciate your view of how the following interactions might affect your conclusions:

> **Dispersal:** Regarding long-range movements leading to new pack establishment, your model assumes no "moves" prior to age 1 year but the data from other studies does not support this assumption. Indeed, wolves in the 6-12 month age class do engage in long-range, long-lasting extra-territorial movements (e.g., Fuller et al. 2003 [Table 6.6], Treves et al. 2009). Similarly, human-caused mortality was judged to trigger compensatory immigration from neighboring unhunted populations in Alaska (Adams et al. 2008). In several other studies, wolf packs were found to destabilize and even vanish after agency lethal interventions (Bradley 2004; Brainerd et al. 2008; Borg et al. 2015).
>
> **Uniform reproduction:** The Petracca et al. (2024) model also appears to assume equivalent reproduction in any pack regardless of its size, habitat quality, and past effects of human-caused killing. To our minds, these assumptions are unsupported and opposed by years of data. For example, replicated work by independent researchers working in different areas has shown that the complex social organization of wolf packs affects births, deaths, and dispersal. Mortality affects pack stability, dispersal, and breeding (Ausband et al. 2015, 2017, Bassing et al. 2020, Borg et al. 2015, Brainerd et al. 2008, Cassidy et al. 2023, Haber et al. 1996, Rutledge et al. 2010, Sand et al. 2006, Schmidt et al. 2008, Stahler et al. 2006, Smith et al. 2020).
>
> **Habitat:** A wolf pack's range and the habitat quality can also affect reproduction, death, and dispersal. Have we misinterpreted the modeling of 0-6 pups per pack? If so, would you be willing to share the data on pup production that were adapted to each pack's history and habitat quality and the algorithm for assigning pup recruitment on a pack-by-pack basis?
>
> **Fine-scale movement:** As Dr. Carroll pointed out, models of wolf movement that do not consider step by step decisions by wolves to cross or not cross obstacles are likely to over-estimate the rate of colonization of new areas. He wrote, "*Because the model cannot consider any type of step-wise dispersal mortality or behavioral response (avoidance, attraction) to individual landscape features, it ignores most sources of variation in dispersal success and destination*."[5] As evidence for his claim, he points to an early prediction (Maletzke et al. 2015) that Washington's Southern Cascades would be colonized by 2021. Given that prediction has still not been met, a more transparent summary of why the

[5] Carroll, C. Scientific peer review of WDFW Draft Periodic Status Review for the Gray Wolf in Washington (Smith et al., 2023) and supporting documents (Petracca et al., 2023a, 2023b). July 2023.

prediction failed and what was done to correct it in Petracca et al. (2024) would be helpful.

Immigration: We do not understand the origins of the assumption that (baseline) immigration from out of state would continue at past rates, and in particular given the population reductions taking place in adjacent states. Can you please share the data on all collared animals including lost to monitoring so we can understand the immigration data on which that assumption rests? This is a general problem with the two pre-prints and Petracca et al. (2024); namely, that data supporting assumptions is unclear or absent.

Bias introduced by censoring lethal removals: The model step that censored removals seems to dismiss additive mortality without discussing the assumption. Additive mortality has been shown for multiple wolf populations (Adams et al. 2008, Murray et al. 2010, Sparkman et al. 2011); and in population-level analyses also (Creel & Rotella 2010; Vucetich 2012). Further, this model step also treats removals as uninformative deductions from N, which is an assumption that would be violated if removals were influenced by population size or if removals influenced other mortality causes (literature supports both expectations; Chapron & Treves 2016, Louchouarn et al. 2021, Santiago-Ávila et al. 2020, 2022). Finally, Washington state's history suggests a correlation between N and lethal removals, as in many states (Fritts et al. 1992, Fig. 7 for Minnesota, Chapron & Treves 2016, Figure1a and 1b for Wisconsin and Michigan respectively). It's possible that in every population of wolves, the perception of the need for lethal removal increases as the number of wolf-human interactions increases or the number of wolves increases. Therefore, assuming a constant number of lethal removals or even a constant proportion of the wolf population merits justification that considers prevailing empirical evidence.

Sex ratio: Your 2024 model assumed a 1:1 sex ratio. Thus, the assumption that a breeding pair will form when >2 wolves are in the same territory seems consequential. Would not an equal sex ratio imply a 50% chance of failure to form a pair bond when two wolves meet in suitable territory, given a 50% chance that both wolves in the territory will be of the same sex? This is cursorily mentioned in the Discussion, but without addressing how important such an assumption may be. We would have liked to see a careful sensitivity analysis for modeling that uncertainty and the impact it may have on population growth and recovery objectives.

Propagating error introduced by measurement bias prior to 2020: It appears that the field monitoring methods employed by WDFW have undisclosed uncertainty and systematic biases that should be considered in your models. Consider the estimates of pup survival to winter, pack sizes, and migration into or out of packs. The methods assume that individuals seen in winter could be differentiated as pups of the year from young immigrants and in turn from residents present the year before. Sexing and age estimation from aerial telemetry observations is dubious, especially given the errors in age estimation from field inspection of carcasses and comparisons of estimates based on size to cementum annuli in tooth enamel (Costello et al. 2004 for bear ages; Treves et al. 2017 for wolf age estimation from size). How could aerial telemetry be justified as more accurate for age estimation than either of those two methods involving hands-on inspection of animals or remains? From there the estimates of migration become even more tentative, and then the relative contributions of birth and migration to population size change become tenuous. Furthermore, when the estimate of census size (N) becomes statistically non-independent from the estimate of reproductive performance — a problem we have pointed out in Wisconsin (Wydeven et al. 2004) — expectations about future pup survival become that much less certain. Claims made about pup production in summer also merit scrutiny. Experimental tests failed to find accuracy and consistency by experts conducting summer howl surveys to estimate pup production by packs (Palacios et al. 2017). Hence, we surmise the uncertainties around all WDFW census data and resulting parameter values have been underestimated. Although your Bayesian approach allows for uncertainty about life-history rate parameters, the methods do not seem to allow a pack-specific approach to uncertainty but instead a population wide handling of uncertainty. Using population-wide parameter values would tend to create more confidence in parameter values than might be warranted when simulating single pack colonization of new areas in the future as Petracca et al. (2024) attempts. By contrast, we expect pack-specific life-history rates. For example, packs in the Colville Reservation would presumably reflect different life history rates than those in other parts of Washington given the tribe's legal killing. Although we do not necessarily expect such sophisticated forecast models beyond 2021, the model you constructed for the period before 2021 seems to demand such care in identifying non-additive, non-linear, and variable temporal changes in life-history parameters for each pack simulated after 2020. Any errors made in estimating life-history, habitat suitability, and pack persistence prior to 2021 are likely to propagate into the forecast models in ways that are hard to predict. Therefore, it seems to us that much more care would be needed in describing uncertainty prior to 2021 and justifying assumptions that propagate error beyond 2020. Under such conditions of uncertainty, claiming 100% confidence in a prediction seems inadequate.

Discarded information: Finally, we are concerned that prior information was discarded without explanation. We perceive an unjustified step in the modeling for years 2009-2020 in Petracca et al. (2024). The data used for the IBM component for spatial projections did not include data on the colonization of new areas (e.g., rates of successful colonization and breeding pair or pack establishment, including conditional on distance from initial territory and RSF covariates prior to 2020). Yet the data were available, at least to the WDFW co-author, we presume. Discarding data on past colonization to predict future colonization seems to deviate from the best practices expected in forecast modeling methods. Modeling unmoored from the constraints of real-world data from the same region and same time period is not best practice. The model your paper states was built from 2009-2020 data should have been validated against the actual time steps of colonization observed from 2009-2020 (internal validation), before being applied to the unknown periods from 2021 onward. If there is scientific justification for ignoring past colonization events, we would be keen to learn of such. Our concern was anticipated in 2023, on p.7 of Carroll's public comment on the PSR: *"Predictive modeling of such events is quite difficult, and even realistic models give results with high uncertainty. It would be informative to apply Petracca et al. (2023a, 2023b) model and more realistic SEPM models to "backcast the observed relatively slow rate of pack establishment in western Washington and the Pacific states."* We echo this request.

We look forward to a healthy, constructive scientific discussion with our goal that public policy can stand on a firm foundation of replicable evidence derived from robust study designs.

Sincerely,

A. Treves, PhD, Professor, Carnivore Coexistence Laboratory, University of Wisconsin-Madison
F. J. Santiago-Ávila, PhD, Project Coyote
B. M. vonHoldt, PhD, Associate Professor of Ecology and Evolutionary Biology, Princeton University

Literature cited

Adams, L. G., Stephenson, R. O., Dale, B. W., Ahgook, R. T., & Demma, D. J. (2008). Population Dynamics and Harvest Characteristics of Wolves in the Central Brooks Range, Alaska. Wildlife Monographs, 170, 1–25.

Ausband, D. E., Mitchell, M. S., Stansbury, C. R., Stenglein, J. L., & Waits, L. P. (2017). Harvest and group effects on pup survival in a cooperative breeder. Proceedings of the Royal Society B: Biological Sciences. https://doi.org/10.1098/rspb.2017.0580

Ausband, D. E., Stansbury, C. R., Stenglein, J. L., Struthers, J. L., & Waits, L. P. (2015). Recruitment in a social carnivore before and after harvest. Animal Conservation, 18(5), 415–423.

Bassing, S. B., Ausband, D. E., Mitchell, M. S., Schwartz, M. K., Nowak, J. J., Hale, G. C., & Waits, L. P. (2020). Immigration does not offset harvest mortality in groups of a cooperatively breeding carnivore. Animal Conservation, 23(6), 750–761.

Bradley, E. H. (2004). *Evaluation of wolf-livestock conflicts and management in the northwestern United States*. University of Montana.

Brainerd, S. M., Andrén, H., Bangs, E. E., Bradley, E. H., Fontaine, J. A., Hall, W., Iliopoulos, Y., Jimenez, M. D., Jozwiak, E. A., Liberg, O., Mack, C. M., Meier, T. J., Niemeyer, C. C., Pedersen, H. C., Sand, H., Schultz, R. N., Smith, D. W., Wabakken, P., & Wydeven, A. P. (2008). The effects of breeder loss on wolves. The Journal of Wildlife Management, 72(1), 89–98.

Borg, B. L., Brainerd, S. M., Meier, T. J., & Prugh, L. R. (2015). Impacts of breeder loss on social structure, reproduction and population growth in a social canid. Journal of Animal Ecology, 84(1), 177–187.

Bryan, H. M., Smits, J. E. G., Koren, L., Paquet, P. C., Wynne-Edwards, K. E., & Musiani, M. (2015). Heavily hunted wolves have higher stress and reproductive steroids than wolves with lower hunting pressure. Functional Ecology, 29(3), 347–356.

Cassidy, K. A., Borg, B. L., Klauder, K. J., Sorum, M. S., Thomas-, R., Dewey, S. R., Stephenson, J. A., Stahler, D. R., Gable, T. D., Bump, J. K., Homkes, A. T., Windels, S. K., & Smith, D. W. (2023). Human- caused mortality triggers pack instability in gray wolves. Frontiers in Ecology and the Environment, 1–7.

Chapron, G., & Treves, A. (2016). Blood does not buy goodwill: allowing culling increases poaching of a large carnivore. Proceedings of the Royal Society of London B: Biological Sciences, 283(1830), 20152939.

Costello CM, Inman KH, Jones DE, Inman RM, Thompson BC, Quigley HB. Reliability of the cementum annuli technique for estimating age of black bears in new mexico Wildlife Society Bulletin 2004;32(1):169-76.

Creel, S., & Rotella, J. J. (2010). Meta-analysis of relationships between human offtake, total mortality and population dynamics of gray wolves (Canis lupus). *PLoS One*, *5*(9). https://doi.org/10.1371/journal.pone.0012918.

Fritts SH, Paul WJ, Mech LD, Scott DP. Trends and management of wolf-livestock conflicts in minnesota: US Fish and Wildlife Service, Resource Publication 181, Washington, DC.; 1992. https://web.archive.org/web/20000831154850/http://www.npwrc.usgs.gov/resource/1998/wolflive/program.htm

Fuller, T. K., Mech, L. D., & Cochrane, J. F. (2003). Wolf population dynamics. In D. Mech & L. Boitani (Eds.), *Wolves: Behavior, Ecology and Conservation* (pp. 161–191). The University of Chicago Press.

Haber, G. C. (1996). Biological, Conservation, and Ethical Implications of Exploiting and Controlling Wolves. Conservation Biology, 10(4), 1068–1081.

Louchouarn, N., Santiago-Ávila, F. J., Parsons, D. R., & Treves, A. (2021). Evaluating how lethal management affects poaching of Mexican wolves. Royal Society Open Science, 8(200330).

Maletzke, B. T., R. B. Wielgus, D. J. Pierce, D. A. Martorello, and D. W. Stinson. (2016). A metapopulation model to predict occurrence and recovery of wolves. The Journal of Wildlife Management 80:368-376.

Murray, D. L., Smith, D. W., Bangs, E. E., Mack, C., Oakleaf, J. K., Fontaine, J., Boyd, D., Jiminez, M., Niemeyer, C., Meier, T. J., Stahler, D., Holyan, J., & Asher, V. J. (2010). Death from anthropogenic causes is partially compensatory in recovering wolf populations. Biological Conservation, 143(11), 2514–2524.

Palacios V, Font E, García EJ, Svensson L, Llaneza L, Frank J, López-Bao JV. Reliability of human estimates of the presence of pups and the number of wolves vocalizing in chorus howls: Implications for decision-making processes. European Journal of Wildlife Research. 2017;63:59-66.

Petracca, L. S., B. Gardner, B. T. Maletzke, and S. J. Converse. (2023a). Merging integrated population models and individual-based models to project population dynamics of recolonizing species. bioRxiv doi: 10.1101/2023.03.14.532675.

Petracca, L. S., Gardner, B., Maletzke, B. T., & Converse, S. J. (2024). Merging integrated population models and individual-based models to project population dynamics of recolonizing species. *Biological Conservation*, *289*, 110340. https://doi.org/10.1016/j.biocon.2023.110340

Petracca L. S., B. Gardner, B. T. Maletzke, and S. J. Converse. (2023b). Forecasting dynamics of a recolonizing wolf population under different management strategies. bioRxiv doi: 10.1101/2023.03.23.534018.

Rutledge, L. Y., Patterson, B. R., Mills, K. J., Loveless, K. M., Murray, D. L., & White, B. N. (2010). Protection from harvesting restores the natural social structure of eastern wolf packs. Biological Conservation, 143(2), 332–339.

Sand, H., Wikenros, C., Wabakken, P., & Liberg, O. (2006). Effects of hunting group size, snow depth and age on the success of wolves hunting moose. Animal Behaviour, 72, 781–789.

Santiago-Ávila, F. J., Agan, S., Hinton, J. W., & Treves, A. (2022). Evaluating how management policies affect red wolf mortality and disappearance. Royal Society Open Science, 9(210400). https://doi.org/https://doi.org/10.1098/rsos.210400

Santiago-Ávila, F. J., Chappell, R. J., & Treves, A. (2020a). Liberalizing the killing of endangered wolves was associated with more disappearances of collared individuals in Wisconsin, USA. Scientific Reports, 1–14.

Schmidt, K., Jędrzejewski, W., Theuerkauf, J., Kowalczyk, R., Okarma, H., & Jędrzejewska, B. (2008). Reproductive behaviour of wild-living wolves in Białowieża Primeval Forest (Poland).

Smith DW, Stahler DR, and MacNulty DR (Eds). 2020. Yellowstone wolves: science and discovery in the world's first national park. Chicago, IL: University of Chicago Press.

Sparkman, A. M., Waits, L. P., & Murray, D. L. (2011). Social and Demographic Effects of Anthropogenic Mortality : A Test of the Compensatory Mortality Hypothesis in the Red Wolf. *PLoS ONE*, *6*(6). https://doi.org/10.1371/journal.pone.0020868

Treves, A., Martin, K. A., Wiedenhoeft, J. E., & Wydeven, A. P. (2009). Dispersal of gray wolves in the Great Lakes region. In *Recovery of Gray Wolves in the Great Lakes Region of the United States* (pp. 191–204). Springer. https://doi.org/10.1007/978-0-387-85952-1_12

Treves A, Langenberg JA, López-Bao JV, Rabenhorst MF. Gray wolf mortality patterns in wisconsin from 1979 to 2012. J Mammal. 2017;98(1):17-32.

Vucetich, J. A. (2012). *Appendix: The influence of anthropogenic mortality on wolf population dynamics with special reference to Creel & Rotella (2010) and Gude et al.(2011) in the Final peer review of four documents amending and clarifying the Wyoming gray wolf management plan.* FWS-R6-ES-2011-0039; 92220-1113-0000-C6. https://www.federalregister.gov/documents/2012/05/01/2012-10407/endangered-and-threatened-wildlife-and-plants-removal-of-the-gray-wolf-in-wyoming-from-the-federal

Wydeven AP, Treves A, Brost B, Wiedenhoeft JE. Characteristics of wolf packs in wisconsin: Identification of traits influencing depredation. In: Fascione N, Delach A, Smith ME, editors. People and predators: From conflict to coexistence. Washington, D. C.: Island Press; 2004. p. 28-50.

Appendix 2: Concerns with Petracca et al. (2023b)

Adapted from A. Treves (2023) pre-publication review and comments on Petracca et al. (2023b) Forecasting dynamics of a recolonizing wolf population under different management strategies**. by** Lisanne S. Petracca, Beth Gardner, Benjamin T. Maletzke, Sarah J. Converse. Biorxiv https://www.biorxiv.org/content/10.1101/2023.03.23.534018v1 - comments.

Below our letter to the authors, we also identify several concerns with Petracca et al. (2023b) we have yet to convey to the authors, which can be summarize as: (a) unjustified assumptions in the development of alternative scenarios and use of parameter values which are not supported in the scientific literature and (b) the reliance on value-based decisions focused on increased killing of wolves, as opposed to maximizing the likelihood of reaching recovery goals, for the development of alternative scenarios. We have also appended Treves' critique of Petracca et al. (2023b). For more details the following fice concerns fall into one of the two types of problems identified by (a) and (b) above.

(A) **The baseline scenario is derived from Petracca et al. (2024) and therefore suffers from the same scientific shortcomings.**
All the shortcomings, limitations and unstated assumptions we have identified in Petracca et al. (2024) are reproduced in Petracca et al. (2023b), since the point of departure for the forecasted management scenarios is a baseline scenario derived from the former study. Therefore, Petracca et al. (2023b) also neglected to follow best modelling practices or principles of open science, and is equally unreliable.

(B) **Unjustified assumptions in the use of parameter values for alternative scenarios.**
Treves' commentary on the study highlights the unjustified use of parameter values that are neither supported by the scientific literature nor account for annual variability, such as the use of a single point estimate for lethal removal rate (0.04). The use of that particular value for lethal removal throughout the study period conceals unstated assumptions unsupported by the scientific literature, including: (1) that neither non-lethal nor lethal methods for mitigating conflicts will be effective in reducing conflict, such that constant lethal removals will be permanently necessary, and (2) that predation on domesticated animals is random (since the rate was applied randomly). Moreover, the 'harvest' parameter does not "acknowledge the many sources of evidence for super-additive mortality when the public begins killing wolves".

(C) **The study's scenarios are value-based rather than scientific decisions.**
Value-based decisions reflect personal or organizational preferences or beliefs. Assumptions about parameter values or variable interactions should be transparent and scientifically justified, but this study fails to do so. The absence of stated assumptions in the modeling paper appears to be a scientific misstep, as the circumscribed range of parameter values lacks transparency and peer-reviewed justification. The authors do not consider alternative scenarios for wolf-human coexistence, presenting only a subset biased towards negative views of wolves, with an emphasis on increased killing.

(D) **The study omits inclusion of scenarios that would minimize time to fulfilling the state's commitment to recovery.**
Despite having the explicit goal of exploring the probability of meeting the state's wolf recovery goals, the study fails to explore management scenarios that would minimize time to recovery, such as reducing removals throughout the state (i.e., including Eastern Washington). Failing to explore scenarios that would improve the state's likelihood of meeting its policy goals is a value-based judgment by the authors and WDFW. In doing so, the study presents a biased assessment given it focuses not on what the state can

do to achieve recovery as soon as possible (arguably, its commitment and duty), but rather on how much the state can shirk that commitment and reduce the wolf population while still, eventually, achieving recovery.

(E) The study only models exclusive yet unrealistic scenarios, with no overlap or interaction between them.

Petracca et al. (2023b)'s predictions rely on Petracca et al.'s forecast of 11 management scenarios, developed with WDFW input, each considering a different target for management or consideration of uncertainty ('conditions' hereafter). It is highly unlikely that real 'systems' (i.e., populations) will be subject exclusively to one or other scenario, and especially for the length of the time period forecasted (2020-2070). Instead, realistic modeling should have included multiple scenarios with overlap between conditions, such as simultaneous increases in 'harvests' and 'removals' known to occur with increasing wolf populations, with various levels of disease, in addition to reductions in immigration (given substantial reductions in adjacent wolf populations). Such interaction between scenarios would have surely reduced the probability of wolf recovery in the state. For example, it is easy to envision a scenario where there is increased harvest mortality, lethal removals and reduced immigration, all three of which, individually, "resulted in low probabilities (0.11, 0.18, and 0.27, respectively) of meeting recovery goals across all years (2021-2070)." It is concerning and scientifically irresponsible that such an obvious biological reality was not explored or even discussed, especially when scenarios were developed alongside WDFW staff for the explicit intent of exploring how state wolf recovery would be impacted by such conditions. The certain overlap of such conditions with reduced state protections suggest recovery is still precarious, and downlisting premature.

Additionally, WDFW's response that "this exercise was a sensitivity analysis to examine sensitivity to particular kinds of threats or management to better understand which may or may not significantly affect the wolf population." is evasive. Sensitivity to exclusive kinds of threats or management cannot appropriately account for sensitivity to how their interaction would affect the wolf population since the above sources of mortality compound each other in a social species, affecting everything from colonization to pair bonding to recruitment. It is also contradictory that, while WDFW suggests that the study was "not intended to predict the future", the agency is relying on its forecasting precisely in that way by anticipatorily reducing protections instead of exploring how to reduce such conditions to achieve the state's objective.

Pre-publication review of "Forecasting dynamics of a recolonizing wolf population under different management strategies" by Petracca et al. https://doi.org/10.1101/202...

Reviewed by
Adrian Treves, PhD
Professor of Environmental Studies, Founder and Director of the Carnivore Coexistence Lab, University of Wisconsin-Madison
+1-608-890-1450
https://faculty.nelson.wisc.edu/treves/CCC.php
Direct inquiries to atreves@wisc.edu
11 May 2023

I appreciate that Dr. Petracca and colleagues posted their manuscript to a preprint server to facilitate independent review and scientific debate. Such preprints are a healthy step in our field to improve the reliability of science.

Also I acknowledge the risk posed by preprints, such as policy-makers or the public running with results or inferences before they have been approved by qualified peer scientists. I think two aspects of the preprint process guard against such undesirable outcomes: (a) peer reviews attached to the preprint as a comment should serve to caution against such precipitous use of preprints, and (b) the authors can reinforce the need for caution in subsequent revisions to the preprint, even citing their pre-reviewers. The science-policy interface in which this work lies is fraught with difficulties.

Also I acknowledge these sorts of models are complex and difficult to parameterize realistically with confidence. None of my comments or criticisms below are meant to undermine the hard work put in, but rather they are meant to improve the final product, improve outcomes for wolves, and improve the policy that may result from applied research. Thanks in advance for reading my comments in that spirit.

I have chosen not to cite much research below, instead calling the authors' attentions to our website (above) where peer-reviewed substantiation of all my assertions can be found. I welcome peers' emails to atreves@wisc.edu if anyone has trouble finding the evidence.

Most of my comments relate to Tables 1 and 2 and the associated scenarios.
A question about Table 1: the caption includes "Lethal removal rate was calculated directly from state agency records." Please provide those with annual numbers and locations (East or West) to help the reader understand the geographic and spatial context of that assertion.

The annual lethal removal rate was a single point estimate of 0.04. I don't understand why this was treated as a constant not bracketed by annual variability? Later, the authors wrote "In scenario 1 ("Baseline") we simulated all relevant factors, as described below, at levels observed in the data collection period (2009-2020)." All factors include those affecting the human-caused mortality, right? There are numerous studies documenting a variable annual rate of lethal removal. There seem to me to be other issues with assuming a constant annual lethal removal rate in baseline and the scenario for increased removals below.

The assumptions that seem to be made about constant annual lethal removal in the baseline or the increased removal scenarios might be summed up as "livestock losses will never get better or worse so long as the current rate of removal is applied randomly to wolf packs and entire packs are removed." I don't mean to caricature the assumption, I mean to make it plainer so it can be scrutinized.

1. If lethal removal is assumed to be effective in preventing livestock loss as WDFW has implied in the past, then it seems surprising that the model would treat it as ineffective or needing constant renewal. Can this be justified scientifically and by reference to articles that have not themselves been undermined by subsequent work? I call your attention to recent reviews of the literature on lethal removal which indicates unpredictable effects of lethal removal of wolves, resulting in increases, decreases and no change in livestock losses depending on study and site and years (the latter of no effect in the majority of cases, see studies of wolf removal by Grente, Krofel, and Santiago-Ávila.

2. Is predation on livestock random? If not, how does the imposition of a random scheme affect the model (a sensitivity analysis would be useful); many studies reveal that predation on livestock is not spatially random or uniform. Rather livestock losses are sometimes highly predictable from spatial features and wolf pack demographics. Therefore, I also call your attention to risk models that are analogous to resource selection functions, which have been used to model livestock loss in our region among others (see my lab website and search for "risk" and "forecast" please).

3. Has WDFW lethal removal eliminated entire packs and in what percentage of cases? This baseline information might be helpful in interpreting the scenarios. I discuss partial or entire pack removal further below.

I was confused by the increased removals scenario and the harvest scenario. Given they are differentiated I have to assume increased removals is NOT public hunting, trapping, hounding, etc. It is unclear what conditions might lead to such an increase in lethal removals. The authors wrote "In scenarios 4 and 5 ("Increased removals"), we simulated an increased number of lethal management removals such that 30% of the wolf population[*] would be removed every four years, corresponding to an annual removal rate of 8.5%." Does this replace the baseline removal rate or supplement it? I didn't see a scientific justification for the value of 30% and I don't understand where 8.5% came from (30 /4 = 7.5%). Even if I add the baseline it does not reach 8.5%. I'm sure I'm missing something but the calculation could be clarified.

Another concern about this scenario is that it uses a flat mortality rate (% of population) regardless of conditions. That seems to simulate population reduction (sometimes called culling) but applied randomly to entire packs. Given that is a highly unusual pattern of management, it would help to understand the rationale behind it. See below where other more common scenarios are NOT considered. Therefore, I do not understand the criteria applied when selecting scenarios that deserve modeling and scenarios that do not deserve modeling.

"Harvest"
See issues with terminology in the section on Minor comments below.
Every 6 months: This is an unusual off-take pattern. Readers may be tempted to assume that the policy-makers among the authors or their superiors in state agencies are planning two seasons of wolf-killing per year. The authors might wish to address why such an unusual wolf-killing system was included in this paper. Also, the method that allows only adults or juveniles yet simulates twice-a-year 'harvest' assumes the public can avoid killing pups. Is there evidence for that assumption? The assumption seems dubious on its face but regardless it requires some consideration of methods of 'harvest' and accidental non-target killing.

Additive: While this is more conservative than any compensatory scenarios, it still does not acknowledge the many sources of evidence for super-additive mortality when the public begins killing wolves: Creel, Vucetich, Chapron, or when wolf-killing is liberalized in general: Santiago-Avila, Louchouarn, Suutarinen, Liberg, Treves. There are now more than ten studies quantifying the super-additive effects on population dynamics or the undocumented losses of wolves when killing is liberalized (I.e., undocumented deaths that can be attributed to policies of liberalized killing).

The OMISSION of any alternative scenario with super-additive mortality and the OMISSIOn of alternative scenarios with increases in illegal killing triggered by the harvest and increased removals scenarios are problematic. I capitalized the word OMISSION to emphasize that they are not scientific decisions but value-based decisions about which scenarios to publish and which not to publish.

Value-based decisions are akin to unstated assumptions derived from personal or organizational preferences / beliefs / policies. Assumptions about parameter values or interactions between variables should be transparently stated and usually justified scientifically. Unstated assumptions in a modeling paper seem to me to be scientific missteps because the range of possible parameter values was circumscribed for reasons that are not transparent or justified by peer-reviewed research.

Also, please note that an attempt to scientifically justify circumscribed parameter values might require an even-handed summary of evidence for and against the assumed constraints on parameter values. For example, the increased removal scenarios (currently unjustified) might be paired with a lowered removal scenario or a scenario that curbs ongoing mortality sources such as poaching or vehicle collisions, hypothetically. To me it seems easier to evaluate alternative scenarios even-handedly than to justify the current ones.

Furthermore, my concern is that the decisions about which scenarios to publish in the current manuscript leave unanswered 'why these scenarios and not others?' And the authors do not touch upon alternative scenarios for how wolf-human coexistence might play out differently. Instead, the scenarios presented in this paper are a subset of wolf-human coexistence and that subset is slanted towards negative views of wolves (more killing). For example, there is nothing scientific telling us to simulate lethal removal at level x or y. We explored this problem in sustainable use models in Frontiers in Conservation Science in 2021.

My criticism is meant to be constructive as it is not too late to adapt your models to positive wolf-human coexistence scenarios, such as those involving provisioning to improve wolf reproduction or survival, increasing wild prey bases in regions with low prey, better enforcement against unregulated, human-caused mortality, use of non-lethal methods to protect livestock etc. I understand WDFW might never undertake such actions but that does not constrain scientists seeking approximations of reality. Also, administrations change, private actors / organizations sometimes step in, and background conditions change especially for a simulation run for 50 years.

I hope you see how a subset of scenarios was presented for non-scientific reasons.

Please remind readers that the selection of scenarios is value-based not science-based. Moreover, the selection of parameters within scenarios may also be value-based. For example, partial pack removals — simulated in your methods when "excess" removals are randomly assigned to another pack short of full pack removal — is NOT suggested to be effective in any study, even Bradley et al. 2015. Moreover, can the latter study even be used to justify the effectiveness of removal of entire wolf packs? I don’t think so. Consider that Santiago-Ávila et al. 2018 showed Bradley et al. 2015 was not reproducible until and unless the methods are clarified. Also, the 2018 article identified a possible statistical bias favoring lethal removal. If the data were to be shared (another hallmark of reproducibility), the bias minimized, and the methods clarified, one might argue that full pack removal has a scientific basis. But we’re not there yet.

Because I noticed omissions of scenarios and circumscribed parameter values without explicit statement of assumptions and missing literature, I offer a comment on potentially competing interests.

The scientific community has changed position on this in recent years and is increasingly recognizing the potentially distorting effects of values and ideology on scientific research. Nothing is necessarily disqualifying but all should be disclosed fully and transparently.

Ideological commitments expressed through memberships in civil society and professional societies (e.g., TWS or AFWA), institutional policy positions (e.g., WDFW's current policies), and personal affiliations or rivalries, might all place pressures on individuals that reflect competing interests. These can affect the unstated assumptions, literature reviews (what is cited and summarized versus omitted) and the methods chosen and analyses used, in addition to the traditional issues relating to financial interests. I am not referring to one or two articles being missed but a pattern of omitting peer-reviewed research in highly ranked international journals as I noted here. I emphasize the issue of potentially competing interests as a way to inspire greater public confidence in the scientific endeavor. Thanks for your kind attention.

Again I admire your decision to publish preprints so that pre-publication review has an opportunity to influence the future manuscript and perhaps public policy.

Minor concerns
Terminology:
The term "recovery" has a meaning in US federal and state endangered species law as you all no doubt are aware. Recovery in its legal sense may lead policy-makers to shift regulatory schemes to down- or delist wolves. Therefore it is not a value-neutral scientific term and could be viewed as prejudicial. I see passages in your text where recover(y) is appropriate but others where it was used to refer to recolonization or population growth. There I recommend instead using recolonizing or geographic spread or numerical rebound which do not imply a legal status. This seems especially relevant when scenario outcomes suggest a low likelihood of achieving legal recovery.

Relatedly, I recommend careful consideration of certain jargon words that may be mainstream in wildlife management but are not commonplace in ecological sciences or policy among all publics – and may have value-based or moral connotations, e.g., harvest and depredation. In place of harvest I suggest "permitted, regulated wolf-killing by the public", because harvest is a euphemism that holds implicit assumptions about the values of wolves and motivations of humans who participate. To see why not to use 'depredation', look at the first definition in the Oxford English Dictionary. I used it for years but now see the error.

Finally, the discussion of non-lethal methods might benefit from updating to include studies since 2010 on livestock-guarding dogs, and systematic reviews of effectiveness 2016-2021.