Elsevier Editorial System(tm) for Science of the Total Environment
Manuscript Draft

Manuscript Number: STOTEN-D-16-00847R1

Title: Predicting Environmental Mitigation Requirements for Hydropower Projects through the Integration of Biophysical and Socio-Political Geographies

Article Type: Research Paper

Keywords: hydropower; mitigation; modeling; prediction; environmental; sociopolitical

Corresponding Author: Mr. Christopher DeRolph,

Corresponding Author's Institution: Oak Ridge National Laboratory

First Author: Christopher DeRolph

Order of Authors: Christopher DeRolph; Michael P Schramm; Mark S Bevelhimer

Abstract: ABSTRACT
Uncertainty about environmental mitigation needs at existing and proposed hydropower projects makes it difficult for stakeholders to minimize environmental impacts. Hydropower developers and operators desire tools to better anticipate mitigation requirements, while natural resource managers and regulators need tools to evaluate different mitigation scenarios and order effective mitigation. Here we sought to examine the feasibility of using a suite of multi-faceted explanatory variables within a spatially explicit modeling framework to fit predictive models for future environmental mitigation requirements at hydropower projects across the conterminous U.S. Using a database comprised of mitigation requirements from more than 300 hydropower project licenses, we were able to successfully fit models for nearly 50 types of environmental mitigation and to apply the predictive models to a set of more than 500 non-powered dams identified as having hydropower potential. The results demonstrate that mitigation requirements are functions of a range of factors, from biophysical to socio-political. Project developers can use these models to inform cost projections and design considerations, while regulators can use the models to more quickly identify likely environmental issues and potential solutions, hopefully resulting in more timely and more effective decisions on environmental mitigation.

Response to Reviewers: 1. Reviewer #1:
This paper tries to develop a model for forecasting hydropower development environmental mitigation requirements.The idea is interesting, but the goal is not clear. If the model focus on a small area or watershed, it will be useful for hydropower development planning and environmental mitigation. Specifically, the paper also has some relatively large defects.
1.1. data processing: how the original data be processed as input of these models. Basic characteristics of power plant must be descipted is detai, also the mitigation requirement.

• Author response 1: the basic characteristics of a power plant are described in detail in many other reference materials and we do not feel it is prudent to add details on power plant characteristics to this manuscript. The mitigation requirements are described in detail in an appendix to a recently published article (see Schramm et al. 2016) on development of the mitigation database by the same authors of this manuscript. A reference to this paper and appendix has been added to Section 2.1.

1.2. evaluation the model preformance.The author believe that reducing the prediction variables is the highlight of this paper, but didn't explain clearly what variables are removed, why get rid of, the differences between different variables in this model?

• Author response 2: we don’t fully understand this comment, but have responded to the best of our ability as follows. Section 2.2 of MATERIALS AND METHODS explains our process of using a literature review in conjunction with expert opinion to select predictor variable sets for each of the six main Tier 1 mitigation categories, and we feel this is sufficiently explained in the existing text. Table 1 shows which model(s) each candidate predictor was included in. In Section 3.2 of RESULTS, we explain our implementation of Parisien et al.’s (2011) approach of identifying a variable as important if its relative influence as output by the BRT model was >=5%. We then summarize the important variables in a couple of different ways to elucidate the frequency and magnitude of the important variables across all of our models.

1.3. lacking of the model test results.Difference among these model results should be given a reasonable explanation.Does the change of the predicted result consistent with the actual results? Which requirements are involved in this mode in right ways?

• Author response 3: we don’t fully understand this comment, but have responded to the best of our ability as follows. As described in Section 2.3, we split the binary presence/absence data for each mitigation that we modeled into training (80%) and validation (20%) datasets. We then used the ROC value calculated on the validation dataset to assess model predictive performance. This process estimates model fit on a dataset that was not used to build the model. Given the novelty of the mitigation database, we were unable to obtain an independent validation dataset so we split the data and calculated model performance on a separate validation dataset.

1.4. scales: the mitigation requirements will change with time and space, most environmental impact mitigation requirements may come from local groups, but the model didnot consider the change of space scale.

• Author response 4: drivers of mitigation requirements operates at different scales, and we did our best to capture that with local, state, and regional variables. The ‘Spatial Scale’ column in Table 1 shows the variety of scales that the predictors were derived at. We also added some text on possible limitations to our approach resulting from disparity of scales between mitigation requirements and some of the predictor variables to the last paragraph of the DISCUSSION.

1.5. as a conclusion, all factors influent the requirement does not make sense.

• Author response 5: the guide for authors for this journal asks that authors present concise conclusions in a narrative from about what was found and what it means in the broader scheme. We feel that our study shows that a multitude of physical and social landscape factors are important contributors to required mitigation, and that this is an area that needs further research to assess causality of different drivers. We feel this is an important high-level finding that needs to be included in our concise CONCLUSIONS section.

2. Reviewer #2:
Graphical Abstract:
2.1. There is a step in the last line (between "Planning" and "Recreational management...". I think, it is not necessary.

- Author response 6: I understand how this could be confusing. The Tier 3 categories shown on the graphical abstract are supposed to only include Tier 3 mitigations that were included in our modeling effort described in this manuscript. However, there are over 130 Tier 3 categories in the mitigation database compiled by Schramm, Bevelhimer, and DeRolph (2016), some of which were not modeled as part of this manuscript because they occurred at less than 5% of the plants in the database. So, we included the 'Planning' Tier 2 category as a separate category because we wanted to maintain the hierarchical structure of the Schramm database and to show how the Tier 1 categories were divided into Tier 2, and Tier 2 into Tier 3.

2.2. Maybe the editing of the "Predictor variables" could be improved a bit. Sometimes the text fields are not centered.

- Author response 7: I improved the horizontal and vertical centering of the 'Predictor variables' portion of the diagram.

Introduction:
2.3. It would be fine if you could integrate some information on similar scientific approaches (incl. references) evaluating environmental mitigation.

- Author response 8: to our knowledge, there aren't other similar studies on predicting environmental mitigation. As stated in the first sentence of the DISCUSSION, the spatial modeling approach developed here integrates GIS techniques, novel data, machine-learning algorithms, and niche modeling concepts common in landscape ecology. We added a reference to this sentence to a paper by Guisan and Thuiller (2005) that is an oft-cited reference on ecological niche modeling that can serve as a nice reference for readers that are unfamiliar with element distribution modeling concepts.

Materials and Methods:
2.4. I think, the readability of Table 1 could be improved a bit.

- Author response: see Author response 9.

Results:
2.5. The readability of Table 2 should be improved too. It is a bit too much information for one Table. Could you summarize some data?

- Author response 9: we recognize that there is a lot of information presented in Tables 1 and 2, and that's the reason we used a color scheme for the variable categories - to add a way for the reader to visually assess what types of variables are important for different categories of models without having to focus on specific predictors unless they desire to do so. We used an application called COLORBREWER http://colorbrewer2.org/ to identify a color scheme for the different variable categories that would allow readers to easily distinguish between them, and then we carried this color scheme throughout the paper to maintain this additional visual dimension of the data. Great care was taken to distill the information presented in Tables 1 and 2 to the most basic but necessary information to reproduce our study, and we adopted data presentation approaches from other similar modeling studies to accomplish this.

2.6. In my opinion, Figure 3 looks like a table. Moreover, the listed designations of the variables have a cryptic character (e.g. MAVELU). The understandability of the illustrated information could be improved by changing / clarifying the wording.

• Author response 10: we understand that Figure 3 kind of looks like a table, but it is not a table so we would like to leave it as a figure. There is no relation between the information across rows, and we feel it is better left as a figure and not converted to a table. The variable names have to be abbreviated somehow (this is quite common in other similar papers – see Arganaraz et al. in STOTEN 2015 for an example), and we did our best to use logical abbreviations that make sense at least after reading the variable description.

2.7. When printing the article in black and white, the different colours used in the tables and figures are sometimes looking quite similar. Could you please improve this too?

• Author response 11: as stated in Author response 9, we used an application called COLORBREWER http://colorbrewer2.org/ to identify a color scheme for the different variable categories that would allow readers to easily distinguish between them, and then we carried this color scheme throughout the paper to maintain this additional visual dimension of the data. If a reader is printing in black and white, all of the essential information is still decipherable but the added visual dimension of the color scheme is lost. There is really no way to use a gray scale to add this visual dimension, so we feel it is best to leave the color scheme as is for readers that print in color or view the article on-screen.

Discussion:

2.8. I think, there are two spaces on page 20 (24/25) between "autocorrelation." and "Since".

• Author response 12: Thanks for catching this. We fixed this error.

Conclusions:

2.9. This chapter is very short. Is it possible to draw some additional conclusions?

• Author response 13: the guidance on conclusions in the Guide for Authors for this journal states that 1) The abstract, highlights and conclusions of papers in this journal must contain clear and concise statements as to why the study was done and how readers will benefit from the results, 2)the main conclusions of the study may be presented in a short Conclusions section, which may stand alone or form a subsection of a Discussion or Results and Discussion section, and 3)the conclusions should be a narrative about what you found and what it means in the broader scheme. We feel there is significant detail in the DISCUSSION section on each of the points we make in our CONCLUSIONS and that we closely followed the requirements in the Guide for Authors on conclusions.

3. Reviewer #3:

The authors made an effort to build a predictive model by coupling it with biophysical to socio-political factors, for future environmental mitigation requirements at hydropower projects across the conterminous U.S. It would be meaningful for hydropower project developers, regulators to inform and identify the environmental issues and potential solutions on environmental mitigation. However, some adjustments and revisions are necessary. More detailed comments are as follows:

3.1. Page 4-6, the section 2.1 is missing.

• Author response 14: Thanks for catching this. We updated the section headings. The old 2.2 is the new 2.1, and so on.

3.2. In the explanatory variables, the human dimensions, who are the human, the local residents, hydropower developers, or all public? Since the different stakeholders may have different survey responses, the author should make it clear.

• Author response 15: we updated the text in Section 2.2.3 to add clarification. The aim of the human dimensions variable category was to consider variables that could be derived for the whole U.S. that would describe some aspect of the socio-political or regulatory landscape thought to be important in hydropower mitigation deliberations.

3.3. I have some confused about the relationship between the response variables (biodiversity, fish passage, habitat, hydrology, recreation, and water quality) and the explanatory variables (biological, facility, human, hydrologic, landscape, locational, and stream network characteristics). The author may introduce how these variables are contributed to your model in the beginning of section 2.

• Author response 16: we don't fully understand this comment, but have responded to the best of our ability as follows. We feel that Section 2.2 Explanatory Variables describes our variable selection process in detail, including literature references that support our approach and table references that convey our variable selections.

3.4. In this study, "predictive models were built only if a mitigation type was required for at least 5% of the plants in the mitigation database", why choose 5%? Are there some similar studies or designed in this study?

• Author response 17: we adopted the 5% approach from Rickbeil et al. 2014, but forgot to include this reference. The text in Sections 2.1 and 3.1 has been updated with this reference.

3.5. Some uncertainties and limitations about the predictive model and predictor database should be analyzed in the manuscript.

• Author response 18: We agree and added some additional information on limitations of our modeling approach to the last paragraph of the DISCUSSION.

3.6. The authors may refer to some latest publications:
Bing Yu, Linyu Xu. Review of ecological compensation in hydropower development. Renewable & Sustainable Energy Reviews, 2016,55:729-738;
Bing Yu, Linyu Xu, Zhifeng Yang. Ecological compensation for inundated habitats in hydropower developments based on carbon stock balance. Journal of Cleaner Production, 2016,114:334-342.

• Author response 19: Thank you for these recommendations. Reference added for Yu and Xu 2016 to paragraph 3 of INTRODUCTION. Reference added for Yu et al. to paragraph 5 of DISCUSSION.

4. Reviewer #4:
Even if the argument of the paper about the prediction of environmental mitigation requirements of hydropower projects is very relevant due to the renaissance of the global interests on dams' development, in my opinion the paper in the current version needs major changes before it can be considered for publication (see specific comments below). An improved explanation of the variables used and how these can inform decision makers is required to better understand the practical implications of the model to inform decision makers on specific interventions useful to mitigate the environmental and social impacts of dams. Specific comments:

4.1. The authors use impacts on recreational opportunity and humans as if they are complementary. In the introduction (line 24, page 3), recreational opportunity is considered as one of the aspects which need to be included together with wildlife and environmental quality in the identification of mitigation strategies. Later on in the introduction (line 32, page 4) the human dimension is introduced. However the authors did not provide an explanation of what they mean about recreational opportunity and how this is linked to the analysis of the human dimension. Moreover, in line 56 in the introduction the authors use recreational resources; however no explanation of recreational resources is provided. The authors should provide in the introduction a brief explanation of what they mean about recreational resources and the human dimension and explicitly explain how these are linked together.

- Author response 20: we don’t fully understand this comment, but have responded to the best of our ability as follows. The phrase ‘recreation opportunity’ is a phrase used in the Electric Consumers Protection Act of 1986 that we simply adopted. Recreation is one of six main categories of mitigation requirements, while human dimensions is one of seven main categories of potential predictor variables used to predict mitigation requirements. Given the goal of our study was to predict numerous mitigation requirements across each of the six main mitigation categories, we do not feel it is warranted to place particular emphasis in the INTRODUCTION on the relationship between recreation or human dimensions. Appendix B presents partial dependence plots for the three variables with the highest relative influence for each model, which allows for examination of the direction of variable influence for the most important variables in each model. So if a reader is interested in specific linkages between recreation mitigation and human dimensions variables, these plots can be explored.

4.2. In the introduction (line 5, page 4) the authors state "evaluating the environmental costs of hydropower projects with the variety of societal benefits hydropower projects provide" however the paper includes in its analysis the human dimension as well, wouldn't be better to refer to environmental and social costs?

- Author response 21: we agree and made a change to this sentence in the text.

4.3. Figure 2 in the materials and methods sections provides some variables to demonstrate the environmental and cultural heterogeneity of the conterminous United States. However it is not clear why the authors use the variables "average percent democratic votes" and "unemployment" to explain cultural heterogeneity. I don't think political votes can explain cultural diversity or different cultural characteristics. The use of these variables needs a clear explanation and justification.

- Author response 22: The intent of this set of maps was to provide visual examples of different social and physical landscape characteristics that are later used as mitigation predictor variables and that also demonstrate the heterogeneity of the U.S. social and physical landscape. Unemployment is a very common demographic characteristic derived from census data and is a common descriptor of social landscapes. As stated in Section 2.2.3, Kosnik (2010) found the largest influences on FERC’s regulatory decisions to be congressional politics and regulatory tendencies. Our collinearity analysis for all candidate predictors presented in Table 1 showed significant correlation between presidential voting and congressional politics at the state-scale, so we elected to include presidential political voting tendencies as a human-related predictor.

4.4. In Table 1 the authors report the input variables they used in their mitigation prediction model. Under the human dimension they report different variables which are not explained in the text and it is also not clear how they are linked to the human dimension of the mitigation strategies. For example: National Audubon Society chapters (what is this? not explained in the text, how is it relevant for predicting mitigation strategies linked to the human dimension of the impacts?); Dam removals (how is it relevant for predicting mitigation strategies linked to the human dimension of the impacts?); TU and CCA chapters (what is this? not explained in the text, how is it relevant for predicting mitigation strategies?); Sierra Club chapters (what is this? not explained in the text, how is it relevant for predicting mitigation strategies linked to the human dimension of the impacts?); usHouse and usSenate (how are these relevant for predicting mitigation strategies linked to the human dimension of the impacts?). A better explanation of the variables used and their relevance should be included in section 2.3.3 on human dimensions. Citations to other papers and analyses are included but it is not enough to understand how these variables are important for the specific study realized in the paper.

- Author response 23: thank you for this suggestion. We added significant text to the end of Section 2.2.3 to describe these different human dimensions variables.

4.5. Why recreation is not mentioned in Table 1 and it appears only in the results section (see table 2 model results summary)? Please explain how the variables used for the human dimension are linked with recreation.

- Author response 24: recreation is mentioned in Table 1. It is abbreviated as 'R' in the 'Models' column.

4.6. In the discussion and conclusion sections the paper lacks a clear discussion on how the variables, that from the implementation of the proposed model result to be relevant for the design of mitigation strategies, can actually be linked to specific mitigation strategies. Some examples of mitigation strategies linked to the human and biological variables identified as relevant by the application of the model could help the reader to better understand the practical implications of the model to inform decision makers on specific interventions useful to mitigate those impacts.

- Author response 25: As described in the last paragraph of the INTRODUCTION, our primary goal was to build statistical models to predict future mitigation requirements at hydropower project sites, while the secondary goal was to gain some understanding into potential key environmental and social drivers of these requirements that may warrant additional future research. In essence we have developed a screening tool that can be used by developers to reduce uncertainty in possible mitigation requirements and by regulators interested in identifying a preliminary portfolio of possible environmental issues and associated mitigation options. Further research is needed to establish robust links between specific explanatory variables, mitigation requirements, and mitigation strategies. We added the previous sentence to the text in the CONCLUSIONS to clarify the possible applications of the models and the limitations.

OAK RIDGE National Laboratory

PO Box 2008
Oak Ridge, TN 37831-6038
(865) 574-0974 | derolphcr@ornl.gov

11 May 2016

Dear Professor Pollard:

Please find attached our responses to reviewer comments and revision to manuscript STOTEN-D-16-00847 entitled "Predicting Environmental Mitigation Requirements for Hydropower Projects through the Integration of Biophysical and Socio-Political Geographies". We thank the editor and the reviewers for their attention to this manuscript. We believe that addressing the comments have led to a further improved manuscript. We have responded point-by-point to reviewer comments.

We look forward to continue working with you towards publication of this manuscript. Please feel free to contact me via telephone at 865-574-0974 or at the email address provided below if you require additional information or clarification.

Sincerely,

Christopher R. DeRolph, M.S., GISP
Geospatial Scientist

Authors: Christopher R. DeRolph (**first and corresponding author**), derolphcr@ornl.gov
Michael P. Schramm, mpschramm@gmail.com
Mark S. Bevelhimer, bevelhimerms@ornl.gov

***Responses to Reviewers Comments**

1. Reviewer #1:

   This paper tries to develop a model for forecasting hydropower development environmental mitigation requirements.The idea is interesting, but the goal is not clear. If the model focus on a small area or watershed, it will be useful for hydropower development planning and environmental mitigation. Specifically, the paper also has some relatively large defects.

   1.1. data processing: how the original data be processed as input of these models. Basic characteristics of power plant must be descipted is detai, also the mitigation requirement.

   - Author response 1: the basic characteristics of a power plant are described in detail in many other reference materials and we do not feel it is prudent to add details on power plant characteristics to this manuscript. The mitigation requirements are described in detail in an appendix to a recently published article (see Schramm et al. 2016) on development of the mitigation database by the same authors of this manuscript. A reference to this paper and appendix has been added to Section 2.1.

   1.2. evaluation the model preformance.The author believe that reducing the prediction variables is the highlight of this paper, but didn't explain clearly what variables are removed, why get rid of, the differences between different variables in this model?

   - Author response 2: we don't fully understand this comment, but have responded to the best of our ability as follows. Section 2.2 of MATERIALS AND METHODS explains our process of using a literature review in conjunction with expert opinion to select predictor variable sets for each of the six main Tier 1 mitigation categories, and we feel this is sufficiently explained in the existing text. Table 1 shows which model(s) each candidate predictor was included in. In Section 3.2 of RESULTS, we explain our implementation of Parisien et al.'s (2011) approach of identifying a variable as important if its relative influence as output by the BRT model was >=5%. We then summarize the important variables in a couple of different ways to elucidate the frequency and magnitude of the important variables across all of our models.

   1.3. lacking of the model test results.Difference among these model results should be given a reasonable explanation.Does the change of the predicted result consistent with the actual results? Which requirements are involved in this mode in right ways?

   - Author response 3: we don't fully understand this comment, but have responded to the best of our ability as follows. As described in Section 2.3, we split the binary presence/absence data for each mitigation that we modeled into training (80%) and validation (20%) datasets. We then used the ROC value calculated on the validation dataset to assess model predictive performance. This process estimates model fit on a dataset that was not used to build the model. Given the novelty of the mitigation database, we were unable to obtain an independent validation dataset so we split the data and calculated model performance on a separate validation dataset.

   1.4. scales: the mitigation requirements will change with time and space, most environmental impact mitigation requirements may come from local groups, but the model didnot consider the change of space scale.

   - Author response 4: drivers of mitigation requirements operates at different scales, and we did our best to capture that with local, state, and regional variables. The 'Spatial Scale' column in Table 1 shows the variety of scales that the predictors were derived at. We also added some text on possible limitations to our approach resulting from disparity of scales between mitigation requirements and some of the predictor variables to the last paragraph of the DISCUSSION.

   1.5. as a conclusion, all factors influent the requirement does not make sense.

- Author response 5: the guide for authors for this journal asks that authors present concise conclusions in a narrative from about what was found and what it means in the broader scheme. We feel that our study shows that a multitude of physical and social landscape factors are important contributors to required mitigation, and that this is an area that needs further research to assess causality of different drivers. We feel this is an important high-level finding that needs to be included in our concise CONCLUSIONS section.

2. Reviewer #2:

Graphical Abstract:

2.1. There is a step in the last line (between "Planning" and "Recreational management...". I think, it is not necessary.

- Author response 6: I understand how this could be confusing. The Tier 3 categories shown on the graphical abstract are supposed to only include Tier 3 mitigations that were included in our modeling effort described in this manuscript. However, there are over 130 Tier 3 categories in the mitigation database compiled by Schramm, Bevelhimer, and DeRolph (2016), some of which were not modeled as part of this manuscript because they occurred at less than 5% of the plants in the database. So, we included the 'Planning' Tier 2 category as a separate category because we wanted to maintain the hierarchical structure of the Schramm database and to show how the Tier 1 categories were divided into Tier 2, and Tier 2 into Tier 3.

2.2. Maybe the editing of the "Predictor variables" could be improved a bit. Sometimes the text fields are not centered.

- Author response 7: I improved the horizontal and vertical centering of the 'Predictor variables' portion of the diagram.

Introduction:

2.3. It would be fine if you could integrate some information on similar scientific approaches (incl. references) evaluating environmental mitigation.

- Author response 8: to our knowledge, there aren't other similar studies on predicting environmental mitigation. As stated in the first sentence of the DISCUSSION, the spatial modeling approach developed here integrates GIS techniques, novel data, machine-learning algorithms, and niche modeling concepts common in landscape ecology. We added a reference to this sentence to a paper by Guisan and Thuiller (2005) that is an oft-cited reference on ecological niche modeling that can serve as a nice reference for readers that are unfamiliar with element distribution modeling concepts.

Materials and Methods:

2.4. I think, the readability of Table 1 could be improved a bit.

- Author response: see Author response 9.

Results:

2.5. The readability of Table 2 should be improved too. It is a bit too much information for one Table. Could you summarize some data?

- Author response 9: we recognize that there is a lot of information presented in Tables 1 and 2, and that's the reason we used a color scheme for the variable categories - to add a way for the reader to visually assess what types of variables are important for different categories of models without having to focus on specific predictors unless they desire to do so. We used an application called COLORBREWER http://colorbrewer2.org/ to identify a color scheme for the different variable categories that would allow readers to easily distinguish between them, and then we carried this color scheme throughout the paper to maintain this additional visual dimension of the data. Great care was taken to distill the information presented in Tables 1 and 2 to the most basic but necessary information to reproduce our study, and we adopted data presentation approaches from other similar modeling studies to accomplish this.

2.6. In my opinion, Figure 3 looks like a table. Moreover, the listed designations of the variables have a cryptic character (e.g. MAVELU). The understandability of the illustrated information could be improved by changing / clarifying the wording.

- Author response 10: we understand that Figure 3 kind of looks like a table, but it is not a table so we would like to leave it as a figure. There is no relation between the information across rows, and we feel it is better left as a figure and not converted to a table. The variable names have to be abbreviated somehow (this is quite common in other similar papers – see Arganaraz et al. in STOTEN 2015 for an example), and we did our best to use logical abbreviations that make sense at least after reading the variable description.

2.7. When printing the article in black and white, the different colours used in the tables and figures are sometimes looking quite similar. Could you please improve this too?

- Author response 11: as stated in Author response 9, we used an application called COLORBREWER http://colorbrewer2.org/ to identify a color scheme for the different variable categories that would allow readers to easily distinguish between them, and then we carried this color scheme throughout the paper to maintain this additional visual dimension of the data. If a reader is printing in black and white, all of the essential information is still decipherable but the added visual dimension of the color scheme is lost. There is really no way to use a gray scale to add this visual dimension, so we feel it is best to leave the color scheme as is for readers that print in color or view the article on-screen.

Discussion:

2.8. I think, there are two spaces on page 20 (24/25) between "autocorrelation." and "Since".

- Author response 12: Thanks for catching this. We fixed this error.

Conclusions:

2.9. This chapter is very short. Is it possible to draw some additional conclusions?

- Author response 13: the guidance on conclusions in the Guide for Authors for this journal states that 1) *The abstract, highlights and conclusions of papers in this journal must contain clear and concise statements as to why the study was done and how readers will benefit from the results, 2)the main conclusions of the study may be presented in a short Conclusions section, which may stand alone or form a subsection of a Discussion or Results and Discussion section, and 3)the conclusions should be a narrative about what you found and what it means in the broader scheme.* We feel there is significant detail in the DISCUSSION section on each of the points we make in our CONCLUSIONS and that we closely followed the requirements in the Guide for Authors on conclusions.

3. Reviewer #3:

The authors made an effort to build a predictive model by coupling it with biophysical to socio-political factors, for future environmental mitigation requirements at hydropower projects across the conterminous U.S. It would be meaningful for hydropower project developers, regulators to inform and identify the environmental issues and potential solutions on environmental mitigation. However, some adjustments and revisions are necessary. More detailed comments are as follows:

3.1. Page 4-6, the section 2.1 is missing.

- Author response 14: Thanks for catching this. We updated the section headings. The old 2.2 is the new 2.1, and so on.

3.2. In the explanatory variables, the human dimensions, who are the human, the local residents, hydropower developers, or all public? Since the different stakeholders may have different survey responses, the author should make it clear.

- Author response 15: we updated the text in Section 2.2.3 to add clarification. The aim of the human dimensions variable category was to consider variables that could be derived for the whole U.S. that would describe some aspect of the socio-political or regulatory landscape thought to be important in hydropower mitigation deliberations.

3.3. I have some confused about the relationship between the response variables (biodiversity, fish passage, habitat, hydrology, recreation, and water quality) and the explanatory variables (biological, facility, human, hydrologic, landscape, locational, and stream network characteristics). The author may introduce how these variables are contributed to your model in the beginning of section 2.

- Author response 16: we don't fully understand this comment, but have responded to the best of our ability as follows. We feel that Section 2.2 Explanatory Variables describes our variable selection process in detail, including literature references that support our approach and table references that convey our variable selections.

3.4. In this study, "predictive models were built only if a mitigation type was required for at least 5% of the plants in the mitigation database", why choose 5%? Are there some similar studies or designed in this study?

- Author response 17: we adopted the 5% approach from Rickbeil et al. 2014, but forgot to include this reference. The text in Sections 2.1 and 3.1 has been updated with this reference.

3.5. Some uncertainties and limitations about the predictive model and predictor database should be analyzed in the manuscript.

- Author response 18: We agree and added some additional information on limitations of our modeling approach to the last paragraph of the DISCUSSION.

3.6. The authors may refer to some latest publications:
Bing Yu, Linyu Xu. Review of ecological compensation in hydropower development. Renewable & Sustainable Energy Reviews, 2016,55:729-738;
Bing Yu, Linyu Xu, Zhifeng Yang. Ecological compensation for inundated habitats in hydropower developments based on carbon stock balance. Journal of Cleaner Production, 2016,114:334-342.

- Author response 19: Thank you for these recommendations. Reference added for Yu and Xu 2016 to paragraph 3 of INTRODUCTION. Reference added for Yu et al. to paragraph 5 of DISCUSSION.

4. Reviewer #4:
Even if the argument of the paper about the prediction of environmental mitigation requirements of hydropower projects is very relevant due to the renaissance of the global interests on dams' development, in my opinion the paper in the current version needs major changes before it can be considered for publication (see specific comments below). An improved explanation of the variables used and how these can inform decision makers is required to better understand the practical implications of the model to inform decision makers on specific interventions useful to mitigate the environmental and social impacts of dams. Specific comments:

4.1. The authors use impacts on recreational opportunity and humans as if they are complementary. In the introduction (line 24, page 3), recreational opportunity is considered as one of the aspects which need to be included together with wildlife and environmental quality in the identification of mitigation strategies. Later on in the introduction (line 32, page 4) the human dimension is introduced. However the authors did not provide an explanation of what they mean about recreational opportunity and how this is linked to the analysis of the human dimension. Moreover, in line 56 in the introduction the authors use recreational resources; however no explanation of recreational resources is provided. The authors should provide in the introduction a brief explanation of what they mean about recreational resources and the human dimension and explicitly explain how these are linked together.

- Author response 20: we don’t fully understand this comment, but have responded to the best of our ability as follows. The phrase ‘recreation opportunity’ is a phrase used in the Electric Consumers Protection Act of 1986 that we simply adopted. Recreation is one of six main categories of mitigation requirements, while human dimensions is one of seven main categories of potential predictor variables used to predict mitigation requirements. Given the goal of our study was to predict numerous mitigation requirements across each of the six main mitigation categories, we do not feel it is warranted to place particular emphasis in the INTRODUCTION on the relationship between recreation or human dimensions. Appendix B presents partial dependence plots for the three variables with the highest relative influence for each model, which allows for examination of the direction of variable influence for the most important variables in each model. So if a reader is interested in specific linkages between recreation mitigation and human dimensions variables, these plots can be explored.

4.2. In the introduction (line 5, page 4) the authors state "evaluating the environmental costs of hydropower projects with the variety of societal benefits hydropower projects provide" however the paper includes in its analysis the human dimension as well, wouldn't be better to refer to environmental and social costs?

- Author response 21: we agree and made a change to this sentence in the text.

4.3. Figure 2 in the materials and methods sections provides some variables to demonstrate the environmental and cultural heterogeneity of the conterminous United States. However it is not clear why the authors use the variables "average percent democratic votes" and "unemployment" to explain cultural heterogeneity. I don't think political votes can explain cultural diversity or different cultural characteristics. The use of these variables needs a clear explanation and justification.

- Author response 22: The intent of this set of maps was to provide visual examples of different social and physical landscape characteristics that are later used as mitigation predictor variables and that also demonstrate the heterogeneity of the U.S. social and physical landscape. Unemployment is a very common demographic characteristic derived from census data and is a common descriptor of social landscapes. As stated in Section 2.2.3, Kosnik (2010) found the largest influences on FERC’s regulatory decisions to be congressional politics and regulatory tendencies. Our collinearity analysis for all candidate predictors presented in Table 1 showed significant correlation between presidential voting and congressional politics at the state-scale, so we elected to include presidential political voting tendencies as a human-related predictor.

4.4. In Table 1 the authors report the input variables they used in their mitigation prediction model. Under the human dimension they report different variables which are not explained in the text and it is also not clear how they are linked to the human dimension of the mitigation strategies. For example: National Audubon Society chapters (what is this? not explained in the text, how is it relevant for predicting mitigation strategies linked to the human dimension of the impacts?); Dam removals (how is it relevant for predicting mitigation strategies linked to the human dimension of the impacts?); TU and CCA chapters (what is this? not explained in the text, how is it relevant for predicting mitigation strategies?); Sierra

Club chapters (what is this? not explained in the text, how is it relevant for predicting mitigation strategies linked to the human dimension of the impacts?); usHouse and usSenate (how are these relevant for predicting mitigation strategies linked to the human dimension of the impacts?). A better explanation of the variables used and their relevance should be included in section 2.3.3 on human dimensions. Citations to other papers and analyses are included but it is not enough to understand how these variables are important for the specific study realized in the paper.

- Author response 23: thank you for this suggestion. We added significant text to the end of Section 2.2.3 to describe these different human dimensions variables.

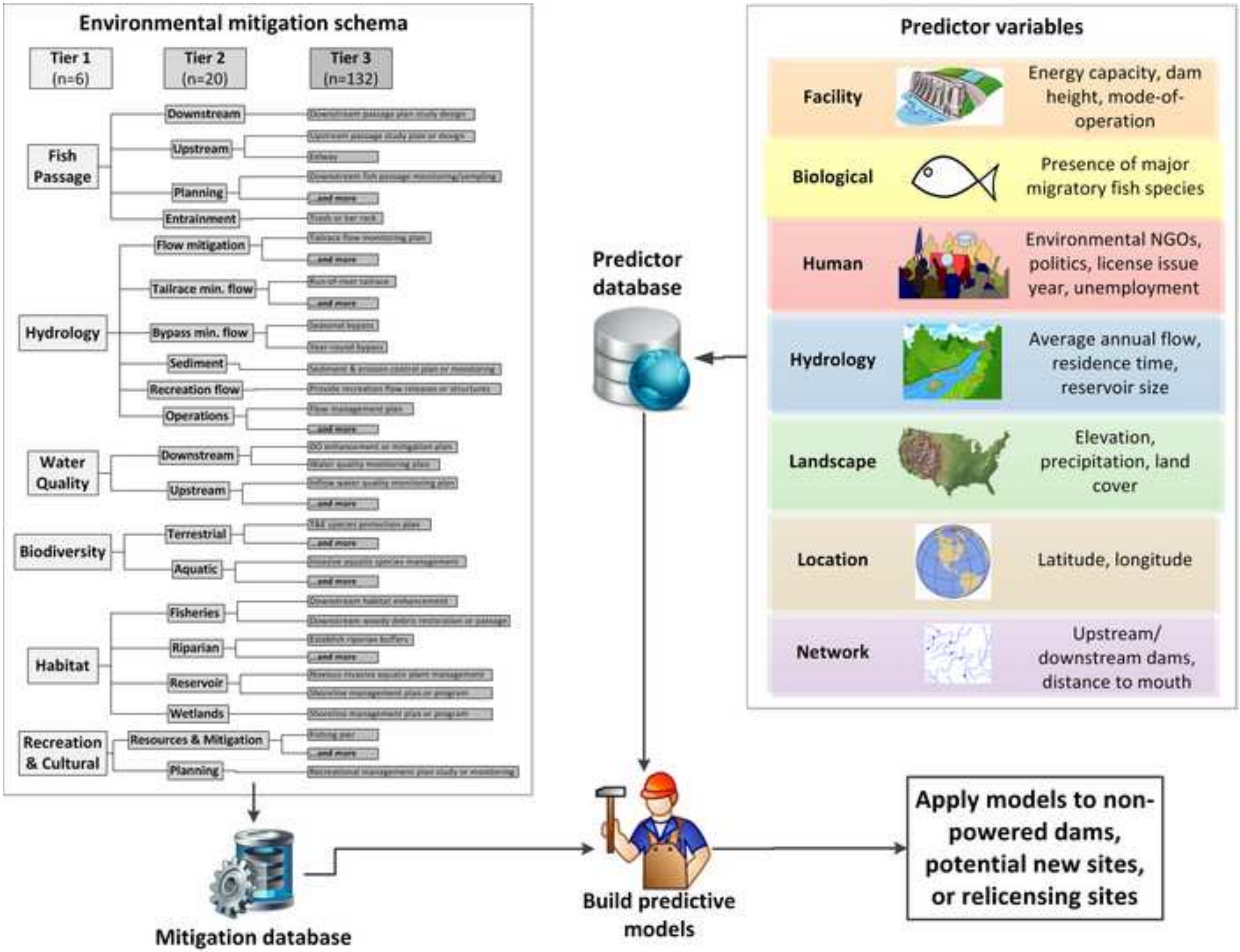
4.5. Why recreation is not mentioned in Table 1 and it appears only in the results section (see table 2 model results summary)? Please explain how the variables used for the human dimension are linked with recreation.

- Author response 24: recreation is mentioned in Table 1. It is abbreviated as 'R' in the 'Models' column.

4.6. In the discussion and conclusion sections the paper lacks a clear discussion on how the variables, that from the implementation of the proposed model result to be relevant for the design of mitigation strategies, can actually be linked to specific mitigation strategies. Some examples of mitigation strategies linked to the human and biological variables identified as relevant by the application of the model could help the reader to better understand the practical implications of the model to inform decision makers on specific interventions useful to mitigate those impacts.

- Author response 25: As described in the last paragraph of the INTRODUCTION, our primary goal was to build statistical models to predict future mitigation requirements at hydropower project sites, while the secondary goal was to gain some understanding into potential key environmental and social drivers of these requirements that may warrant additional future research. In essence we have developed a screening tool that can be used by developers to reduce uncertainty in possible mitigation requirements and by regulators interested in identifying a preliminary portfolio of possible environmental issues and associated mitigation options. Further research is needed to establish robust links between specific explanatory variables, mitigation requirements, and mitigation strategies. We added the previous sentence to the text in the CONCLUSIONS to clarify the possible applications of the models and the limitations.

***Graphical Abstract**



***Highlights (for review)**

- Statistical models were built to predict required mitigation at hydropower sites
- Predictors describe biosphere, lithosphere, hydrosphere, and anthroposphere
- BRT model results show biophysical and sociopolitical factors influence mitigation
- Hydropower developers can use models to reduce cost and inform design decisions
- Resource managers can use models to identify likely environmental issues

Predicting Environmental Mitigation Requirements for Hydropower Projects through the Integration of Biophysical and Socio-Political Geographies

Christopher R. DeRolph*[a], Michael P. Schramm[a], and Mark S. Bevelhimer[a]

[a] Oak Ridge National Laboratory, Environmental Sciences Division, PO Box 2008, 1 Bethel Valley Road, Oak Ridge, TN 37831-6038

*corresponding author: derolphcr@ornl.gov

ABSTRACT

Uncertainty about environmental mitigation needs at existing and proposed hydropower projects makes it difficult for stakeholders to minimize environmental impacts. Hydropower developers and operators desire tools to better anticipate mitigation requirements, while natural resource managers and regulators need tools to evaluate different mitigation scenarios and order effective mitigation. Here we sought to examine the feasibility of using a suite of multi-faceted explanatory variables within a spatially explicit modeling framework to fit predictive models for future environmental mitigation requirements at hydropower projects across the conterminous U.S. Using a database comprised of mitigation requirements from more than 300 hydropower project licenses, we were able to successfully fit models for nearly 50 types of environmental mitigation and to apply the predictive models to a set of more than 500 non-powered dams identified as having hydropower potential. The results demonstrate that mitigation requirements are functions of a range of factors, from biophysical to socio-political. Project developers can use these models to inform cost projections and design considerations, while regulators can use the models to more quickly identify likely environmental issues and potential solutions, hopefully resulting in more timely and more effective decisions on environmental mitigation.

Keywords hydropower, mitigation, modeling, prediction, environmental, sociopolitical

1. INTRODUCTION

Hydroelectric power is currently the largest of the renewable energy resources worldwide, contributing to electricity generation in 160 countries (Manzano-Agugliaro et al., 2013). The environmental impacts of hydropower are well established (Liermann et al., 2012; Nilsson et al., 2005; Poff et al., 1997; Poff et al., 2007), and are mitigated with mixed success (Trussart et al., 2002). In the United States (U.S.), the authority to issue 30-50 year licenses for the operation of non-federal hydropower facilities belongs to the U.S. Federal Energy Regulatory Commission (FERC). The passage of the Electric Consumers Protection Act of 1986 (ECPA) substantially changed FERC's consideration of environmental impacts with the requirement that equal consideration be given to the protection and enhancement of, and mitigation of damage to, wildlife, environmental quality, and recreational opportunity. Furthermore, a string of court rulings eroded FERC's singular authority to prescribe environmental mitigation by requiring FERC to include fishway prescriptions from the National Marine Fisheries Service or U.S. Fish and Wildlife Service, as well as minimum streamflow requirements included as part of state water quality certificates (Blumm and Nadol, 2001; Tarlock, 2012). The result was a significant increase in the number of mitigation requirements included in FERC licenses and a growing role of other federal and state agencies in the licensing process (Blumm and Nadol, 2001; Deshazo and Freeman, 2005; Kosnik, 2010).

FERC and the hydropower industry have suggested that this instable policy context and increased regulatory plurality have resulted in increased licensing time and increased uncertainty in mitigation requirements (FERC, 2001; U.S.-Congress, 2012). Original licenses for new projects and relicensing of existing projects provide a once in every 30 to 50 year opportunity to address environmental concerns at hydropower projects. FERC addresses potential environmental impacts by incorporating license conditions (mitigation requirements) where evidence shows project operations will impact environmental or recreational resources. With over 300 relicense applications anticipated between 2016 and 2026 (FERC, 2015), there is new urgency to integrate sustainability practices into future hydropower development by evaluating and balancing the environmental and social costs of hydropower projects with the variety of societal potential benefits hydropower projects provide.

Hydropower developers and owners desire some certainty and ability to better anticipate mitigation requirements. Similarly, resource managers and regulators must be able to evaluate likely mitigation scenarios and determine the relative effectiveness of mitigation implemented at similar projects. While each hydropower project is unique, Yu and Xu (2016) recommend development of common approaches and principles for designing ecological and social compensation mechanisms for hydropower development. The authors of this manuscriptA recently developed a database of environmental mitigation requirements in FERC licenses (Schramm et al., 2016) that presents new opportunities for analyzing past environmental mitigation requirements (Schramm et al., 2016)and predicting future mitigation requirements.

In this study we sought to examine the feasibility of using a suite of multidisciplinary explanatory variables to fit predictive models for environmental mitigation requirements at hydropower projects across the conterminous U.S. We developed a spatially explicit framework (applying niche modeling concepts common in landscape ecology) to predict nearly 50 types of environmental mitigation requirements using biological, facility, human, hydrologic, landscape, locational, and stream network characteristics. Our primary goal was to build statistical models to predict future mitigation requirements at hydropower project sites, while the secondary goal was to gain some understanding into potential key environmental and social drivers of these requirements that may warrant additional future research. As an example of how the models can be applied, we made predictions to a set of non-powered dams (NPDs) across the U.S. that were previously identified as having considerable energy potential.

2. MATERIALS AND METHODS

The conterminous United States (Fig. 1) is environmentally and culturally heterogeneous (Fig. 2), containing diverse physiographic regions ranging from mountains to inland and coastal plains, and encompassing examples of nearly every global climate. There is also considerable geographic variation in socio-political, economic, and cultural characteristics.

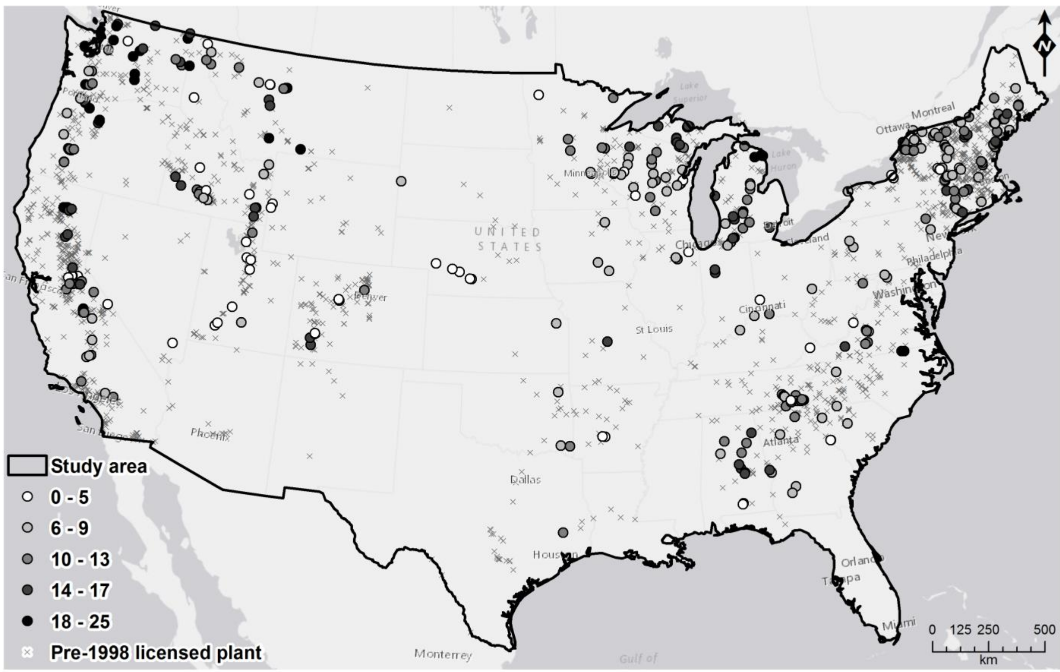


Fig. 1. Study area showing location of 463 hydropower plants licensed from 1998 through September 2015. Color of plant locations indicates number of mitigation requirements for mitigation categories selected for statistical modeling.

Formatted: Indent: First line: 0 cm

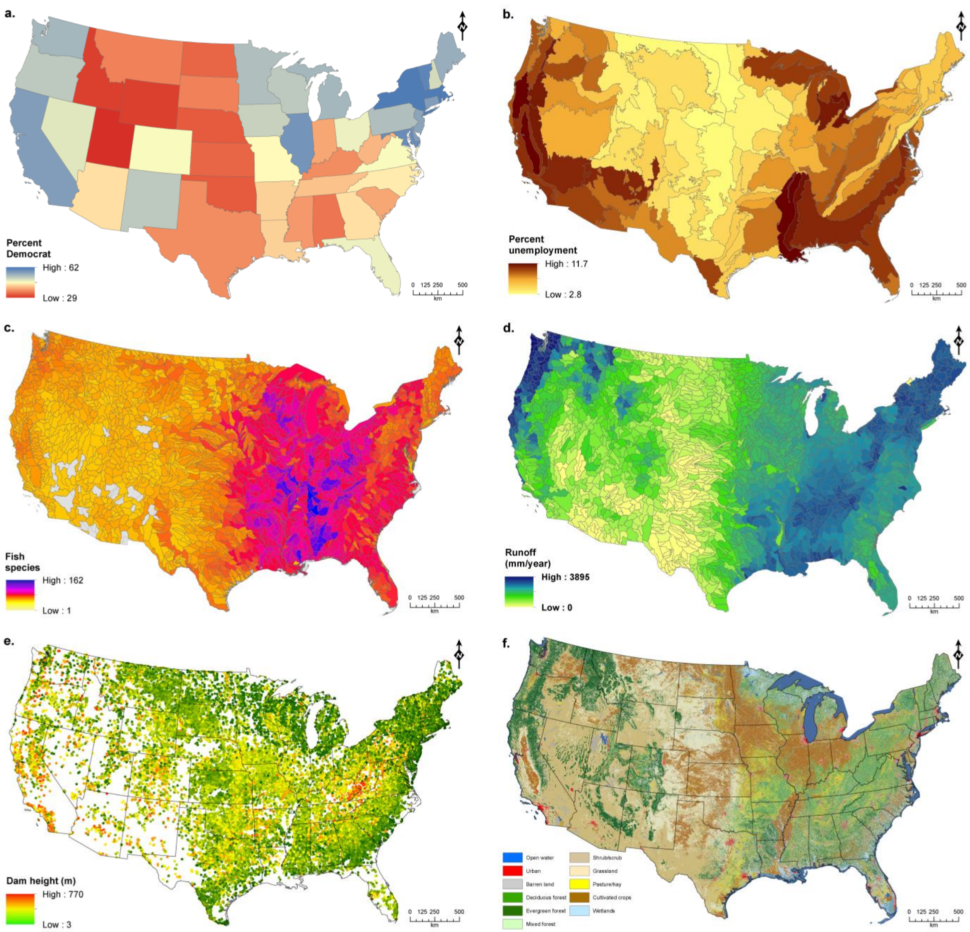


Fig. 2. a. Average percent democratic votes cast in U.S. presidential elections from 1996 to 2012. b. Percent unemployment from 2010 U.S. census, aggregated to physiographic region. c. Number of freshwater fish species per HUC8 watershed. d. Average annual runoff per HUC8 watershed. e. Dam locations symbolized by height. f. 2011 land cover.

Formatted: None, Indent: First line: 0 cm, Space Before: 0 pt, Line spacing: single, Don't keep with next

Formatted: Indent: First line: 0 cm, Line spacing: single

Formatted: Indent: First line: 0 cm, Line spacing: single

Formatted: Indent: First line: 0 cm, Line spacing: single

~~1.1.~~2.1. Mitigation database and response variables

A database of environmental mitigation requirements was compiled for FERC licenses issued from 1998 through September 2015 (Schramm et al., 2016). Since our goal is prediction of future mitigation requirements, the manual review of licenses was limited to those issued from 1998 through 2015 with an assumption that more recently issued licenses would better reflect future mitigation requirements. The database includes Bernoulli distributed presence-absence mitigation data at 463 hydropower plants in the

Comment [DCR1]: This was in revision when I initially submitted this paper, but has now been published.

study area from 316 licenses. Six broad categories (Tier 1) of mitigation (biodiversity, fish passage, habitat, hydrology, recreation, and water quality) and 20 subcategories (Tier 2) were used to classify specific mitigation types in the hierarchical database. A full list of each of the mitigation types catalogued in the database and the percent of times each was required, including each of the 132 Tier 3 categories, is presented in Appendix A. Descriptions of each of the Tier 3 categories is provided in Appendix A of Schramm et al. (2016). Predictive models were built only if a mitigation type was required for at least 5% (Rickbeil et al., 2014) of the plants in the mitigation database. Models were not built for the very broad Tier 1 categories.

Comment [DCR2]: Added reference for 5% threshold.

~~1.1.~~2.2. Explanatory variables

Given that hydropower project licensing is influenced by a suite of biophysical and socio-economic factors, the candidate predictor variables (Table 1) employed here were selected based on expert opinion and on previous research by Kosnik (2010) and Trussart (2002) as broad-scale measures of biological, facility, human, hydrologic, landscape, locational, and stream network characteristics thought to have some bearing on mitigation requirements. The models that each candidate predictor was included in are indicated in Table 1. We used expert opinion to identify candidate predictors for each of the six Tier 1 categories, and these six predictor sets were then used to build models for each Tier 2 and Tier 3 model nested within the Tier I categories. Given that our goal was prediction and not explanation, we did not delve into the exact causal role of each potential predictor. Instead, we selected predictors based on hypothesized quality of association between the predictor and the response, data quality, and data availability (Shmueli, 2010).

~~1.1.1.~~2.2.1. Biological

The presence or absence of important fish species can influence not only fish passage mitigation requirements but also other measures related to biological conservation (Cada, 1998; Fraley et al., 1989; Renofalt et al., 2010). We used conservation status in concert with expert opinion to compile a list of high profile migratory fish species supported by policy protections (McManamay et al., 2015). We then mapped distributions of each of these species using the NatureServe (2010) database of current

Table 1. Summary and description of input variables for the boosted regression tree models. Variables or units in bold and underlined indicate remaining predictor variables after collinearity analysis.

Variable	Description	Units	Spatial scale	Source	Models
Biological					
bigPlyrSum	Major migratory fish species	Count	HUC8 watershed	NatureServe fish distributions, expert opinion	P, H, B, A, R
Facility					
Height	Dam height	Feet	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
HY_MW	Generation capacity	Megawatts	Hydropower dam	ORNL NHAAP database	--
HY_MWh	Generation	Megawatt-hours	Hydropower plant	ORNL NHAAP database	--
Length	Dam length	Feet	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
Mode	Dam mode-of-operation	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
owner	Ownership type	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
PrmyPurps	Dam primary purpose	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
Human					
birdG_xx	National Audubon Society chapters	Count, PA, PC	State	National Audubon Society	--
damR_xx	Dam removals	**Count**, PA, PC	State	American Rivers	P, H, W, B, A, R
education	Education attainment - percent bachelor's degree or higher	Percent	USEPA Level 3 Ecoregion	US Census	P, H, W, B, A, R
FishG_xx	TU and CCA chapters	**Count**, PA, **PC**	State	TU, CCA	P, H, B, A, R
hshldincm	Mean household income	US dollars	USEPA Level 3 Ecoregion	US Census	--
IssueYear	FERC hydropower project license issue year	Year	Hydropower plant	ORNL NHAAP database	P, H, W, B, A, R
LandG_xx	Land trusts	Count, PA, PC	State	Land Trust Alliance	--
politics	see note*	Difference	State	US Federal Election Commission	P, H, W, B, A, R
xx_POPDENS	2000 population density	Individuals/km2 (L and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A, R
q12_avg	Survey response on environmental impact of dams	Rank	Geographic region	MIT Energy Survey, 2008	P, H, W, B, A, R
q16_avg	Survey response on increasing or decreasing hydro power	Rank	Geographic region	MIT Energy Survey, 2008	P, H, W, B, A, R
SierG_xx	Sierra Club chapters	Count, **PA**, **PC**	State	Sierra Club	P, H, W, B, A, R
unemplymnt	Unemployment	Percent	USEPA Level 3 Ecoregion	US Census	P, H, W, B, A, R
usHouse	LCV US House of Rep. mean scorecard for 1998 to 2013	Percent	State	League of Conservation Voters	--
usSenate	LCV US Senate mean scorecard for 1998 to 2013	Percent	State	League of Conservation Voters	--
wshed_xx	Local watershed associations	**Count, PA, PC**	State	USEPA	P, H, W, B, A, R
Hydrology					
ADRAIN	Total artificial drainage area	Square meters	NHD Plus V1 Catchment	USGS	H, W, A
BFI_MEAN	Mean base-flow index for GW discharge into streams	Percent	NHD Plus V1 Catchment	USGS	H, W, A
CNTC_MEAN	Baseflow residence time in the subsurface	Days	NHD Plus V1 Catchment	USGS	H, W, A
DITCHES	Estimated area subject to the practice of ditches	Square meters	NHD Plus V1 Catchment	USGS	--
FlowYr	Average annual flow	Cubic feet per second	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
IRRIG	Estimated area subject to the practice of irrigation	Square meters	NHD Plus V1 Catchment	USGS	--
KFACT	Soil erodibility factor	Dimensionless	NHD Plus V1 Catchment	USGS	H, W, A
MAVELU	Mean Annual Velocity (fps) at bottom of flowline	Cubic feet per second	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
MEAN_IEOF	Mean value for infiltration-excess overland flow	Percent	NHD Plus V1 Catchment	USGS	H, W, A
MEAN_RCHRG	Mean annual natural groundwater recharge	Millimeters	NHD Plus V1 Catchment	USGS	--
nidStorSum	Accumulated upstream storage	Acre-feet	Hydropower dam	National Anthropogenic Barriers Dataset	H, W, A
ResDay	Reservoir residence time	Days	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
SATOF_MEAN	Average value of saturation overland flow	Percent	NHD Plus V1 Catchment	USGS	H, W, A
SfArea	Reservoir surface area	Acres	Hydropower dam	National Inventory of Dams	P, H, W, B, A, R
Stor	Reservoir storage	Acre-feet	Hydropower dam	National Inventory of Dams	--
TILES	Estimated area of tile drains	Square meters	NHD Plus V1 Catchment	USGS	--
Landscape					
xx_CROPS	Land cover classified as cultivated crops	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_NPDES	Number of NPDES sites	Count (**L** and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W
xx_PASTURE	Land cover classified as pasture/hay	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_ROADCRC	Road-stream crossings	Count (L and N)	NHD Plus V1 Catchment	USGS, National Fish Habitat Partnership	--
xx_URBANHC	Land cover classified as high intensity urban	Percent (L and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	--
xx_URBANL	Land cover classified as low intensity urban	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_URBANM	Land cover classified as medium intensity urban	Percent (**L** and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
CNPY_MEAN	Mean canopy cover	Percent	NHD Plus V1 Catchment	USGS	H, W, B, A
CROP_AREA	Total crop area for fertilizer/manure derived from land use	Square meters	NHD Plus V1 Catchment	USGS	H, W, B, A
d303_count	Impaired or threatened waters	Count	NHD Plus V1 Catchment	USEPA 303(d) list	H, W, B, A
IMPV_MEAN	Mean impervious surface	Percent	NHD Plus V1 Catchment	USGS	--
L_MINES	Number of mines or mineral processing plants	Count	NHD Plus V1 Catchment	National Fish Habitat Partnership	W
L_ROADLEN	Length of roads	Meters	NHD Plus V1 Catchment	National Fish Habitat Partnership	P, H, W, B, A
MAXELEVSMO	Maximum elevation	Meters	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
PPT30MEAN	30-year (1971-2000) average annual precipitation	Millimeters	NHD Plus V1 Catchment	USGS	P, H, W, B, A, R
SLOPE	Slope of stream reach	Unitless	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
SLP_PERC	Landscape slope	Percent	NHD Plus V1 Catchment	USGS	P, H, W, B, A, R
TMAX30_MEA	30-year (1971-2000) average annual maximum temperature	Celsius	NHD Plus V1 Catchment	USGS	--
Location					
POINT_X	Longitude	Decimal degrees	Hydropower dam	National Anthropogenic Barriers Dataset	P, H, W, B, A, R
POINT_Y	Latitude	Decimal degrees	Hydropower dam	National Anthropogenic Barriers Dataset	P, H, W, B, A, R
Stream network					
dist2Mouth	Stream network distance to network mouth	Meters	Entire downstream flow path	Calculated from NHD Plus V1 flowlines	P, H, W, B, A, R
DrArea	Drainage area upstream of dam	Square miles	Hydropower dam	National Inventory of Dams	--
dsDams	Downstream dams on flow path to network mouth	Count	Entire downstream flow path	Calculated from NHD Plus V1 and NABD	P, H, W, B, A, R
N_DAMSC	Number of dams within network catchment	Count	NHD Plus V1 Catchment	National Fish Habitat Partnership	P, H, W, B, A, R
SO	Strahler stream order	Strahler number	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R

PA = per area; PC = per capita; L = local catchment; N = entire network catchment; xx indicates variable derived for multiple units; P = fish passage; H = hydrology; W = water quality; B = biodiversity; A = habitat; R = recreation; *politics is the difference between mean percent democrat and republican from 1996 to 2012 presidential elections; LCV = League of Conservation Voters; TU = Trout Unlimited; CCA = Coastal Conservation Association.

Formatted: Centered, Indent: First line: 0 cm

Variable	Description	Units	Spatial unit	Source	Models
Biological					
bigPlyrSum	Major migratory fish species	Count	HUC8 watershed	NatureServe fish distributions, expert opinion	P, H, B, A, R
Facility					
Height	Dam height	Feet	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
HY_MW	Generation capacity	Megawatts	Hydropower dam	ORNL NHAAP database	--
HY_MWh	Generation	Megawatt-hours	Hydropower plant	ORNL NHAAP database	--
Length	Dam length	Feet	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
Mode	Dam mode-of-operation	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
owner	Ownership type	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
PrmyPurps	Dam primary purpose	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
Human					
birdG_xx	National Audubon Society chapters	Count, PA, PC	State	National Audubon Society	--
damR_xx	Dam removals	**Count**, PA, PC	State	American Rivers	P, H, W, B, A, R
education	Education attainment - percent bachelor's degree or higher	Percent	USEPA Level 3 Ecoregion	US Census	P, H, W, B, A, R
FishG_xx	TU and CCA chapters	**Count**, PA, **PC**	State	TU, CCA	P, H, B, A, R
hshldincm	Mean household income	US dollars	USEPA Level 3 Ecoregion	US Census	--
IssueYear	FERC hydropower project license issue year	Year	Hydropower plant	ORNL NHAAP database	P, H, W, B, A, R
LandG_xx	Land trusts	Count, PA, PC	State	Land Trust Alliance	--
politics	see note*	Difference	State	US Federal Election Commission	P, H, W, B, A, R
xx_POPDENS	2000 population density	Individuals/km2 (L and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A, R
q12_avg	Survey response on environmental impact of dams	Rank	Geographic region	MIT Energy Survey, 2008	P, H, W, B, A, R
q16_avg	Survey response on increasing or decreasing hydro power	Rank	Geographic region	MIT Energy Survey, 2008	P, H, W, B, A, R
SierG_xx	Sierra Club chapters	Count, **PA, PC**	State	Sierra Club	P, H, W, B, A, R
unemplymnt	Unemployment	Percent	USEPA Level 3 Ecoregion	US Census	P, H, W, B, A, R
usHouse	LCV US House of Rep. mean scorecard for 1998 to 2013	Percent	State	League of Conservation Voters	--
usSenate	LCV US Senate mean scorecard for 1998 to 2013	Percent	State	League of Conservation Voters	--
wshed_xx	Local watershed associations	**Count, PA, PC**	State	USEPA	P, H, W, B, A, R
Hydrology					
ADRAIN	Total artificial drainage area	Square meters	NHD Plus V1 Catchment	USGS	H, W, A
BFI_MEAN	Mean base-flow index for GW discharge into streams	Percent	NHD Plus V1 Catchment	USGS	H, W, A
CNTC_MEAN	Baseflow residence time in the subsurface	Days	NHD Plus V1 Catchment	USGS	H, W, A
DITCHES	Estimated area subject to the practice of ditches	Square meters	NHD Plus V1 Catchment	USGS	--
FlowYr	Average annual flow	Cubic feet per second	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
IRRIG	Estimated area subject to the practice of irrigation	Square meters	NHD Plus V1 Catchment	USGS	--
KFACT	Soil erodibility factor	Dimensionless	NHD Plus V1 Catchment	USGS	H, W, A
MAVELU	Mean Annual Velocity (fps) at bottom of flowline	Cubic feet per second	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
MEAN_IEOF	Mean value for infiltration-excess overland flow	Percent	NHD Plus V1 Catchment	USGS	H, W, A
MEAN_RCHRG	Mean annual natural groundwater recharge	Millimeters	NHD Plus V1 Catchment	USGS	--
nidStorSum	Accumulated upstream storage	Acre-feet	Hydropower dam	National Anthropogenic Barriers Dataset	H, W, A
ResDay	Reservoir residence time	Days	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
SATOF_MEAN	Average value of saturation overland flow	Percent	NHD Plus V1 Catchment	USGS	H, W, A
SfArea	Reservoir surface area	Acres	Hydropower dam	National Inventory of Dams	P, H, W, B, A, R
Stor	Reservoir storage	Acre-feet	Hydropower dam	National Inventory of Dams	--
TILES	Estimated area of tile drains	Square meters	NHD Plus V1 Catchment	USGS	--
Landscape					
xx_CROPS	Land cover classified as cultivated crops	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_NPDES	Number of NPDES sites	Count (**L** and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W
xx_PASTURE	Land cover classified as pasture/hay	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_ROADCRC	Road-stream crossings	Count (L and N)	NHD Plus V1 Catchment	USGS, National Fish Habitat Partnership	--
xx_URBANHC	Land cover classified as high intensity urban	Percent (L and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	--
xx_URBANL	Land cover classified as low intensity urban	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_URBANM	Land cover classified as medium intensity urban	Percent (**L** and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
CNPY_MEAN	Mean canopy cover	Percent	NHD Plus V1 Catchment	USGS	H, W, B, A
CROP_AREA	Total crop area for fertilizer/manure derived from land use	Square meters	NHD Plus V1 Catchment	USGS	H, W, B, A
d303_count	Impaired or threatened waters	Count	NHD Plus V1 Catchment	USEPA 303(d) list	H, W, B, A
IMPV_MEAN	Mean impervious surface	Percent	NHD Plus V1 Catchment	USGS	--
L_MINES	Number of mines or mineral processing plants	Count	NHD Plus V1 Catchment	National Fish Habitat Partnership	W
L_ROADLEN	Length of roads	Meters	NHD Plus V1 Catchment	National Fish Habitat Partnership	P, H, W, B, A
MAXELEVSMO	Maximum elevation	Meters	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
PPT30MEAN	30-year (1971-2000) average annual precipitation	Millimeters	NHD Plus V1 Catchment	USGS	P, H, W, B, A, R
SLOPE	Slope of stream reach	Unitless	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
SLP_PERC	Landscape slope	Percent	NHD Plus V1 Catchment	USGS	P, H, W, B, A, R
TMAX30_MEA	30-year (1971-2000) average annual maximum temperature	Celsius	NHD Plus V1 Catchment	USGS	--
Location					
POINT_X	Longitude	Decimal degrees	Hydropower dam	National Anthropogenic Barriers Dataset	P, H, W, B, A, R
POINT_Y	Latitude	Decimal degrees	Hydropower dam	National Anthropogenic Barriers Dataset	P, H, W, B, A, R
Stream network					
dist2Mouth	Stream network distance to network mouth	Meters	Entire downstream flow path	Calculated from NHD Plus V1 flowlines	P, H, W, B, A, R
DrArea	Drainage area upstream of dam	Square miles	Hydropower dam	National Inventory of Dams	--
dsDams	Downstream dams on flow path to network mouth	Count	Entire downstream flow path	Calculated from NHD Plus V1 and NABD	P, H, W, B, A, R
N_DAMSC	Number of dams within network catchment	Count	NHD Plus V1 Catchment	National Fish Habitat Partnership	P, H, W, B, A, R
SO	Strahler stream order	Strahler number	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R

PA = per area; PC = per capita; L = local catchment; N = entire network catchment; xx indicates variable derived for multiple units; P = fish passage; H = hydrology; W = water quality; B = biodiversity; A = habitat; R = recreation; *politics is the difference between mean percent democrat and republican from 1996 to 2012 presidential elections; LCV = League of Conservation Voters; TU = Trout Unlimited;

distributions of freshwater fishes of the U.S. at the 8-digit hydrologic unit code (HUC8) scale to allow for analysis of interactions between these high profile species and hydropower project locations. The biological explanatory variable is a count of the number of key fish species per HUC8 (McManamay et al., 2015).

~~1.1.2.~~2.2.2. Facility characteristics

The Oak Ridge National Laboratory (ORNL) National Hydropower Asset Assessment Program (NHAAP) is an integrated energy, water, and ecosystem research effort for sustainable hydroelectricity generation and water management. The ORNL NHAAP database (http://nhaap.ornl.gov/) integrates data from multiple data sources and provides the most current, detailed, and spatially comprehensive information for analyzing and visualizing existing U.S. hydropower assets. We included hydropower facility characteristics from the NHAAP database thought to be important drivers of prescribed mitigation such as dam height, generation capacity, dam mode-of-operation, and geographic location (Kosnik, 2010).

~~1.1.3.~~2.2.3. Human dimensions

Formatted: Normal, No bullets or numbering

The convergence of different anthropogenic characteristics such as presence of environmental interest groups, political climate, population demographics, and regulatory tendencies can be impactful on mitigation requirements (Kosnik, 2010). Consistent with the interest group theory of regulation (Peltzman, 1976), Knittel (2006) concluded that electricity industry regulators respond to lobbying from interest groups. In research focused on explaining drivers of environmental mitigation requirements at hydropower projects, Kosnik (2010) found the largest influences on FERC's regulatory decisions to be congressional politics and regulatory tendencies. In an attempt to capture the socio-political and regulatory landscape, w~~W~~e included numerous anthropogenic predictors that serve as direct measures or proxies for local, state, or regional political tendencies, environmental awareness, regulatory trends, and public attitudes toward dams. Candidate predictors aimed at capturing political tendencies include presidential election voting averaged over time and congressional politics. Different aspects of environmental awareness were estimated at a state-scale using prevalence of non-profit organizations, including protection of birds and their habitats (using National Audubon Society chapters), fish and their habitats (using Trout Unlimited chapters for freshwater and Coastal Conservation Association chapters for marine), land conservation (using land trusts), water quality conservation (using local watershed associations), and general environmental awareness (using Sierra Club chapters). Regulatory trends were estimated using the issue year of the license. Prevalence of dam removals and citizen survey responses on energy and environmental impacts from dams were used as estimates of public attitudes toward dams.

2.2.4. Hydrology

1.1.4. Hydrology

Formatted: Indent: First line: 0 cm

Formatted: Normal, No bullets or numbering

Operation of a hydropower facility typically involves modifications to hydrologic regimes both upstream and downstream of dams, reservoirs, or river diversions (Fraley et al., 1989; Ligon et al., 1995; Poff et al., 1997). The magnitude of these flow disturbances can be minimized by discharge management, and there is increasing pressure from regulatory agencies to incorporate ecological flow requirements in licenses and operational plans for hydropower projects (Bunn and Arthington, 2002; Renofalt et al., 2010; Trussart et al., 2002). We included a suite of explanatory variables derived at the stream reach and watershed scale that describe different aspects of the hydrologic regime of a given area, including surface water, groundwater, and reservoir storage characteristics.

1.1.5.2.2.5. Landscape

Broad-scale landscape descriptors such as land cover, terrain, and climate can influence prescribed mitigation in all six of the Tier 1 mitigation categories, either directly or indirectly. Thus we included numerous land cover metrics derived at multiple scales (Tong and Chen, 2002; Wang et al., 2001), topographic variables such as slope and elevation (Moore et al., 1991), and the core climatic variables of average annual precipitation and air temperature (Grimm et al., 2008).

~~1.1.6.~~2.2.6. Location

In the U.S., there are tangible trends and patterns in environmental, economic, cultural, and social conditions from east to west and north to south. We included latitude and longitude to account for spatial effects and capture spatial patterns across the large study area that may be insufficiently represented in the other predictors (Fink et al., 2010; Oppel et al., 2012).

~~1.1.7.~~2.2.7. Stream network

Stream network position and the prevalence of upstream and downstream dams are important descriptors of network fragmentation/connectivity (Kuby et al., 2005). Where a hydropower project falls on the stream network in relation to other barriers and the network mouth can have a strong influence on the nature and magnitude of ordered mitigation (Fraley et al., 1989; Kosnik, 2010).

~~1.2.~~2.3. Statistical analyses

Model development was carried out in R version 3.2.2 (R-Core-Team). Boosted regression trees (a machine-learning technique) were used to develop the predictive models, as this method has been demonstrated to have high predictive performance with presence-absence response variables, allows for complex regression analyses of complex responses, and can handle continuous and categorical explanatory variables (Abram et al., 2015; Arganaraz et al., 2015; Elith et al., 2006; Elith et al., 2008). Before running the models, all predictor variables were assessed for collinearity using Pearson's correlation coefficients (*r*). When *r* values exceeded 0.7 (Dormann et al., 2013), the variable deemed more functionally applicable to hydropower mitigation (Arganaraz et al., 2015; Rickbeil et al., 2014) or that was derived at a higher spatial resolution was retained (Table 1). The data were split into training (80%) and validation (20%) data using the caret package in R, which creates random splits within each class so that the overall class distribution is preserved as well as possible (Kuhn, 2008).

Given the novelty of the mitigation database, we were unable to obtain an independent validation dataset as recommended by Araujo and Guisan (2006). The optimal number of trees was determined using 10-fold cross validation (CV), with the bag fraction set to 0.5 and the learning rate set to 0.001 to ensure that each model had at least 1,000 trees (Elith et al., 2008). The area under the receiver-operating characteristic curve (ROC) calculated on the validation dataset was used to assess predictive performance. We implemented the ROC interpretation presented by Hosmer et al. (2013) where an ROC value of 0.7-0.8 is considered an acceptable prediction, 0.8-0.9 is excellent, and >0.9 is outstanding. For a model to be deemed acceptable, both the internal CV ROC and the validation ROC had to be ≥0.7. We generated partial dependence plots to examine the nature of the models and to interpret the effect of a variable on the response after accounting for the average effects of all other variables in the model (Elith et al., 2008). Spatial autocorrelation of model residuals was evaluated using Moran's *I* statistics (Dormann et al., 2007) calculated with the Spatial Autocorrelation tool in ArcGIS version 10.2.2 (ESRI).

~~1.3.~~2.4. Example model application at non-powered dams

While approximately 2,500 dams in the U.S. provide 78 gigawatts (GW) of conventional and 22 GW of pumped-storage hydropower, there are hundreds of NPDs originally built for other purposes that may be retrofitted for hydropower to produce an additional 12 GW of estimated renewable energy for the U.S (Hadjerioua et al., 2012). While many of the monetary costs and environmental impacts have already been incurred at these sites, our models can be used as a tool to assess potential environmental mitigation requirements that may arise during the hydropower licensing process. As an example of how the modeling can be applied, we made predictions for each of the acceptable models to 568 NPDs estimated by Hadjerioua et al. (2012) to have >1 megawatt (MW) in potential capacity. We used the optimal threshold function in the R package SDMtools (VanDerWal et al., 2012) to identify the value on the ROC curve that is closest to a perfect model fit, and then we applied that value as the predicted present/absent threshold when making predictions to the NPDs.

~~2.~~3. RESULTS

~~2.1.~~3.1. BRT models

Predictive models were built only if a mitigation type was required for at least 5% (Rickbeil et al., 2014) of the plants in the mitigation database, resulting in 57 Tier 3 mitigation types being modelled and all 20 of the Tier 2 mitigations being modelled (see Table 2 for modeling results). Eight of the 57 Tier 3 models were rejected due to either a CV ROC or validation ROC <0.7, leaving 49 Tier 3 models with at least an acceptable fit. All 20 of the Tier 2 models had an ROC ≥0.7. Significant spatial autocorrelation of model residuals was detected in 4 of 20 Tier 2 models and 11 of 49 Tier 3 models.

Comment [DCR3]: Added reference for 5% threshold.

Table 2. Model results summary.

Tier 1	Tier 2	Tier 3	ModelID	Trees	CV ROC	V ROC	OT	MI	Influential Variable 1	Influential Variable 2	Influential Variable 3
Fish Passage	DS	NA (see Tier 2 category)	F101	5550	0.867	0.916	0.36	-0.165	POINT_Y (13)	dist2Mouth (10.7)	FlowYr (8)
	Fish Passage	DS Passage Plan Study Design	F101010	5300	0.892	0.829	0.30	-0.309	POINT_Y (11.1)	dist2Mouth (9.9)	MAXELEVSMO (9.1)
	US	NA (see Tier 2 category)	F102	5250	0.899	0.896	0.29	-0.436	MAXELEVSMO (14.5)	POINT_Y (13)	bigPlyrSum (8.3)
	Fish Passage	Eelway	F102017	4350	0.956	0.966	0.33	*-1.178*	MAXELEVSMO (36.5)	POINT_X (11.9)	bigPlyrSum (7.7)
		US passage study plan or design	F102023	5000	0.909	0.854	0.29	*-0.703*	MAXELEVSMO (12.3)	POINT_X (10.5)	POINT_Y (9.7)
	Passage	NA (see Tier 2 category)	F103	2850	0.780	0.856	0.27	*-0.467*	POINT_Y (9.3)	dsDams (8.1)	MAXELEVSMO (7.5)
	Planning	DS fish passage mon. sampling	F103029	3200	0.888	0.924	0.22	*-0.908*	MAXELEVSMO (18.2)	Height (10)	POINT_X (8.9)
		Fish passage & operations plan	F103031	1050	0.739	0.749	0.08	-0.246	wshed_PC (11.8)	L_ROADLEN (10.5)	dsDams (9.9)
		Fish stranding plan mon. evaluation	F103033	1100	0.712	0.605	--		--	--	--
		US fish passage mon. sampling	F103036	3050	0.891	0.865	0.18	*-0.504*	MAXELEVSMO (16.2)	Height (11.9)	POINT_Y (8.7)
	Entrainment	NA (see Tier 2 category)	F104	3450	0.849	0.756	0.29	-0.222	ResDay (9.8)	SierG_PC (7.8)	politics (7)
		Trash or bar rack	F104043	3700	0.917	0.833	0.22	0.147	POINT_X (10.4)	SierG_PC (9.6)	fishGroups (8.8)
Hydrology	Flow	NA (see Tier 2 category)	F205	4700	0.785	0.787	0.56	*-1.413*	PrmryPurps (7.7)	N_URBANLC (5.3)	Height (5.1)
	Mitigation	Tailrace flow mon. plan	F205045	5850	0.822	0.920	0.39	-0.076	POINT_X (8.5)	IssueYear (4.8)	politics (4.7)
		Tailrace flow or stage mon. equipment	F205048	2650	0.784	0.867	0.17	-0.191	N_CROPSC (9)	Length (7.3)	dist2Mouth (4.7)
		Tailrace ramping rate restriction	F205050	2400	0.790	0.834	0.19	-0.333	SfArea (8.3)	CNPY_MEAN (5.9)	nidStorSum (5.8)
		Bypass flow mon. plan	F205052	2900	0.802	0.853	0.20	*-0.522*	politics (6.9)	Length (6.7)	SfArea (5.6)
		Bypass flushing or flood flow	F205054	3750	0.890	0.951	0.15	0.187	POINT_X (24.6)	N_PASTUREC (6.8)	PPT30MEAN (6.2)
		Bypass flow or stage mon. equipment	F205055	1500	0.735	0.779	0.13	*-0.679*	SLP_PERC (18)	POINT_X (7)	bigPlyrSum (5.2)
		Bypass ramping rate restriction	F205057	2900	0.878	0.802	0.16	0.170	POINT_X (28.6)	BFI_MEAN (6.7)	CNTC_MEAN (6.5)
	Tailrace	NA (see Tier 2 category)	F206	3800	0.863	0.845	0.63	-0.198	POINT_X (10.4)	Mode (10.3)	SLP_PERC (7.2)
	Minimum	Run-of-river Tailrace	F206058	3700	0.904	0.911	0.37	-0.349	Mode (39.2)	Height (7.8)	POINT_X (4.8)
	Flow	Seasonal Tailrace	F206059	2700	0.850	0.846	0.20	0.087	Mode (24.6)	POINT_Y (9.2)	N_PASTUREC (5.7)
		Year-round Tailrace	F206061	1500	0.787	0.899	0.19	-0.312	Mode (22.7)	owner (9.2)	Length (4.3)
	Bypass	NA (see Tier 2 category)	F207	3250	0.808	0.771	0.46	*-0.676*	SfArea (16.3)	MAXELEVSMO (6.6)	MAVELU (4.7)
	Minimum	Seasonal Bypass	F207063	1450	0.678	0.668	--		--	--	--
	Flow	Year-round Bypass	F207065	1200	0.720	0.805	0.23	-0.339	SfArea (15.6)	Height (10.2)	Length (6.9)
	Sediment	NA (see Tier 2 category)	F208	4850	0.767	0.851	0.49	-0.364	IssueYear (6.1)	CNPY_MEAN (5.3)	unemplymnt (4.6)
		Sediment & erosion control plan or mon.	F208066	4100	0.778	0.838	0.47	-0.257	IssueYear (6.6)	CNPY_MEAN (5)	dist2Mouth (4.9)
	Recreation	NA (see Tier 2 category)	F209	1550	0.733	0.796	0.17	-0.127	POINT_X (12.3)	Height (6.8)	SierG_PA (6.8)
	Flow	Provide recreational flow releases	F209071	700	0.655	0.713	--		--	--	--
	Operations	NA (see Tier 2 category)	F210	3050	0.734	0.819	0.53	0.040	FlowYr (8.6)	q16_avg (4.5)	N_URBANLC (4.5)
		Flow mgmt. plan	F210073	3350	0.893	0.985	0.09	-0.399	Length (13)	wshedG_PA (11.4)	IssueYear (8.6)
		Operations compliance mon. plan	F210074	5150	0.807	0.913	0.41	0.146	politics (8.8)	PrmryPurps (4.8)	FlowYr (4.7)
		Provide flow or lake levels electronically	F210075	1750	0.795	0.917	0.14	-0.282	SierG_PA (18.8)	POINT_X (6.2)	Mode (6.1)
Water Quality	DS	NA (see Tier 2 category)	F311	4300	0.838	0.887	0.55	-0.240	ResDay (12.3)	SfArea (6.8)	wshedG_PA (4.6)
	Water	Benthic macroinvertebrate mon.	F311077	1500	0.724	0.938	0.12	-0.396	BFI_MEAN (16.9)	unemplymnt (8.3)	POINT_X (7.8)
	Quality	DO enhancement or mitigation plan	F311078	2200	0.832	0.676	--		--	--	--
		Water quality mon. plan	F311086	6000	0.852	0.873	0.50	-0.375	ResDay (7.7)	SfArea (6.9)	IssueYear (5.1)
	US	NA (see Tier 2 category)	F312	4000	0.831	0.860	0.23	-0.320	unemplymnt (7.6)	POINT_Y (6.8)	N_PASTUREC (6.8)
	Water	Fish tissue sampling & analysis	F312087	4500	0.965	0.823	0.39	*-1.704*	unemplymnt (13.4)	wshedG_PA (9.7)	dist2Mouth (8.8)
	Quality	Impoundment sediment analysis	F312088	4100	0.993	0.999	0.19	0.334	wshedG_PA (21)	dist2Mouth (13.3)	unemplymnt (11.8)
		Inflow water quality mon. plan	F312090	1650	0.831	0.904	0.11	0.106	wshedG_PA (12.1)	KFACT (9.3)	Length (6)
		Impoundment water quaity mon. plan	F312091	4100	0.828	0.805	0.22	-0.166	N_PASTUREC (10)	unemplymnt (6.5)	CNTC_MEAN (5.6)
Biodiversity	Terrestrial	NA (see Tier 2 category)	F413	4150	0.847	0.832	0.64	*-0.638*	POINT_X (17.8)	SfArea (8.7)	Height (6.4)
		Noxious weed & invasive plant mgmt.	F413094	6650	0.912	0.901	0.39	0.068	POINT_X (13.8)	IssueYear (10.4)	PPT30MEAN (6.2)
		Species conservation mgmt. mon.	F413095	5850	0.832	0.899	0.40	-0.265	damRmvls (8.9)	Length (8.5)	Mode (6.7)
		T&E species protection plan	F413096	3950	0.879	0.905	0.21	*0.965*	L_POPDENS (12.5)	SLP_PERC (6.8)	SfArea (5.8)
		Transmission related avian & bat protection	F413097	6250	0.936	0.941	0.19	-0.109	PPT30MEAN (10.3)	POINT_X (9.7)	dsDams (7.5)
		Wildlife terrestrial habitat mgmt.	F413098	4100	0.844	0.937	0.27	0.146	FlowYr (6.3)	SfArea (5.6)	SierG_PC (5.2)
	Aquatic	NA (see Tier 2 category)	F414	3500	0.791	0.859	0.35	-0.271	FlowYr (11.1)	CNPY_MEAN (6.2)	PPT30MEAN (5.7)
		Aquatic species conservation mgmt. mon.	F414100	3400	0.807	0.869	0.34	-0.336	FlowYr (7.9)	POINT_X (7.1)	dist2Mouth (5.4)
		Diadromous species mgmt. mon.	F414101	3000	0.871	0.901	0.26	0.124	POINT_Y (22.1)	FlowYr (11.2)	PPT30MEAN (8.4)
		Invasive aquatic species mgmt.	F414102	2800	0.800	0.881	0.19	*0.552*	FlowYr (15.3)	L_POPDENS (6.4)	POINT_Y (6.2)
Habitat	Fisheries	NA (see Tier 2 category)	F515	2650	0.776	0.730	0.27	-0.038	POINT_X (8.5)	PPT30MEAN (7.4)	Length (5.6)
		DS habitat enhancement	F515105	1200	0.687	0.680	--		--	--	--
		DS woody debris restoration or passage	F515106	2850	0.863	0.879	0.25	0.071	POINT_Y (7.3)	Length (6.2)	damRmvls (5.5)
	Riparian	NA (see Tier 2 category)	F516	2600	0.771	0.869	0.28	0.084	L_POPDENS (8.6)	BFI_MEAN (7.8)	POINT_X (7.4)
		Establish riparian buffers	F516108	3100	0.866	0.864	0.25	*0.673*	Mode (10)	IssueYear (9)	MAVELU (6.5)
		Riparian habitat mon. or planning	F516110	2300	0.793	0.912	0.22	-0.368	L_POPDENS (8.1)	SierG_PA (7.4)	PPT30MEAN (7)
	Reservoir	NA (see Tier 2 category)	F517	6100	0.858	0.905	0.40	-0.082	ResDay (6.4)	IssueYear (5.8)	SfArea (5.4)
		Noxious invasive aquatic plant mgmt.	F517111	6950	0.928	0.952	0.25	*-0.348*	fishGroups (9.2)	IssueYear (7.6)	Length (4.9)
		Shoreline mgmt. plan or program	F517112	4800	0.856	0.952	0.27	0.192	POINT_Y (10.2)	SfArea (8.4)	Height (7.5)
	Wetlands	NA (see Tier 2 category)	F518	3800	0.828	0.874	0.19	0.207	SierG_PA (10.5)	POINT_Y (10.4)	L_POPDENS (5.4)
		Wetland protection	F518116	3500	0.878	0.875	0.14	0.082	POINT_Y (12.4)	Mode (8.5)	PPT30MEAN (6.6)
Recreation	Resources	NA (see Tier 2 category)	F619	3200	0.741	0.744	0.65	0.089	SfArea (10.6)	FlowYr (7.6)	Length (6.9)
	and	Boating facilities	F619118	1200	0.625	0.660	--		--	--	--
	Mitigation	Canoe portage launch	F619119	5000	0.859	0.773	0.37	-0.161	POINT_X (10.2)	dsDams (8)	N_DAMSC (6.7)
		Fishing pier	F619120	1700	0.797	0.869	0.13	0.055	N_DAMSC (18.3)	SfArea (10)	POINT_Y (8.2)
		Interpretive education sign & displays	F619123	1900	0.720	0.731	0.20	-0.192	MAVELU (7.3)	PPT30MEAN (7.3)	dist2Mouth (6.7)
		Parking	F619125	3550	0.715	0.722	0.32	0.441	MAXELEVSMO (13)	FlowYr (8.2)	ResDay (6.5)
		Shoreline access	F619128	750	0.623	0.759	--		--	--	--
		Stocking recreational fish species	F619129	1250	0.756	0.796	0.09	0.069	FlowYr (14)	PPT30MEAN (9.8)	Height (9.7)
		Trail trailhead or camping areas	F619130	3200	0.781	0.601	--		--	--	--
		Other day use area improvements	F619132	4900	0.750	0.781	0.44	-0.430	ResDay (8.5)	IssueYear (7.3)	PPT30MEAN (6.3)
	Planning	NA (see Tier 2 category)	F620	2450	0.753	0.883	0.74	-0.001	PrmryPurps (15.8)	SfArea (10.6)	Length (9.6)
		Recreational mgmt. plan study or mon.	F620131	2450	0.753	0.883	0.74	-0.001	PrmryPurps (15.8)	SfArea (10.6)	Length (9.6)

See Table 1 for variable descriptions; if no influential variables are shown, model rejected due to poor fit; mgmt. = management; DS = downstream; US = upstream; T&E = threatened and endangered, mon. = monitoring; NA = not applicable; CV ROC = internal cross-validation ROC; V ROC = validation ROC; OT = optimal threshold; MI = Moran's Index; italics indicates spatial autocorrelation detected in training data; color scheme for influential variables corresponds to Table 1 color scheme.

~~2.2.~~3.2. Explanatory variables

The three variables with the highest relative influence in each model are presented in Table 2, and partial dependence plots for these variables are presented in Appendix B. Overall, we considered a variable important if its relative influence was ≥5% (Parisien et al., 2011). A summary of the important variables for the Tier 3 models (Fig. 3) shows that nearly all the categories of variables (i.e. biological, facility, human, hydrologic, landscape, locational, and stream network) were influential within each Tier 1 category.

Fish passage (n=7)

Variable	F	Inf
POINT_X	7	0.63
MAXELEVSMO	6	0.92
POINT_Y	4	0.67
bigPlyrSum	4	0.41
Height	3	0.58
dist2Mouth	3	0.57
dsDams	2	0.74
fishG_PC	2	0.39
FlowYr	2	0.33
wshed_PC	1	1.00
SierG_PC	1	0.92
L_ROADLEN	1	0.89
fishGroups	1	0.84
politics	1	0.74
PrmryPurps	1	0.62
SfArea	1	0.50
SLOPE	1	0.43
MAVELU	1	0.43

Hydrology (n=15)

Variable	F	Inf
Mode	5	0.79
POINT_X	5	0.74
Length	4	0.81
SfArea	3	0.94
nidStorSum	3	0.64
Height	3	0.45
politics	2	1.00
IssueYear	2	0.83
CNPY_MEAN	2	0.73
N_PASTUREC	2	0.25
unemplymnt	2	0.22
SierG_PA	1	1.00
N_CROPSC	1	1.00
SLP_PERC	1	1.00
wshedG_PA	1	0.88
fishG_PC	1	0.63
owner	1	0.40
POINT_Y	1	0.37
FlowYr	1	0.33
bigPlyrSum	1	0.29
dist2Mouth	1	0.29
PPT30MEAN	1	0.25
BFI_MEAN	1	0.24
CNTC_MEAN	1	0.23
MAVELU	1	0.21
SierG_PC	1	0.19

Water quality (n=7)

Variable	F	Inf
unemplymnt	4	0.68
wshedG_PA	3	0.91
BFI_MEAN	3	0.61
dist2Mouth	3	0.57
SierG_PC	3	0.41
SfArea	2	0.87
IssueYear	2	0.56
POINT_X	2	0.44
N_PASTUREC	1	1.00
N_DAMSC	1	1.00
ResDay	1	1.00
KFACT	1	0.76
POINT_Y	1	0.67
CNTC_MEAN	1	0.55
Length	1	0.50
damRmvls	1	0.44
Height	1	0.35

Biodiversity (n=8)

Variable	F	Inf
FlowYr	4	0.88
PPT30MEAN	4	0.66
POINT_X	3	0.95
SfArea	3	0.57
L_POPDENS	2	0.71
POINT_Y	2	0.70
SLP_PERC	2	0.55
damRmvls	1	1.00
Length	1	0.95
SierG_PC	1	0.82
Mode	1	0.75
IssueYear	1	0.75
dsDams	1	0.73
PrmryPurps	1	0.70
dist2Mouth	1	0.69
Height	1	0.38

Habitat (n=6)

Variable	F	Inf
POINT_Y	3	1.00
Mode	3	0.74
PPT30MEAN	3	0.64
IssueYear	2	0.86
fishG_PC	2	0.59
L_POPDENS	1	1.00
fishGroups	1	1.00
SierG_PA	1	0.91
Length	1	0.85
SfArea	1	0.82
damRmvls	1	0.75
Height	1	0.73
MAVELU	1	0.65
education	1	0.59
wshedG_PA	1	0.42
PrmryPurps	1	0.41

Recreation (n=10)

Variable	F	Inf
FlowYr	7	0.72
PPT30MEAN	6	0.72
ResDay	5	0.59
MAXELEVSMO	4	0.80
Height	4	0.66
Length	4	0.66
dist2Mouth	4	0.65
SfArea	4	0.56
L_POPDENS	3	0.73
dsDams	3	0.54
MAVELU	2	0.84
N_DAMSC	2	0.83
IssueYear	2	0.77
SLOPE	2	0.72
POINT_X	2	0.70
POINT_Y	2	0.54
PrmryPurps	1	1.00
unemplymnt	1	0.48
Mode	1	0.47

Fig. 3. Explanatory variables with relative influence ≥5 for Tier 3 models, broken down by Tier 1 category. Relative influence normalized to 0 to 1 scale for each model; Inf = mean relative influence for variable across all models in which relative influence ≥5; F = frequency, or number of times variable had Inf ≥5; color scheme corresponds to Table 1.

Across all Tier 3 models (Fig. 4), the most important variables were longitude (location), reservoir surface area (hydrology), average annual flow (hydrology), precipitation (landscape), and latitude (location). Stream network, facility, human, and biological variables were also important but exceeded the ≥5% relative influence threshold less frequently.

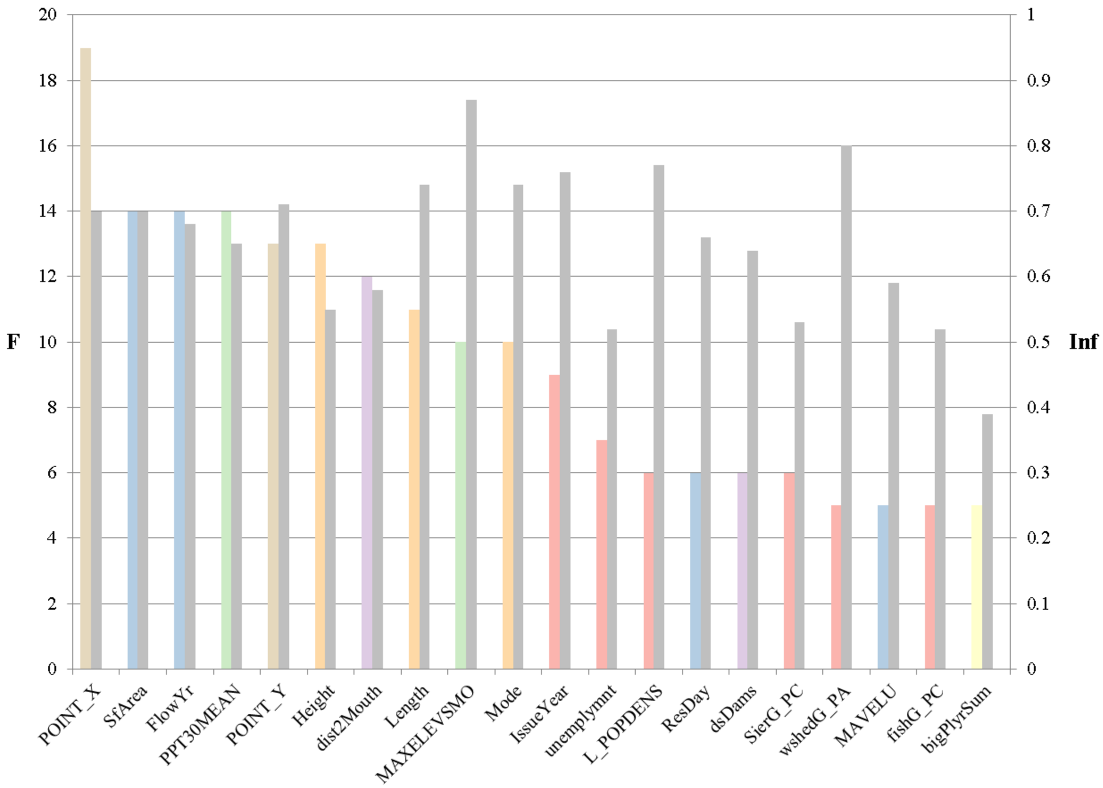


Fig. 4. The 20 most frequently occurring important variables across all Tier 3 models, sorted in descending order from left to right by frequency of occurrence. Colored bars present frequency, while grey bars present the normalized average relative influence for the variable across all of the models in which it was important.

To identify potential key environmental and social drivers of mitigation that may warrant additional future research, we examined important variables across all of our Tier 3 models based on frequency of importance and average relative influence. We grouped important variables into the potential future research areas of socio-political conditions, regional trends, network/landscape position, hydrology/site design, regulatory tendencies, and fisheries (Table 3).

Table 3. The 20 most frequently occurring important variables across all Tier 3 models, with potential future research areas that correspond to each variable. F= frequency; Inf = normalized average relative influence.

Formatted: Font: 9 pt, Bold, Font color: Accent 1

Formatted: Caption

Variable	Category	F	Inf	Future research area
POINT_X	Location	19	0.70	Regional trends
SfArea	Hydrology	14	0.70	Hydrology/site design
FlowYr	Hydrology	14	0.68	Hydrology/site design
PPT30MEAN	Landscape	14	0.65	Hydrology/site design
POINT_Y	Location	13	0.71	Regional trends
Height	Facility	13	0.55	Hydrology/site design
dist2Mouth	Stream network	12	0.58	Network/landscape position
Length	Facility	11	0.74	Hydrology/site design
MAXELEVSMO	Landscape	10	0.87	Network/landscape position
Mode	Facility	10	0.74	Hydrology/site design
IssueYear	Human	9	0.76	Regulatory tendencies
unemplymnt	Human	7	0.52	Socio-political conditions
L_POPDENS	Human	6	0.77	Socio-political conditions
ResDay	Hydrology	6	0.66	Hydrology/site design
dsDams	Stream network	6	0.64	Network/landscape position
SierG_PC	Human	6	0.53	Socio-political conditions
wshedG_PA	Human	5	0.80	Socio-political conditions
MAVELU	Hydrology	5	0.59	Hydrology/site design
fishG_PC	Human	5	0.52	Socio-political conditions
bigPlyrSum	Biological	5	0.39	Fisheries

~~**Table 3.** The 20 most frequently occurring important variables across all Tier 3 models, with potential future research areas that correspond to each variable. F= frequency; Inf = normalized average relative influence.~~

Formatted: Indent: Left: 0 cm

3.3. Predictions to NPDs

We made predictions to 568 NPDs with >1MW potential capacity for each of the 49 acceptable Tier 3 models (Fig. 5). The optimal present/absent threshold for each model is presented in Table 2. The number of predicted mitigation requirements ranged from 9 to 34.

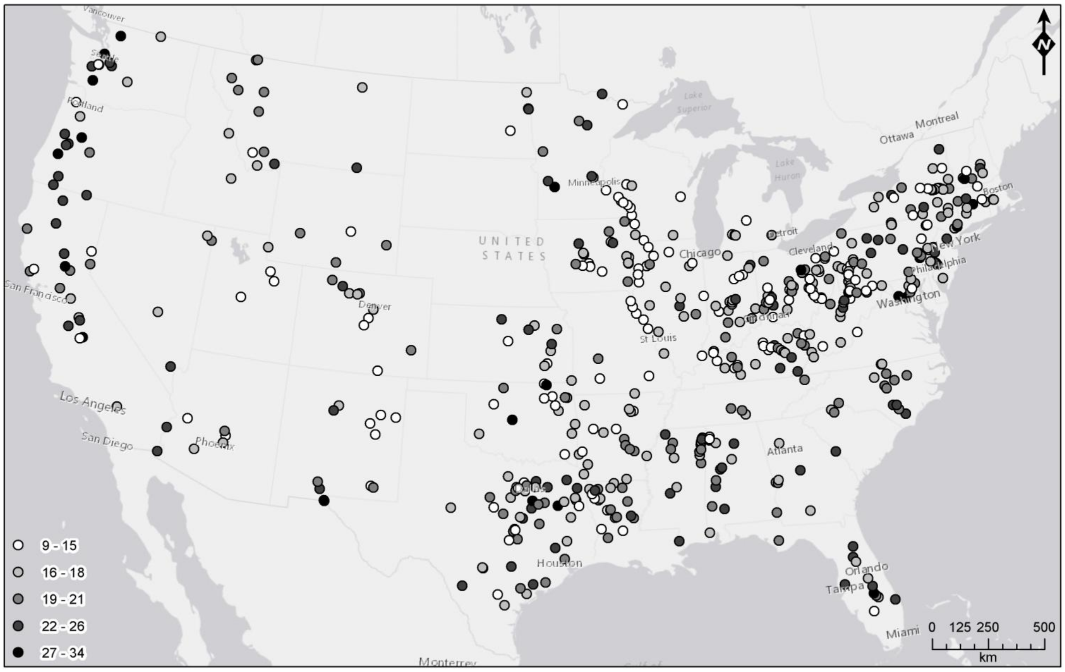


Fig. 5. Number of predicted Tier 3 mitigation requirements at NPDs with >1MW energy potential.

Formatted: Indent: First line: 0 cm

Formatted: Indent: Left: 0 cm

4. DISCUSSION

The spatial modeling approach developed here integrates GIS techniques, novel data, machine-learning algorithms, and niche modeling concepts common in landscape ecology (see Guisan and Thuiller, 2005) to predict environmental mitigation requirements at hydropower project sites. Given the multifaceted, complex nature of demonstrated (Kosnik, 2010) and hypothesized (FERC, 2001) drivers of environmental mitigation requirements, we were uncertain of their predictability. However, we have demonstrated that a broad-scale, multidisciplinary geographical predictor dataset can effectively predict many environmental mitigation requirements across an environmentally and culturally heterogeneous study area.

We summarized and evaluated the influence of the important (relative influence ≥5%) explanatory variables at several different levels of aggregation (Table 2, Fig. 3, and Fig. 4). Since nearly all the categories (e.g. biological, facility, etc.) of variables were influential within each Tier 1 category (Fig. 3) and every Tier 3 model had at least two variable categories represented in the top 3 influential variables (Table 2), it appears that the multi-faceted nature of the predictor dataset we compiled was a key to our modeling success.

Based on our analysis of the top predictor variables across all of the Tier 3 models (Fig. 4), the most common important predictors include metrics of project location (latitude and longitude), project size (annual flow, reservoir size, dam height, and dam length), stream network position (distance along the stream network to network mouth), and climate (precipitation). Elevation above mean sea level, statewide prevalence of local watershed associations, local population density, license issue year, dam length, and dam mode-of-operation were the predictors with the highest average relative influence (Fig. 4) among the most important predictors. Given that the study area is large, environmentally and culturally heterogeneous, and comprised of many diverse physiographic regions, we anticipated latitude and longitude would be valuable predictors that capture regional trends across the U.S. For example, the inclusion of latitude and longitude in models predicting fish passage are related to the fact that most mitigation for passage occurs in the US northeast and northwest. We expected variables related to project size, facility characteristics, and hydrology to be important, given that larger projects are likely to have a higher impact to the environmental and social landscape than smaller projects. Elevation is a proxy for head and a measure of landscape position, and is a very powerful descriptor of landscape context. Stream network position, such as distance to river mouth, can explain presence of diadromous fish species, network connectivity, and existing hydrologic alteration, all of which can heavily influence decisions on mitigation requirements. It is well known in the U.S. that environmental stakeholder groups can be influential in ordered mitigation, so it was not surprising that anthropogenic variables, such as the prevalence of environmental groups, were important. Previous research (Kosnik, 2010) has shown that regulatory trends can influence hydropower mitigation requirements, and the license issue year proved to be an important variable in several models.

Examination of partial dependence plots to assess the direction of variable influence (Appendix B) seems to show that, while there appears to be some consistent direction of influence for important predictors, particularly in the fish passage and water quality models, there are as many examples of contrasting direction of influence within the six broad mitigation categories. This underscores the complexity of the interplay of the nature and magnitude of a given mitigation requirement with the environmental, economic, political, cultural, and social conditions that coalesce at a project and also underscores the need for further investigation into the causality of different drivers of mitigation.

While it is impractical to research causality for all specific mitigation requirements given the sheer number of different types, we identified several potential future research areas (Table 3) that warrant further investigation. One approach to prioritizing future research into mitigation requirement causality would be to delve further into the socio-political and environmental concerns of non-governmental organizations and environmental resource agencies regarding hydropower development, and how those concerns are manifest in prescribed mitigation. These stakeholder groups have a powerful voice and are important to engage early and throughout the project development process if hydropower's contribution to the U.S. renewable energy portfolio is going to be optimized (Fu et al., 2014). A high-level review of The Nature Conservancy's Hydropower by Design strategy (The Nature Conservancy, 2015) and American Rivers Hydropower Reform Coalition platform (Hydropower Reform Coalition, 2016) reveals a common theme of maximizing hydropower sustainability through 1) careful selection of dam location within river networks to optimize both hydropower and conservation objectives, 2) implementing cumulative watershed-scale mitigation strategies, 3) reducing uncertainty and risk associated with project development by directing dam development away from environmentally and socially sensitive areas, and 4) improved outcomes for ecosystem services. Future research into the interplay between socio-political demographics, stream connectivity, ecosystem services, and watershed-scale mitigation approaches and their influence on project siting and ultimate success or failure could serve to catalyze future sustainable hydropower development in the 21st century (Crook et al., 2015; Fu et al., 2014; Karjalainen and Jarvikoski, 2010; Yu et al., 2016).

Another future direction of this research space is the inclusion of cost estimates for different mitigations, which could inform a cost-based approach for identifying priority mitigation types for future investigation of causality. Cost data would also provide a useful constraint for model predictions. Hydropower projects included in the mitigation database (Fig. 1) have a maximum number of 25 mitigation requirements (of the 49 that we modeled), while the model predictions to NPDs included as many as 34 mitigation requirements. Incorporating cost data would allow for additional realism to be integrated into the predictions by sequentially predicting mitigation types from most to least costly with a control on cost.

Our results should be interpreted with caution given that several models showed significant spatial autocorrelation. Since one of our goals was spatial prediction beyond the spatial extent of our dataset, we did not implement methods for accounting for spatial autocorrelation because previously developed methods do not allow for prediction beyond the dataset (Dormann et al., 2007; Rickbeil et al., 2014). We recognize that our models did not use an independent validation dataset, but rather a split of our original dataset. Since there is no comparable dataset available, we argue that our data split combined with tenfold internal cross-validation allowed for reliable evaluation of model performance to be made (Rickbeil et al., 2014).

The BRT models could potentially be improved by improving some of the more coarse resolution predictors – such as those derived at the state-scale – to represent a more refined local scale. A disconnect may exist between the spatial scale at which mitigation requirements are ordered and the scale at which some of our explanatory variables are derived. This disparity of scales and varying resolution of predictors can affect the apparent importance of a predictor variable (Brewer et al., 2007). Schramm et al. (2016) described several possible limitations to the development of the mitigation database, which was limited to a review of mitigation prescribed explicitly in FERC licenses issued from 1998 to 2015. More specifically, some of the reviewed licenses were for relicensing of existing projects and thus may not include previously required mitigation under the original license. Also, FERC encourages the use of settlement agreements (legal agreements developed between hydropower developers, agencies, and other stakeholders on project operations and environmental conditions) that may include mitigation not included in the final license.

5. CONCLUSION

We demonstrated in this study an approach including specific statistical models that can be used by developers and regulators alike to identify and anticipate likely environmental mitigation at existing and proposed hydropower projects in the U.S. The results demonstrate that mitigation requirements in existing licenses have been a result of a range of factors from biological and hydrological to political and cultural. That such a range of variable types is needed to predict mitigation requirements explains much of the difficulty and uncertainty that surrounds the development of effective environmental mitigation during the licensing process in the U.S. Further research is needed to establish robust links between specific explanatory variables, mitigation requirements, and mitigation strategies. However, ~~U~~use of these models by developers can reduce uncertainty with regards to cost projections and inform decisions about project design. Regulators will be able to use the models to more quickly identify likely environmental issues and potential solutions, hopefully resulting in more timely and more effective decisions on environmental mitigation.

6. ACKNOWLEDGEMENTS

We would like to thank R. McManamay for statistical advice, suggestions for predictor variables, and valuable comments on this manuscript. We also thank S.C. Kao for assistance with the NHAAP database and FERC licenses. This study was funded by the U.S. Department of Energy (DOE) Energy Efficiency and Renewable Energy Office, Wind and Water Power Technologies Program through Oak Ridge

National Laboratory, which is managed by UT-Battelle, LLC, for the DOE under contract DE-AC05-00OR22725.

7. REFERENCES

Abram NK, Meijaard E, Wells JA, Ancrenaz M, Pellier A-S, Runting RK, et al. Mapping perceptions of species' threats and population trends to inform conservation efforts: the Bornean orangutan case study. Diversity and Distributions 2015; 21: 487-499.

Araujo M, Guisan A. Five (or so) challenges for species distribution modelling. Journal of Biogeography 2006; 33: 1677-1688.

Arganaraz J, Pizarro G, Zak M, Landi M, Bellis L. Human and biophysical drivers of fires in Semiarid Chaco mountains of Central Argentina. Science of the Total Environment 2015; 520: 1-12.

Blumm MC, Nadol VA. The decline of the hydropower czar and the rise of agency pluralism in hydroelectric licensing. Columbia Journal of Environmental Law 2001; 26.

Brewer SK, Rabeni CF, Sowa SP, Annis G. Natural landscape and stream segment attributes influencing the distribution and relative abundance of riverine smallmouth bass in Missouri. North American Journal of Fisheries Management 2007; 27: 326-341.

Bunn S, Arthington A. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environmental Management 2002; 30: 492-507.

Cada GF. Fish passage mitigation at hydroelectric power projects in the United States. Fish Migration and Fish Bypasses 1998: 208-219.

Crook D, Lowe W, Allendorf F, Eros T, Finn D, Gillanders B, et al. Human effects on ecological connectivity in aquatic ecosystems: Integrating scientific approaches to support management and mitigation. Science of the Total Environment 2015; 534: 52-64.

Deshazo JR, Freeman J. Public Agencies as Lobbyists. Columbia Law Review 2005; 105.

Dormann C, Elith J, Bacher S, Buchmann C, Carl G, Carre G, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 2013; 36: 27-46.

Dormann C, McPherson J, Araujo M, Bivand R, Bolliger J, Carl G, et al. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 2007; 30: 609-628.

Elith J, Graham C, Anderson R, Dudik M, Ferrier S, Guisan A, et al. Novel methods improve prediction of species' distributions from occurrence data. Ecography 2006; 29: 129-151.

Elith J, Leathwick J, Hastie T. A working guide to boosted regression trees. Journal of Animal Ecology 2008; 77: 802-813.

FERC. Report on hydroelectric licensing policies, procedures, and regulations, http://www.ferc.gov/legal/maj-ord-reg/land-docs/ortc_final.pdf (accessed 1.15.15), 2001.

FERC. workload, http://www.ferc.gov/industries/hydropower/gen-info/licensing/workload.xls (accessed 1.15.15), 2015.

Fink D, Hochachka W, Zuckerberg B, Winkler D, Shaby B, Munson M, et al. Spatiotemporal exploratory models for broad-scale survey data. Ecological Applications 2010; 20: 2131-2147.

Fraley J, Marotz B, Decker-Hess J, Beattie W, Zubik R. Mitigation, compensation, and future protection for fish populations affected by hydropower development in the upper Columbia system, Montana, U.S.A. River Research and Applications 1989; 3: 3-18.

Fu B, Wang Y, Xu P, Yan K, Li M. Value of ecosystem hydropower service and its impact on the payment for ecosystem services. Science of the Total Environment 2014; 472: 338-346.

Grimm N, Foster D, Groffman P, Grove J, Hopkinson C, Nadelhoffer K, et al. The changing landscape: ecosystem responses to urbanization and pollution across climatic and societal gradients. Frontiers in Ecology and the Environment 2008; 6: 264-272.

Guisan A, Thuiller W. Predicting species distribution: offering more than simple habitat models. Ecology Letters 2005; 8: 993-1009.

Hadjerioua B, Wei Y, Kao S-C. An Assessment of Energy Potential at Non-Powered Dams in the United States. ORNL/TM-2011/341. Oak Ridge National Laboratory, 2012.

Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression. Hoboken, NJ, U.S.A.: John Wiley & Sons, 2013.

Hydropower Reform Coalition. Platform, http://hydroreform.org/sites/default/files/HRC%20Platform_2.pdf (accessed 2.15.16), 2016.

Karjalainen TP, Jarvikoski TB. Negotiating river ecosystems: Impact assessment and conflict mediation in the cases of hydro-power construction. Environmental Impact Assessment Review 2010; 30: 319-327.

Knittel CR. The Adoption of State Electricity Regulation: The Role of Interest Groups. Journal of Industrial Economics 2006; 54: 201-222.

Kosnik L. Balancing Environmental Protection and Energy Production in the Federal Hydropower Licensing Process Land Economics 2010; 86: 444-466.

Kuby MJ, Fagan WF, ReVelle CS, Graf WL. A multiobjective optimization model for dam removal: an example trading off salmon passage with hydropower and water storage in the Willamette basin. Advances in Water Resources 2005; 28: 845-855.

Kuhn M. Building Predictive Models in R Using the caret Package. Journal of Statistical Software 2008; 28: 1-26.

Liermann C, Nilsson C, Robertson J, Ng R. Implications of Dam Obstruction for Global Freshwater Fish Diversity. Bioscience 2012; 62: 539-548.

Ligon FK, Dietrich WE, Trush WJ. Downstream ecological effects of dams. Bioscience 1995; 45: 183-192.

Manzano-Agugliaro F, Alcayde A, Montoya F, Zapata-Sierra A, Gil C. Scientific production of renewable energies worldwide: An overview. Renewable & Sustainable Energy Reviews 2013; 18: 134-143.

McManamay R, Samu N, Kao S, Bevelhimer M, Hetrick S. A Multi-scale Spatial Approach to Address Environmental Effects of Small Hydropower Development. Environmental Management 2015; 55: 217-243.

Moore ID, Grayson RB, Ladson AR. Digital Terrain Modeling - A review of hydrological, geomorphological, and biological applications. Hydrological Processes 1991; 5: 3-30.

NatureServe. Digital Distribution Maps of the Freshwater Fishes in the Conterminous United States, Arlington, VA. U.S.A., 2010.

Nilsson C, Reidy C, Dynesius M, Revenga C. Fragmentation and flow regulation of the world's large river systems. Science 2005; 308: 405-408.

Oppel S, Meirinho A, Ramirez I, Gardner B, O'Connell A, Miller P, et al. Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. Biological Conservation 2012; 156: 94-104.

Parisien M, Parks S, Krawchuk M, Flannigan M, Bowman L, Moritz M. Scale-dependent controls on the area burned in the boreal forest of Canada, 1980-2005. Ecological Applications 2011; 21: 789-805.

Peltzman S. Toward a more general theory of regulation. Journal of Law and Economics 1976; 19: 211-240.

Poff N, Allan J, Bain M, Karr J, Prestegaard K, Richter B, et al. The natural flow regime. Bioscience 1997; 47: 769-784.

Poff N, Olden J, Merritt D, Pepin D. Homogenization of regional river dynamics by dams and global biodiversity implications. Proceedings of the National Academy of Sciences of the United States of America 2007; 104: 5732-5737.

R-Core-Team. R: a language and environment for statistical computing, Vienna, Austria. Version 3.2.2. Available at https://www.r-project.org/. R Foundation for Statistical Computing, 2015.

Renofalt B, Jansson R, Nilsson C. Effects of hydropower generation and opportunities for environmental flow management in Swedish riverine ecosystems. Freshwater Biology 2010; 55: 49-67.

Rickbeil G, Coops N, Drever M, Nelson T. Assessing coastal species distribution models through the integration of terrestrial, oceanic and atmospheric data. Journal of Biogeography 2014; 41: 1614-1625.

Schramm MP, Bevelhimer MS, DeRolph CR. A synthesis of environmental and recreational mitigation requirements at hydropower projects in the United States. Environmental Science & Policy 2016; 61: 87-96.

Shmueli G. To Explain or to Predict? Statistical Science 2010; 25: 289-310.

Tarlock D. Hydro law and the future of hydroelectric power generation in the United States. Vanderbilt Law Review 2012; 65: 1723-1767.

The Nature Conservancy. The Power of Rivers, http://www.nature.org/media/freshwater/power-of-rivers-report.pdf (accessed 2.15.16), 2015.

Tong S, Chen W. Modeling the relationship between land use and surface water quality. Journal of Environmental Management 2002; 66: 377-393.

Trussart S, Messier D, Roquet V, Aki S. Hydropower projects: a review of most effective mitigation measures. Energy Policy 2002; 30: 1251-1259.

U.S.-Congress. House Commitee on Natural Resources. Mandatory conditioning requirements on hydropower: How federal resources agencies are driving up electricity costs and decreasing the original green energy, 2012.

VanDerWal J, Falconi L, Januchowski S, Shoo L, Storlie C. Package 'SDMTools', http://cran.r-project.org/web/packages/SDMTools/index.html (accessed August 2015), 2012.

Wang L, Lyons J, Kanehl P. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 2001; 28: 255-266.

Yu B, Xu L. Review of ecological compensation in hydropower development. Renewable & Sustainable Energy Reviews 2016; 55: 729-738.

Yu B, Xu L, Yang Z. Ecological compensation for inundated habitats in hydropower developments based on carbon stock balance Journal of Cleaner Production 2016; 114: 334-342.

Appendix A

Mitigation categories in the hierarchical database and the percent of times each was required. Bold indicates model was fit for mitigation category.

Tier 1	Tier 2	Tier 3	ModelID	Percent required
Fish Passage	--	--	F1	48.8%
	Downstream Fish Passage	--	**F101**	**24.9%**
		Surface Collector	F101001	1.9%
		Trap and Transport	F101002	1.4%
		Modification of Spill or Gate Operation	F101003	3.3%
		Sluiceway	F101004	0.4%
		Bypass Facility	F101005	2.1%
		Conduit	F101006	1.2%
		Fish Friendly Turbine	F101007	0.2%
		Generation Shut Down	F101008	2.1%
		Flashboard Removal or Modification	F101009	0.2%
		Downstream Passage Plan Study Design	**F101010**	**15.5%**
		Modify spill or gate design	F101011	2.7%
		Modify sluiceway	F101012	1.6%
		Modify bypass facility	F101013	0.2%
		Modify intake	F101014	0.2%
	Upstream Fish Passage	--	**F102**	**19.4%**
		Adult fishway	F102015	0.4%
		Conduit	F102016	0.2%
		Eelway	**F102017**	**7.4%**
		Fish Ladder	F102018	3.3%
		Lock or elevator	F102019	2.3%
		Modify spill or gate operation	F102020	0.2%
		Tailrace exclusion device	F102021	2.5%
		Trap and transport	F102022	4.3%
		Upstream passage study plan or design	**F102023**	**12.2%**
		Modify adult fishway	F102024	0.2%
		Modify fish ladder	F102025	0.8%
		Modify lock or lift	F102026	0.2%
		Modify trap and transport	F102027	0.6%
	Passage Planning	--	**F103**	**26.2%**
		Design plan entrainment avoidance system	F103028	1.6%
		Downstream fish passage monitoring sampling	**F103029**	**13.6%**
		Entrainment or turbine mortality monitoring	F103030	3.7%
		Fish passage and operations plan	**F103031**	**7.4%**
		Fish passage feasibility assessment	F103032	3.1%
		Fish stranding plan monitoring evaluation	**F103033**	**7.2%**
		Fisheries disease management	F103034	0.6%
		Hatchery operations and management	F103035	1.9%
		Upstream fish passage monitoring sampling	**F103036**	**10.7%**
	Entrainment	--	**F104**	**23.9%**
		Barrier or guidance net	F104037	1.6%
		Fish screen	F104038	4.3%
		Gatewell exclusion screen	F104039	0.4%
		Perforated plate	F104040	0.2%
		Solid panel and bar rack	F104041	0.4%
		Strobe light	F104042	0.2%
		Trash or bar rack	**F104043**	**17.1%**
Hydrology	--	--	F2	95.1%
	Flow Mitigation	--	**F205**	**61.9%**
		Tailrace adaptive flow management	F205044	1.6%
		Tailrace flow monitoring plan	**F205045**	**34.0%**
		Tailrace flow studies	F205046	3.5%
		Tailrace flushing or flood flows	F205047	1.9%
		Tailrace flow or stage monitoring equipment	**F205048**	**14.2%**
		Tailrace flow control device	F205049	2.7%
		Tailrace ramping rate restriction	**F205050**	**11.1%**
		Bypass adaptive flow management	F205051	1.9%
		Bypass flow monitoring plan	**F205052**	**12.8%**
		Bypass flow study	F205053	2.5%

Appendix A

Mitigation categories in the hierarchical database and the percent of times each was required. Bold indicates model was fit for mitigation category.

Tier 1	Tier 2	Tier 3	ModelID	Percent required
		Bypass flushing or flood flow	**F205054**	**5.6%**
		Bypass flow or stage monitoring equipment	**F205055**	**7.2%**
		Bypass flow control device	F205056	0.8%
		Bypass ramping rate restriction	**F205057**	**6.2%**
	Tailrace Minimum Flow	--	**F206**	**64.5%**
		Run-of-river Tailrace	**F206058**	**39.0%**
		Seasonal Tailrace	**F206059**	**13.6%**
		Seasonal and type of year Tailrace	F206060	1.6%
		Year-round Tailrace	**F206061**	**10.3%**
	Bypass Minimum Flow	--	**F207**	**41.9%**
		Seasonal Bypass	**F207063**	**17.1%**
		Seasonal and type of year Bypass	F207064	4.5%
		Year-round Bypass	**F207065**	**20.2%**
	Sediment	--	**F208**	**42.9%**
		Sediment and erosion control plan or monitoring	**F208066**	**41.6%**
		Dredging	F208067	0.2%
		Install or operate gate to flush sediment	F208068	0.8%
		Sediment flushing flows	F208069	0.8%
	Recreation Flow	--	**F209**	**13.2%**
		Maintain recreational lake levels	F209070	3.3%
		Provide recreational flow releases or structures	**F209071**	**9.7%**
		Recreational flow studies	F209072	4.1%
	Operations	--	**F210**	**54.8%**
		Flow management plan	**F210073**	**6.6%**
		Operations compliance monitoring plan	**F210074**	**40.6%**
		Provide flow or lake levels electronically	**F210075**	**10.7%**
Water Quality	--	--	F3	53.7%
	Downstream Water Quality	--	**F311**	**54.0%**
		Adaptive water quality management	F311076	3.7%
		Benthic macroinvertebrate monitoring	**F311077**	**5.4%**
		DO enhancement or mitigation plan	**F311078**	**5.4%**
		Establish or fund water quality stations and stream gages	F311079	3.3%
		Forebay aeration	F311080	0.2%
		Operational changes	F311081	2.7%
		Powerhouse aeration	F311082	2.1%
		Tailrace structures for aeration	F311083	0.2%
		Temperature regulating device or structure	F311084	0.6%
		Temperature regulation or mitigation plan	F311085	0.4%
		Water quality monitoring plan	**F311086**	**50.3%**
	Upstream Water Quality	--	**F312**	**24.5%**
		Fish tissue sampling and analysis	**F312087**	**8.2%**
		Impoundment sediment analysis	**F312088**	**6.4%**
		Macroinvertebrate monitoring	F312089	0.6%
		Inflow water quality monitoring plan	**F312090**	**8.9%**
		Impoundment water quaity monitoring plan	**F312091**	**17.3%**
Biodiversity	--	--	F4	71.4%
	Terrestrial	--	**F413**	**66.6%**
		Acquistion easements conservation or important habitat	F413092	4.1%
		Install upgrade monitor wildlife crossings	F413093	4.1%
		Noxious terrestrial weed and invasive plant management	**F413094**	**25.6%**
		Species conservation management monitoring	**F413095**	**42.9%**
		Threatened and endangered species protection plan	**F413096**	**10.9%**
		Transmission related avaian and bat protection	**F413097**	**15.5%**
		Wildlife terrestrial habitat management	**F413098**	**27.0%**
	Aquatic	--	**F414**	**35.5%**
		Adaptive fishery management	F414099	3.9%
		Aquatic species conservation management monitoring	**F414100**	**25.6%**
		Diadromous species management monitoring	**F414101**	**7.4%**
		Invasive aquatic species management (fish and molluscs)	**F414102**	**9.3%**
		Stocking fish species of concern	F414103	4.5%

Appendix A

Mitigation categories in the hierarchical database and the percent of times each was required. Bold indicates model was fit for mitigation category.

Tier 1	Tier 2	Tier 3	ModelID	Percent required
Habitat	--	--	F5	57.1%
	Fisheries	--	**F515**	**26.8%**
		Downstream gravel and sediment restoration	F515104	4.5%
		Downstream habitat enhancement	**F515105**	**8.7%**
		Downstream woody debris restoration or passage	**F515106**	**15.9%**
		Reservoir fishery habitat enhancement	F515133	3.3%
	Riparian	--	**F516**	**20.4%**
		Dust control and abatement	F516107	0.6%
		Establish riparian buffers	**F516108**	**7.2%**
		Riparian habitat enhancement	F516109	4.7%
		Riparian habitat monitoring or planning	**F516110**	**12.0%**
	Reservoir	--	**F517**	**31.8%**
		Noxious invasive aquatic plant management	**F517111**	**21.0%**
		Shoreline management plan or program	**F517112**	**16.1%**
	Wetlands	--	**F518**	**11.5%**
		Wetland enhancement	F518113	0.8%
		Wetland mitigation	F518114	4.9%
		Wetland monitoring	F518115	2.9%
		Wetland protection	**F518116**	**6.8%**
Recreation	--	--	F6	82.3%
	Resources and Mitigation	--	**F619**	**66.2%**
		Appoint historic cultural resource coordinator	F619117	0.4%
		Boating facilities	**F619118**	**23.1%**
		Canoe portage launch	**F619119**	**24.1%**
		Fishing pier	**F619120**	**8.9%**
		Floating debris removal	F619121	1.0%
		Install fish attracting structure for recreational fishing	F619122	2.7%
		Interpretive education sign and displays	**F619123**	**15.5%**
		Navigational aids and improvements	F619124	1.0%
		Parking	**F619125**	**26.4%**
		Protection of specific historic cultural resource sites	F619126	3.3%
		Public outreach education programs	F619127	0.6%
		Shoreline access	**F619128**	**17.9%**
		Stocking recreational fish species	**F619129**	**6.6%**
		Trail trailhead or camping areas	**F619130**	**14.2%**
		Other day use area improvements	**F619132**	**35.3%**
	Planning	--	**F620**	**72.8%**
		Recreational management plan study or monitoring	**F620131**	**72.8%**

Appendix B

Partial dependence plots for the three variables with the highest relative influence for each Tier 2 and Tier 3 model with an internal CV ROC and independent ROC ≥0.7

NOTES: The ModelID for each model is shown in the upper-right hand corner of each set of three partial dependence plots; see Table 1 for variable descriptions; see Table 2 for details on mitigation types; ticks across the top of each plot show the distribution of deciles for each predictor variable.

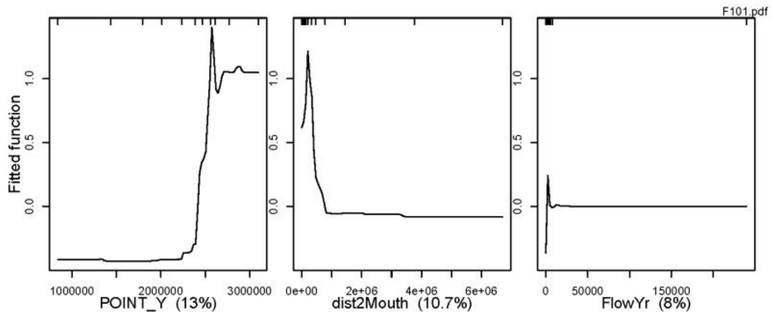
Definitions for categorical variables

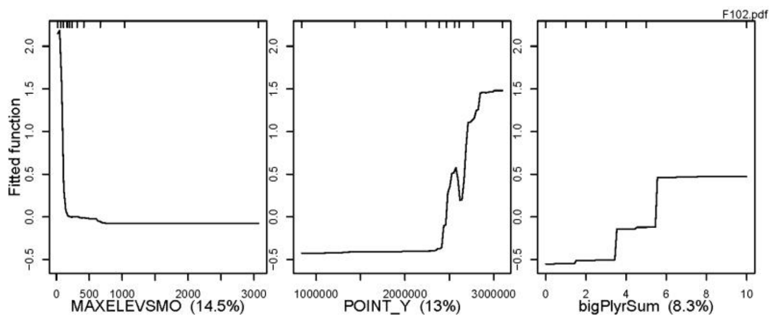
Mode	Value
Canal/Conduit	C
Intermediate Peaking	I
Peaking	P
Pumped Storage	S
Reregulating	R
Run-of-river	O
Run-of-river/Peaking	A
Run-of-river/Upstream Peaking	B

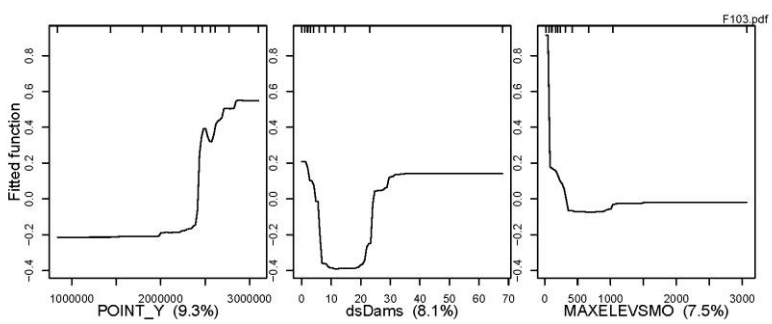
owner	Value
Cooperative	C
Private	P
Public	U
Wholesale Power Marketer	W

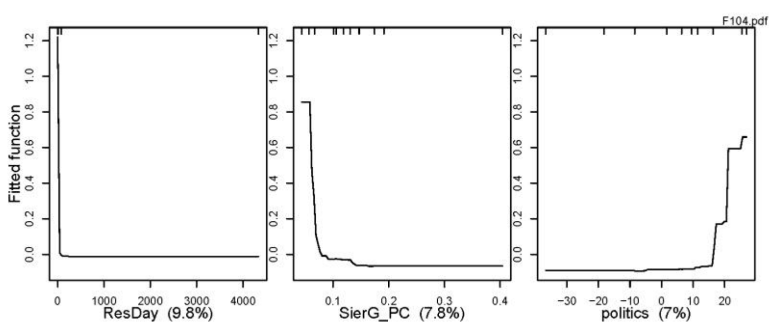
PrmryPurps	Value
Flood control and stormwater management	C
Fish and wildlife pond	F
Hydroelectric	H
Irrigation	I
Navigation	N
Other	O
Recreation	R
Water supply	S

Tier 2 Fish Passage

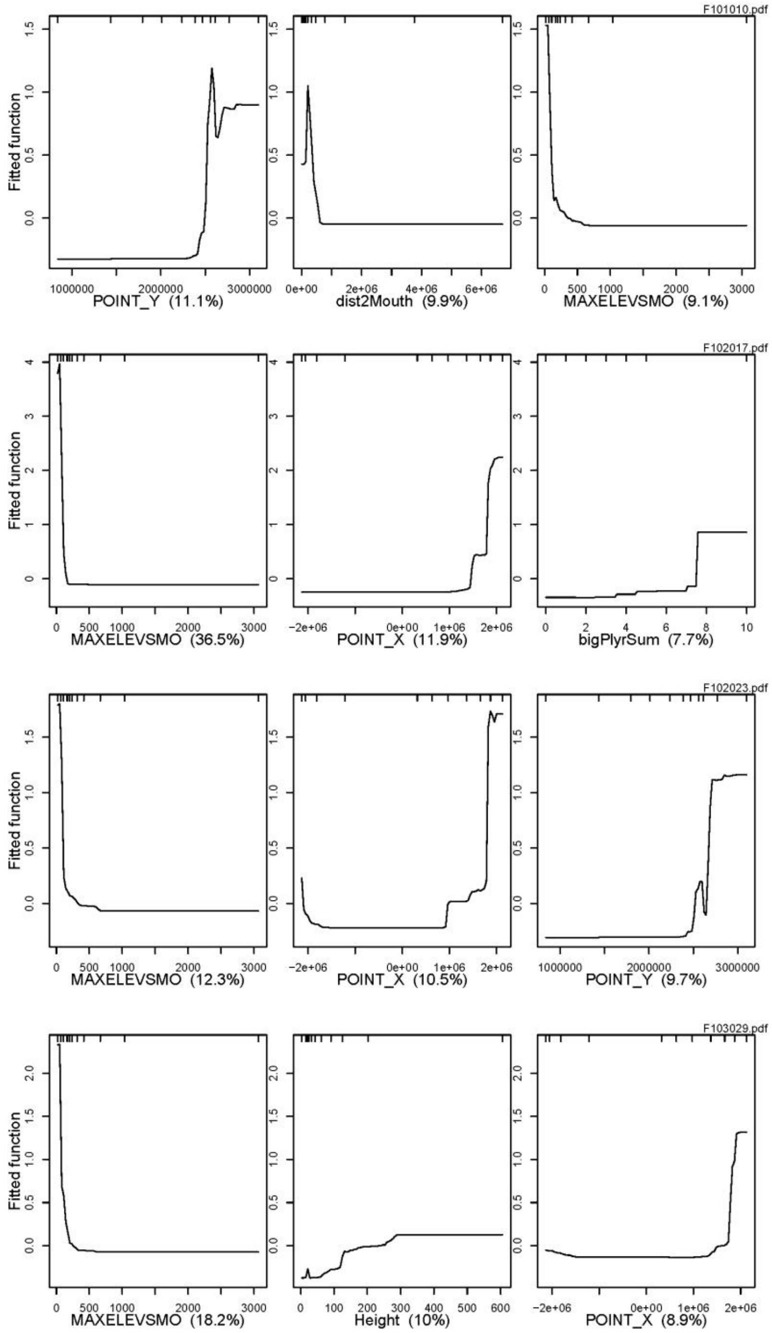


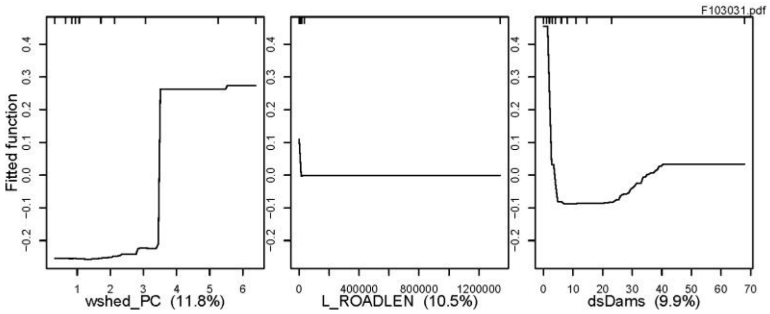


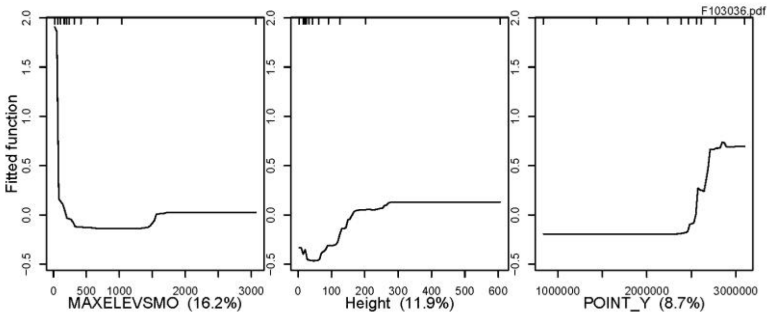


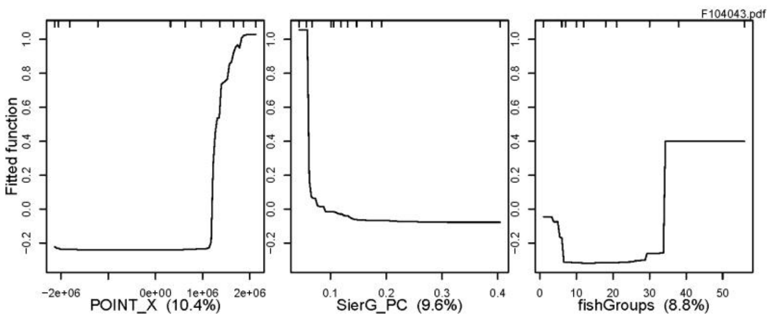


Tier 3 Fish Passage

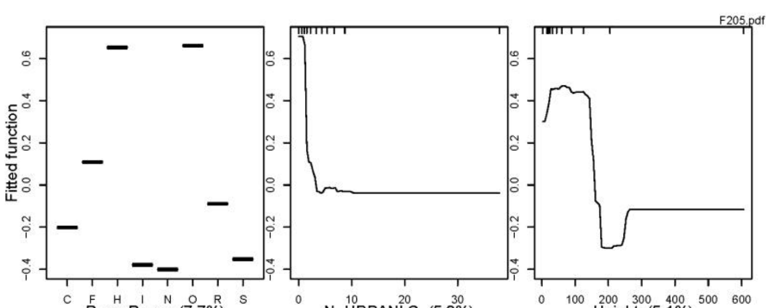


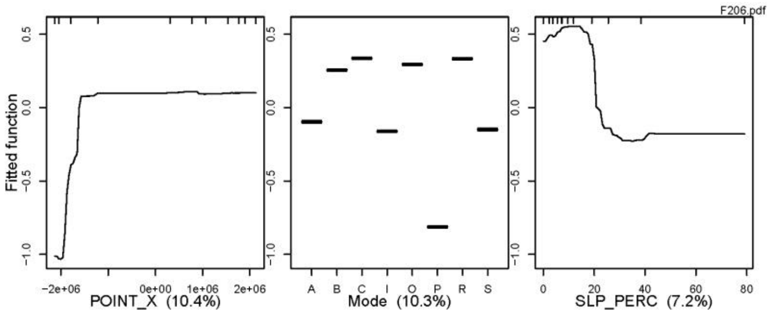


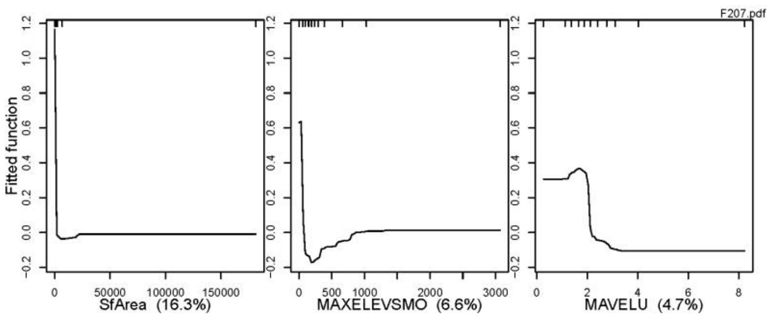


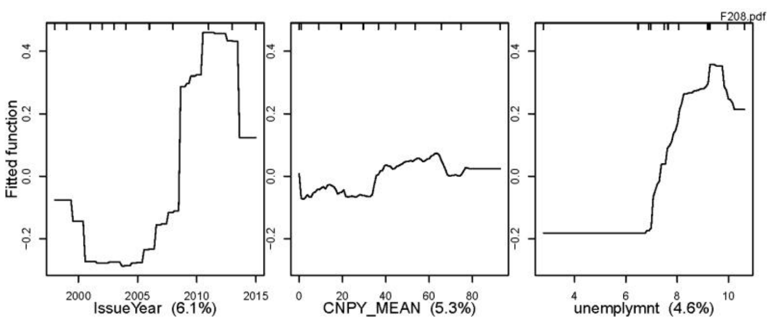


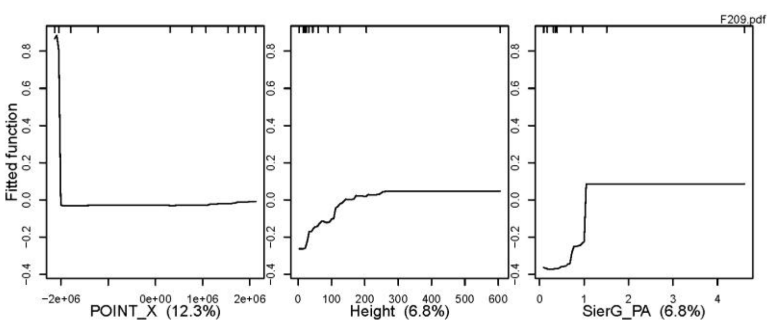
Tier 2 Hydrology

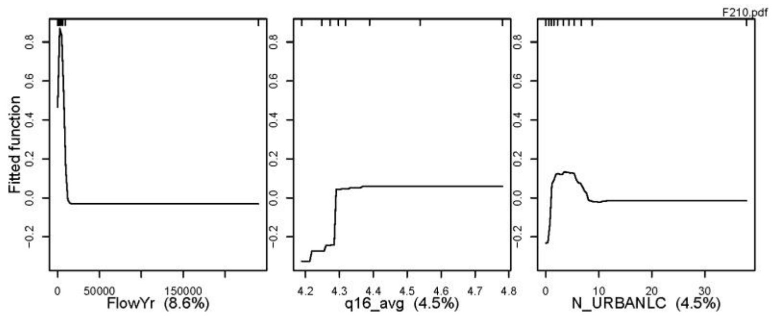



F206.pdf
Fitted function
POINT_X (10.4%)
Mode (10.3%)
SLP_PERC (7.2%)

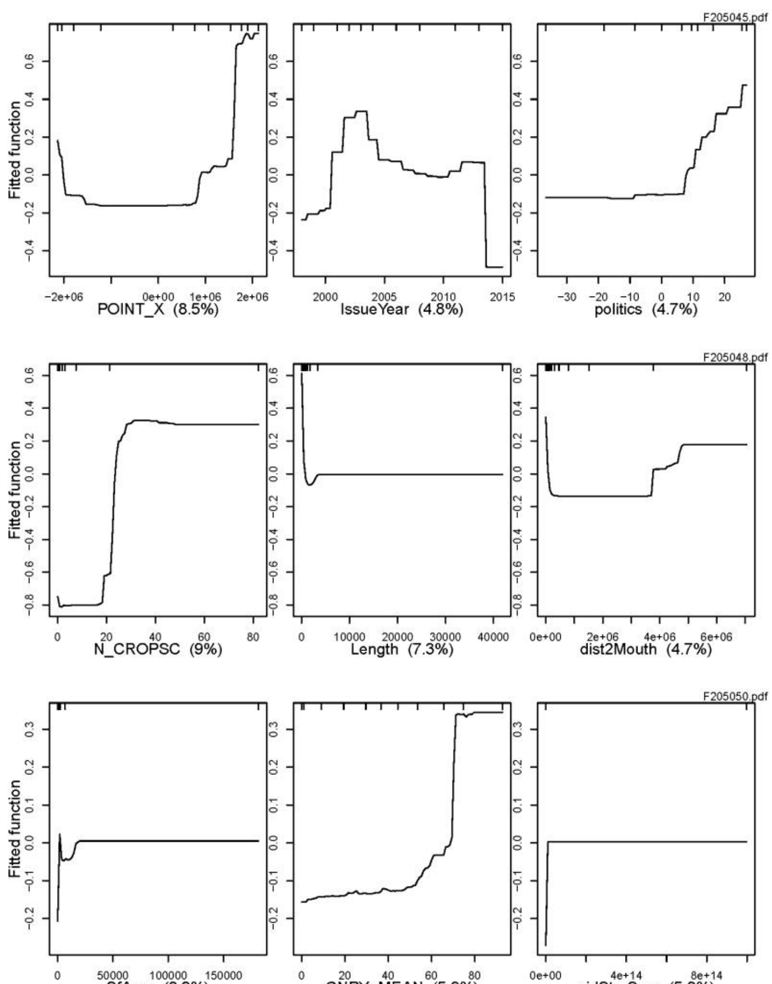

F207.pdf
Fitted function
SfArea (16.3%)
MAXELEVSMO (6.6%)
MAVELU (4.7%)

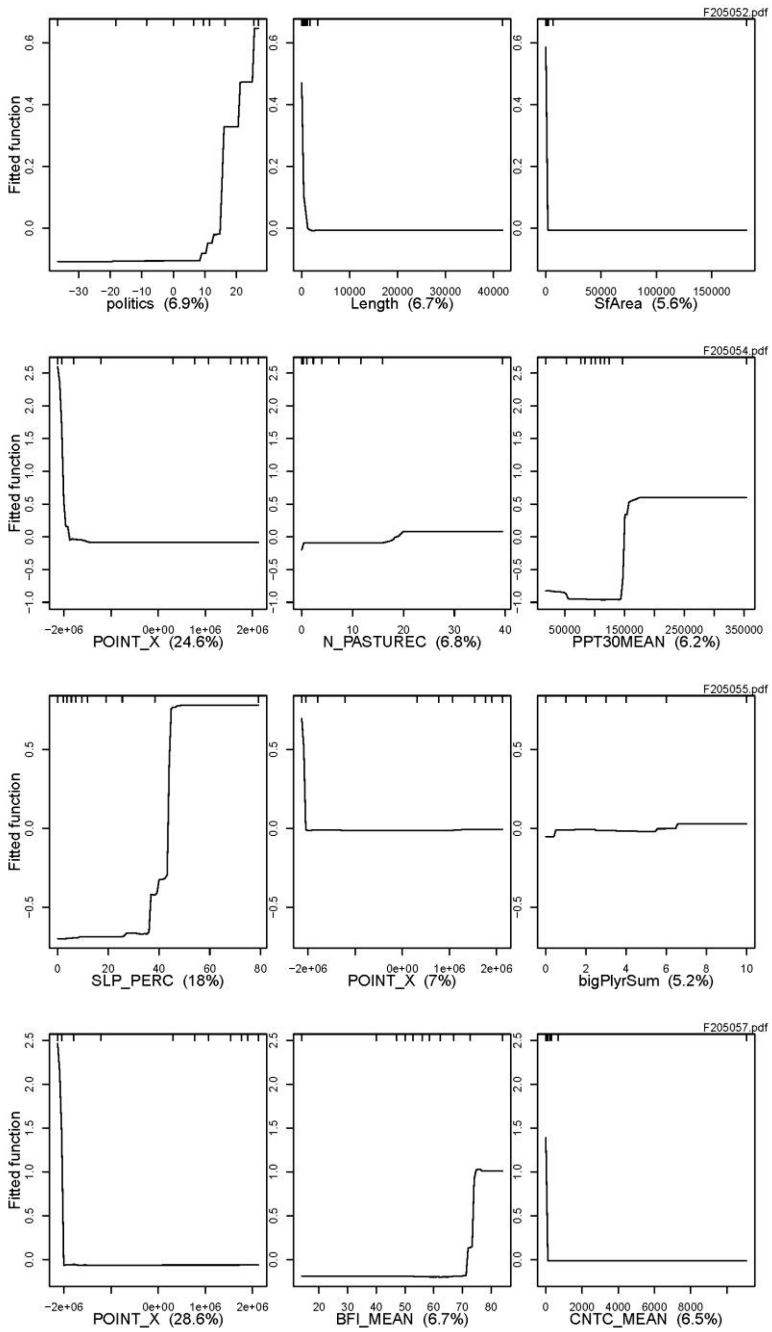

F208.pdf
Fitted function
IssueYear (6.1%)
CNPY_MEAN (5.3%)
unemplymnt (4.6%)

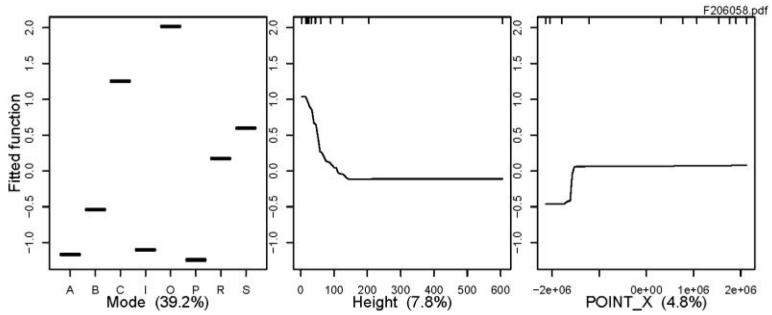

F209.pdf
Fitted function
POINT_X (12.3%)
Height (6.8%)
SierG_PA (6.8%)

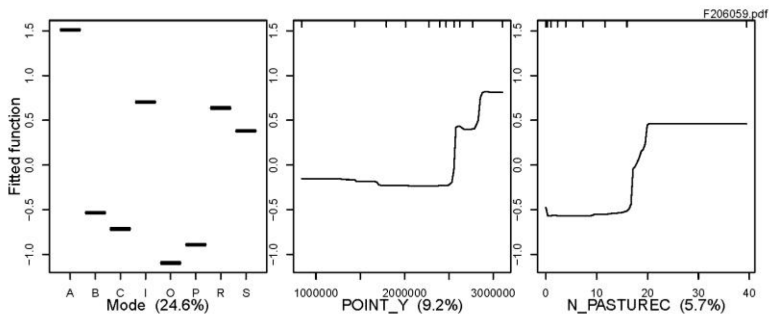


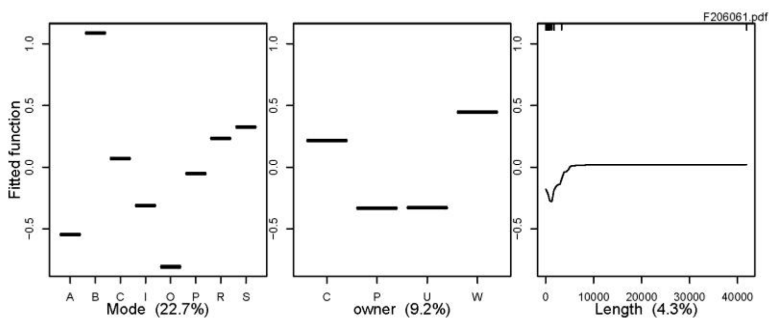
Tier 3 Hydrology

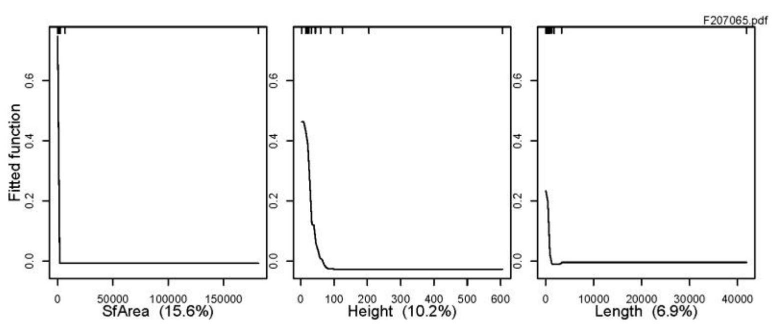


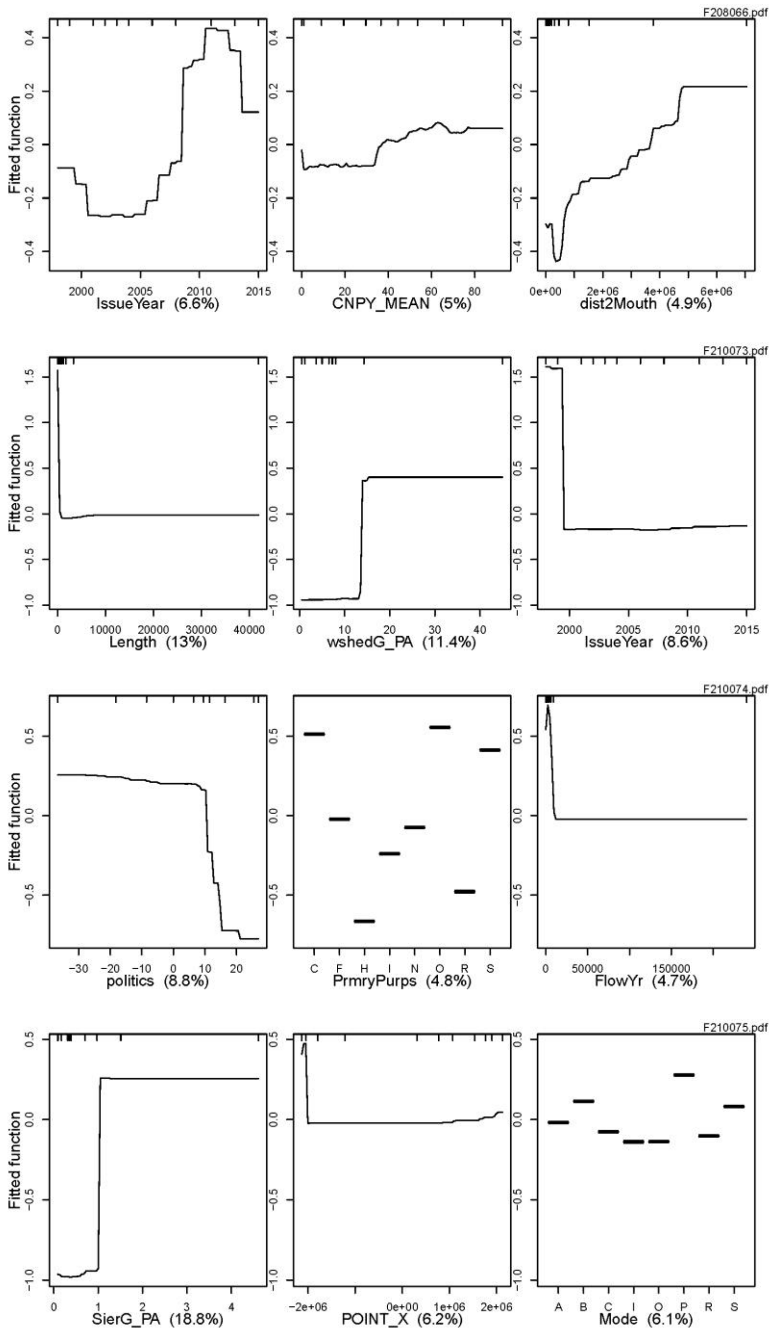

F205052.pdf
Fitted function
politics (6.9%)
Length (6.7%)
SfArea (5.6%)
F205054.pdf
Fitted function
POINT_X (24.6%)
N_PASTUREC (6.8%)
PPT30MEAN (6.2%)
F205055.pdf
Fitted function
SLP_PERC (18%)
POINT_X (7%)
bigPlyrSum (5.2%)
F205057.pdf
Fitted function
POINT_X (28.6%)
BFI_MEAN (6.7%)
CNTC_MEAN (6.5%)


F206058.pdf
Fitted function
Mode (39.2%)
Height (7.8%)
POINT_X (4.8%)

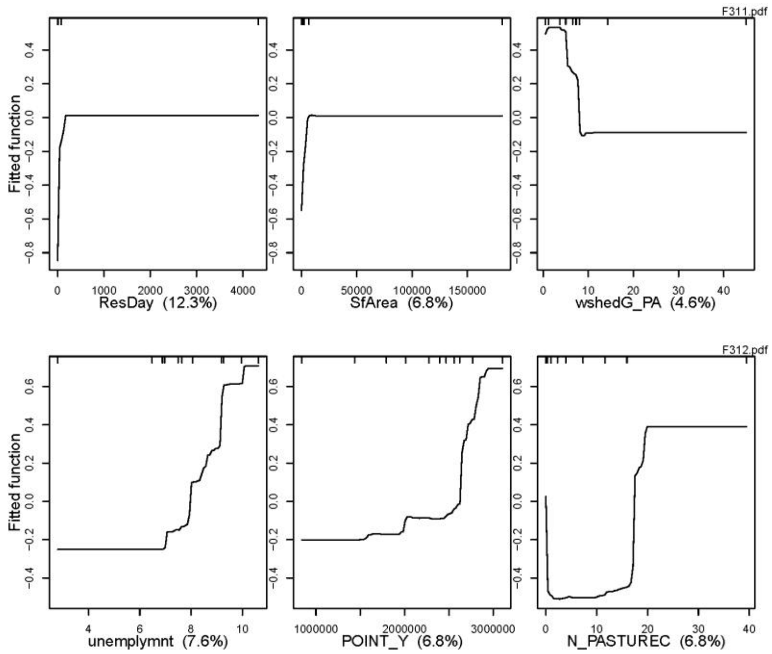

F206059.pdf
Fitted function
Mode (24.6%)
POINT_Y (9.2%)
N_PASTUREC (5.7%)


F206061.pdf
Fitted function
Mode (22.7%)
owner (9.2%)
Length (4.3%)

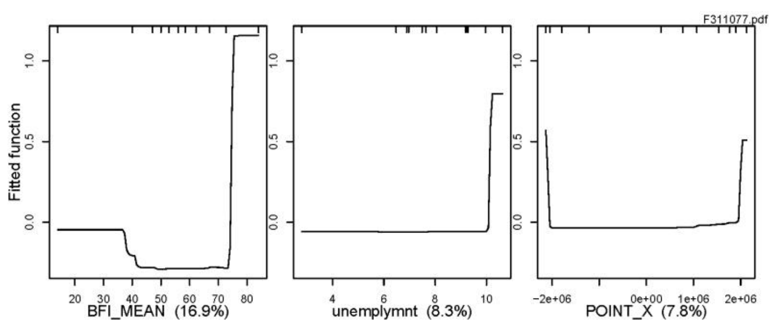

F207065.pdf
Fitted function
SfArea (15.6%)
Height (10.2%)
Length (6.9%)

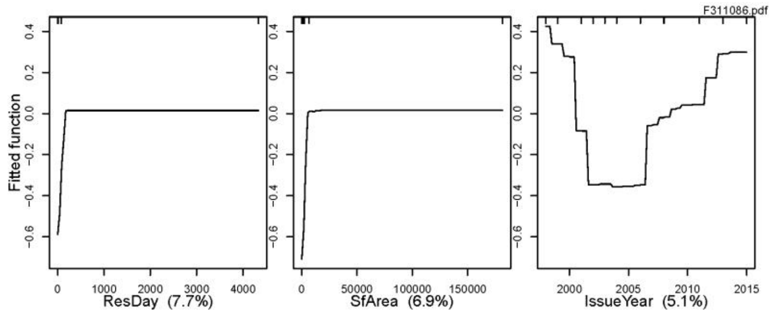

F208066.pdf
Fitted function
IssueYear (6.6%)
CNPY_MEAN (5%)
dist2Mouth (4.9%)
F210073.pdf
Fitted function
Length (13%)
wshedG_PA (11.4%)
IssueYear (8.6%)
F210074.pdf
Fitted function
politics (8.8%)
PrmryPurps (4.8%)
FlowYr (4.7%)
F210075.pdf
Fitted function
SierG_PA (18.8%)
POINT_X (6.2%)
Mode (6.1%)

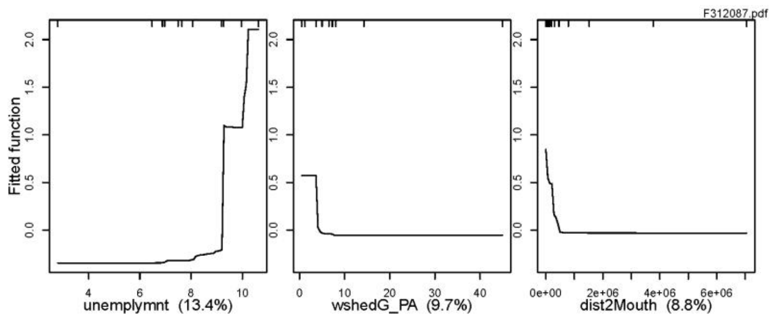
Tier 2 Water Quality

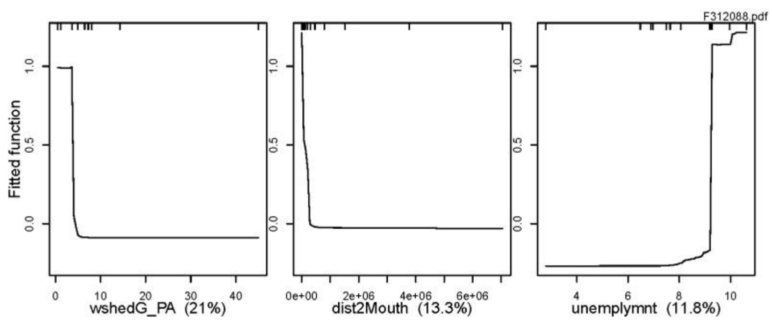


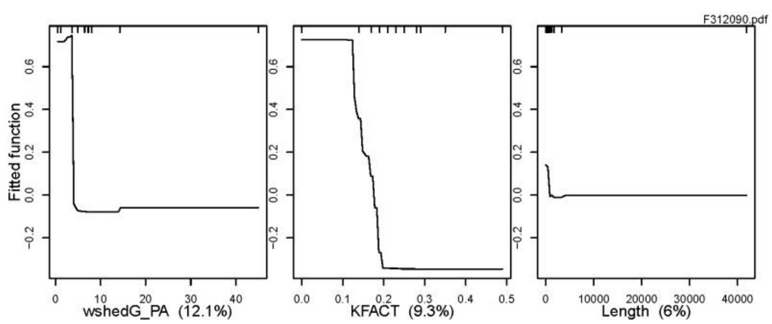
Tier 3 Water Quality

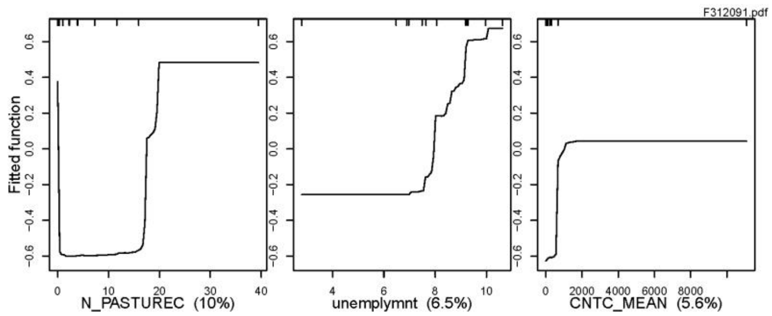



F311086.pdf
Fitted function
ResDay (7.7%)
SfArea (6.9%)
IssueYear (5.1%)

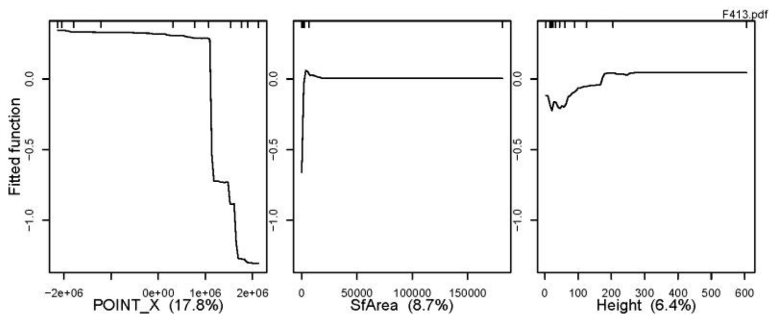

F312087.pdf
Fitted function
unemplymnt (13.4%)
wshedG_PA (9.7%)
dist2Mouth (8.8%)

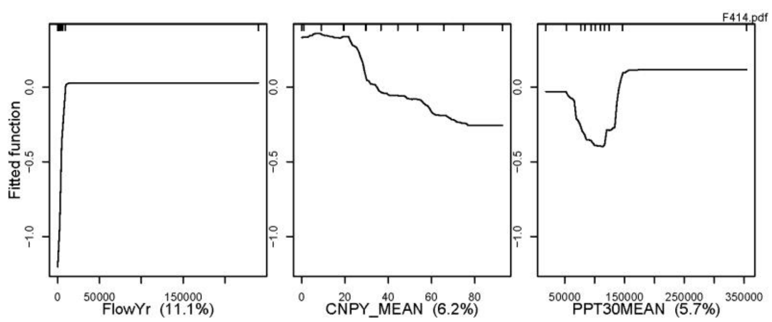

F312088.pdf
Fitted function
wshedG_PA (21%)
dist2Mouth (13.3%)
unemplymnt (11.8%)


F312090.pdf
Fitted function
wshedG_PA (12.1%)
KFACT (9.3%)
Length (6%)

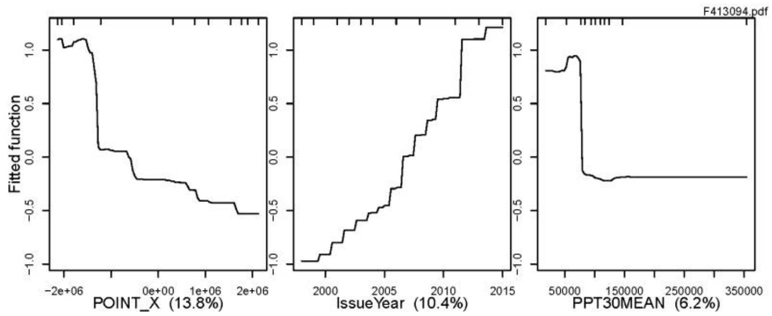


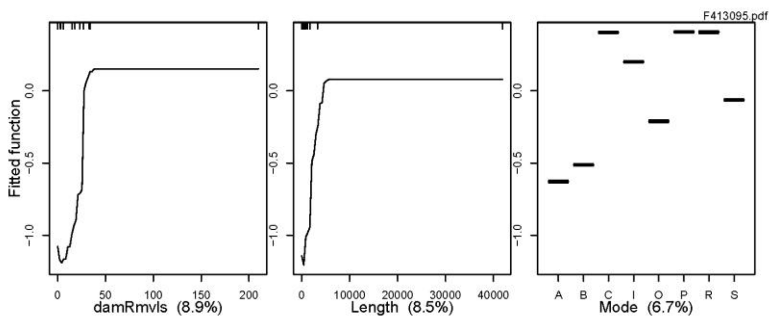
Tier 2 Biodiversity

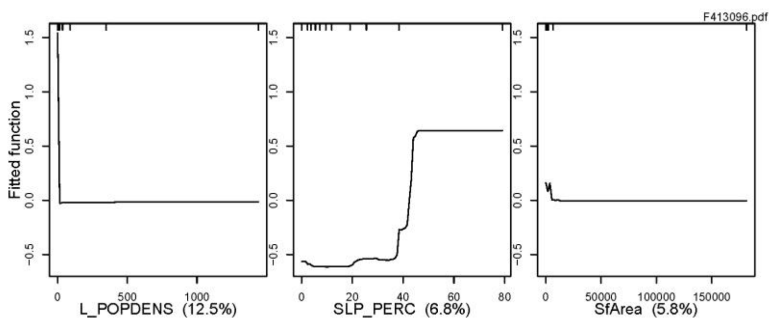


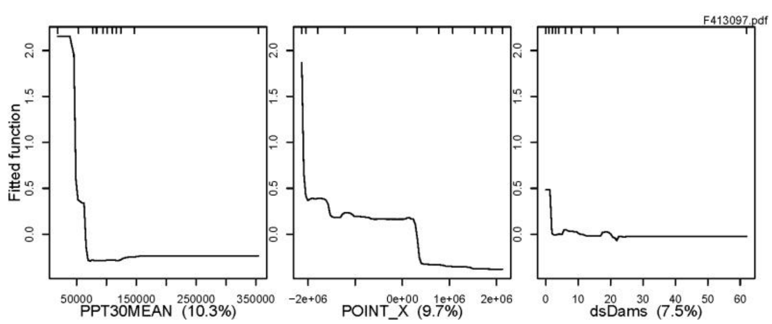


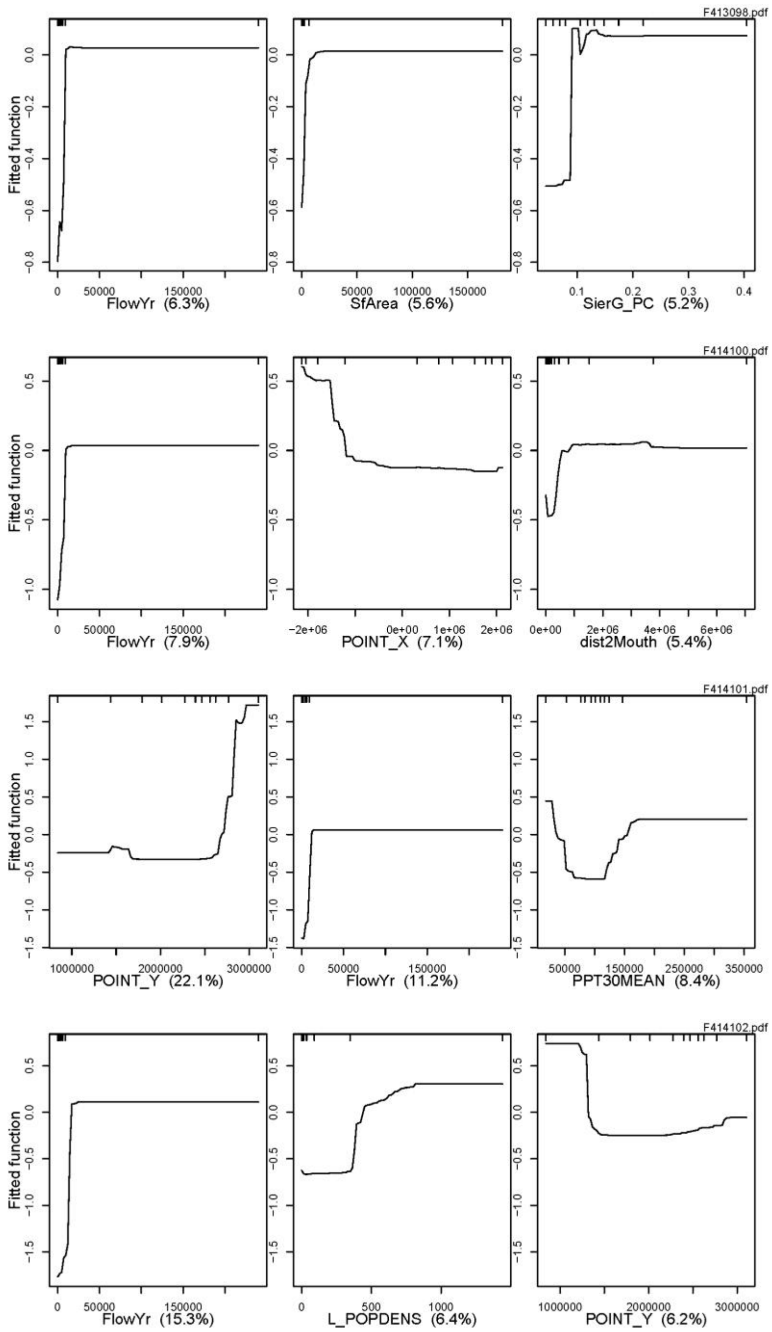
Tier 3 Biodiversity



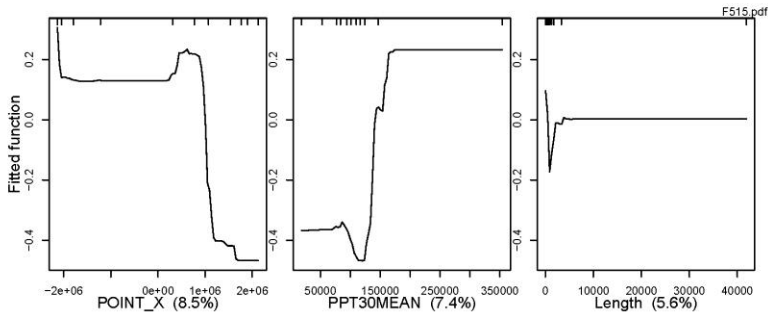


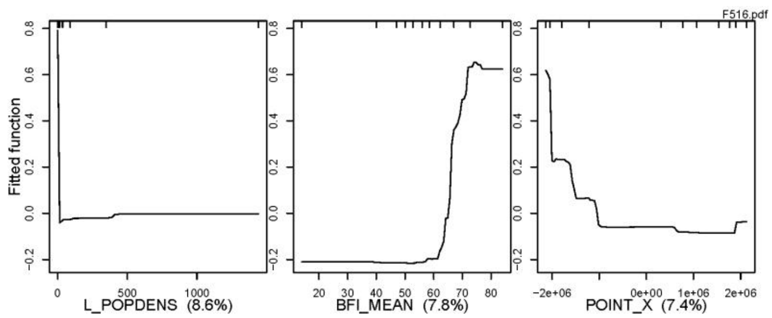


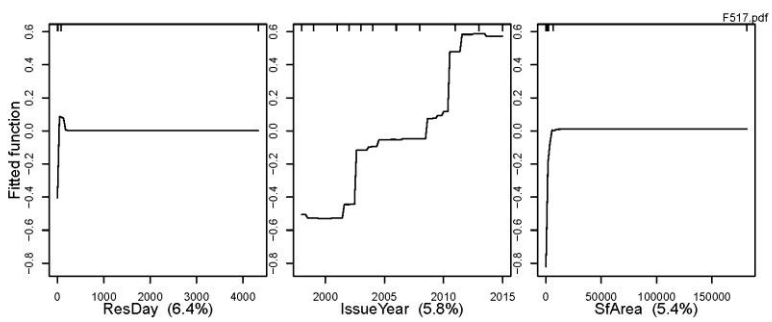


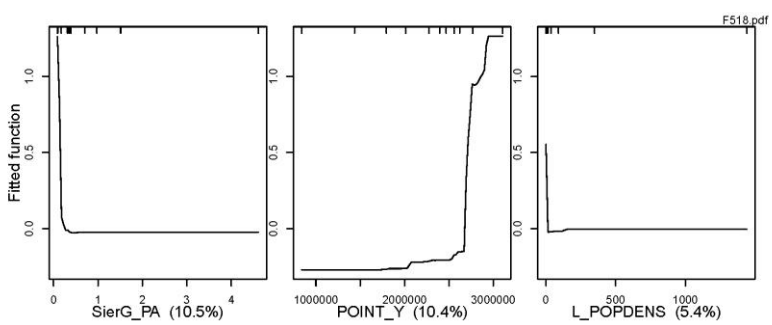

F413098.pdf
Fitted function
FlowYr (6.3%)
SfArea (5.6%)
SierG_PC (5.2%)
F414100.pdf
Fitted function
FlowYr (7.9%)
POINT_X (7.1%)
dist2Mouth (5.4%)
F414101.pdf
Fitted function
POINT_Y (22.1%)
FlowYr (11.2%)
PPT30MEAN (8.4%)
F414102.pdf
Fitted function
FlowYr (15.3%)
L_POPDENS (6.4%)
POINT_Y (6.2%)

Tier 2 Habitat

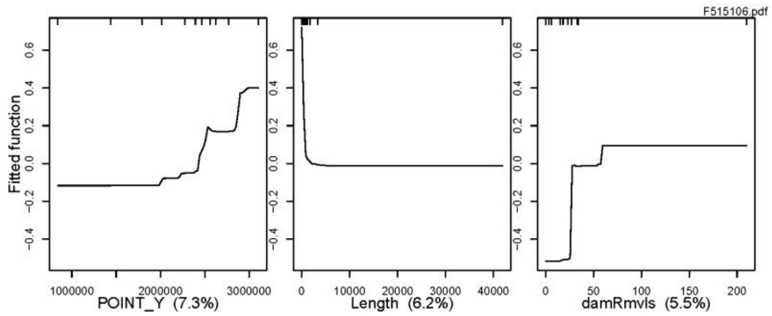


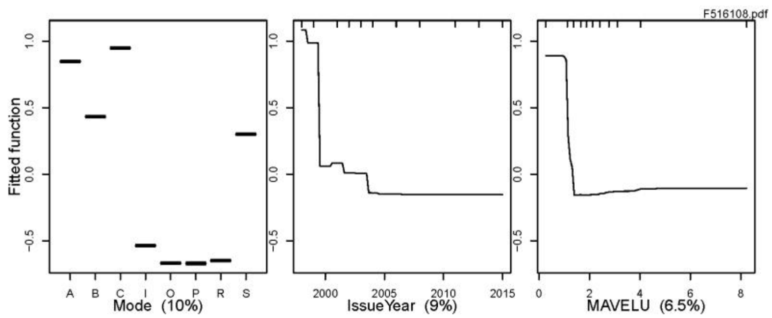


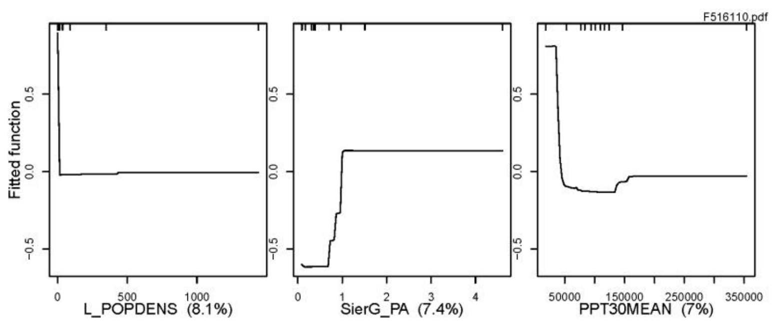


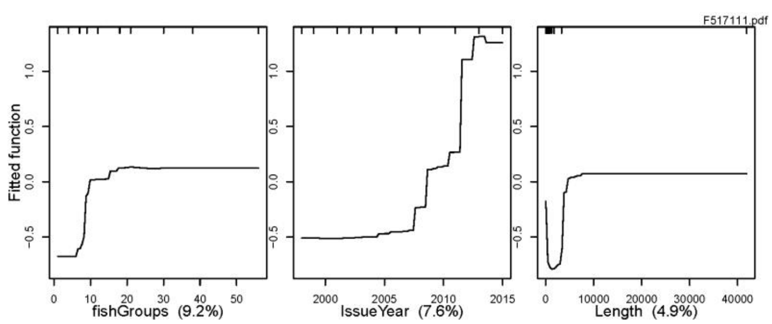


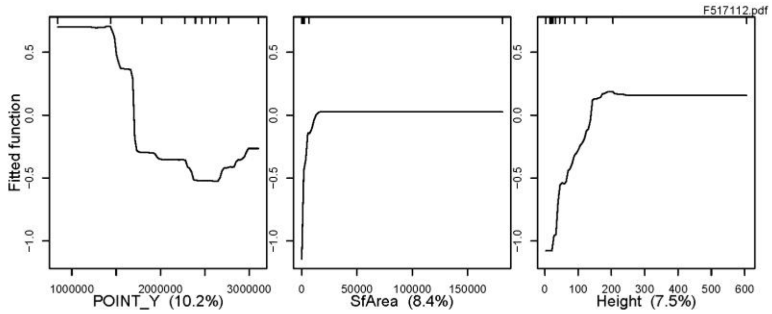
Tier 3 Habitat

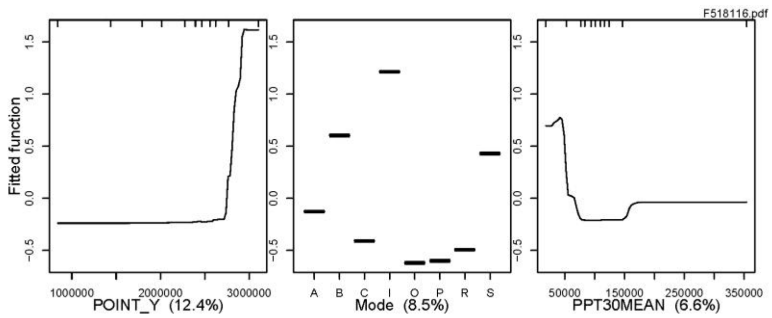




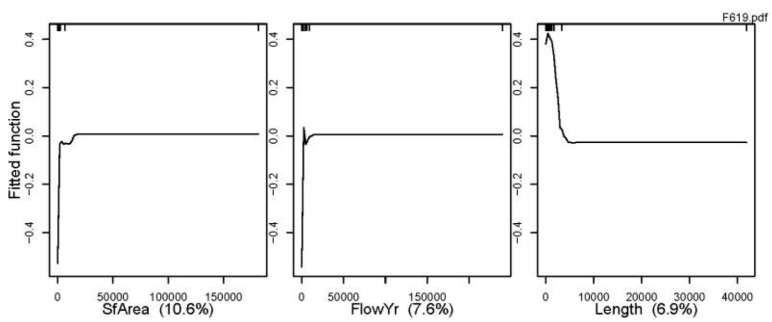


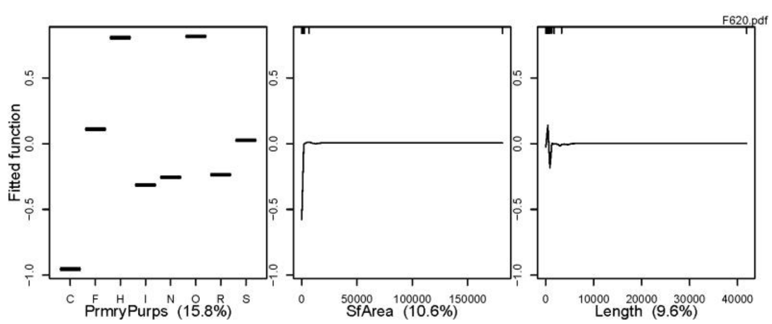




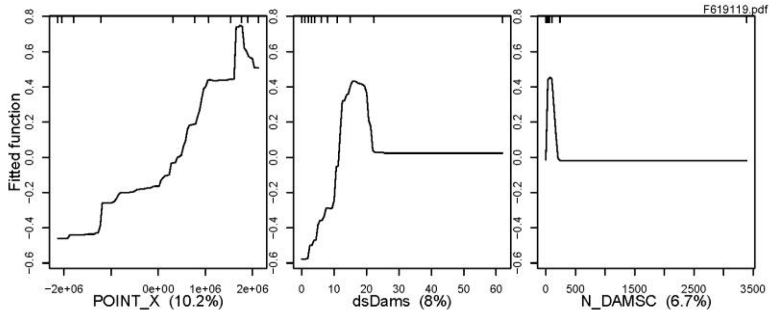


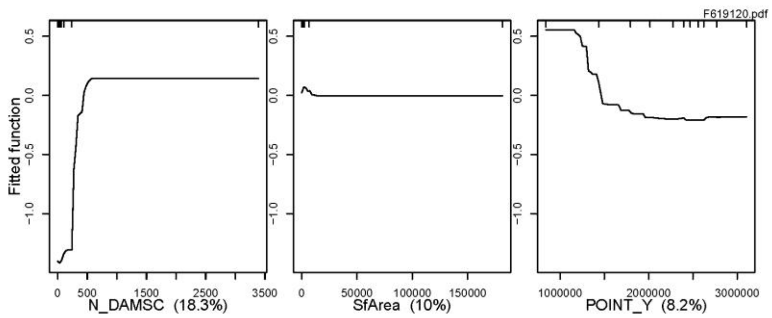
Tier 2 Recreation

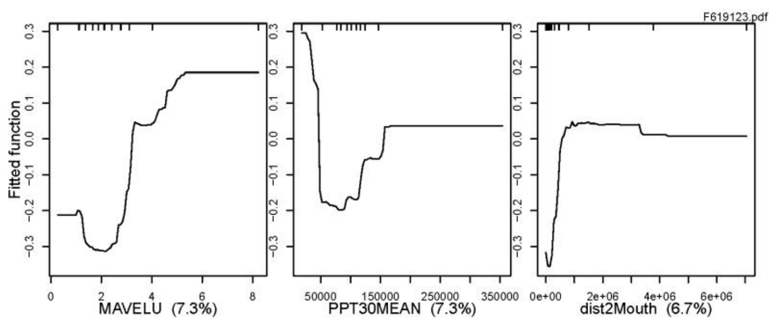


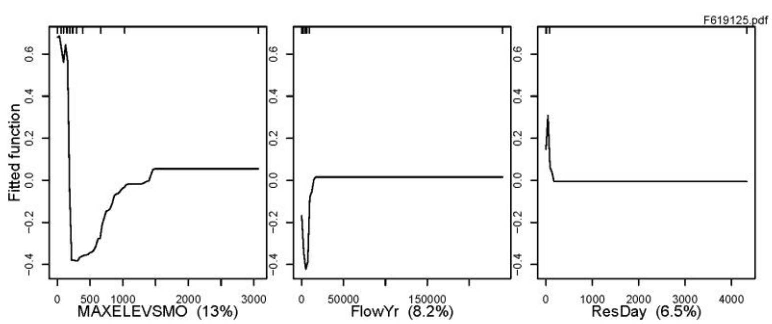


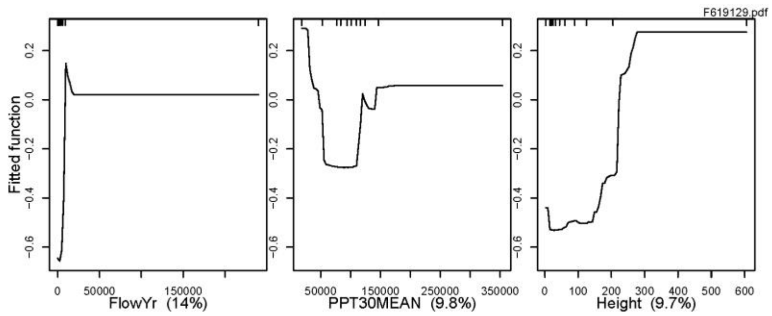
Tier 3 Recreation

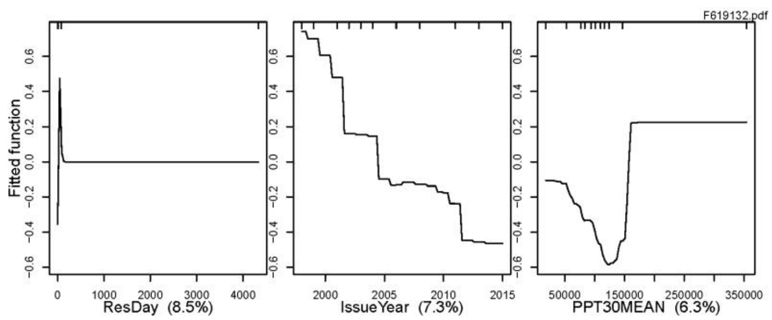


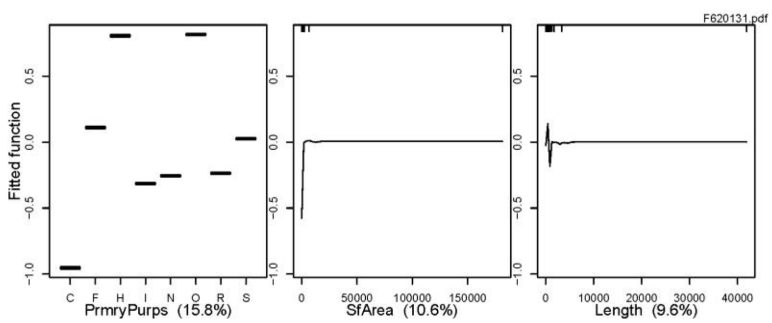







F619129.pdf
Fitted function
FlowYr (14%)
PPT30MEAN (9.8%)
Height (9.7%)


F619132.pdf
Fitted function
ResDay (8.5%)
IssueYear (7.3%)
PPT30MEAN (6.3%)


F620131.pdf
Fitted function
C F H I N O R S
PrmryPurps (15.8%)
SfArea (10.6%)
Length (9.6%)

Predicting Environmental Mitigation Requirements for Hydropower Projects through the Integration of Biophysical and Socio-Political Geographies

Christopher R. DeRolph*[a], Michael P. Schramm[a], and Mark S. Bevelhimer[a]

[a] Oak Ridge National Laboratory, Environmental Sciences Division, PO Box 2008, 1 Bethel Valley Road, Oak Ridge, TN 37831-6038

*corresponding author: derolphcr@ornl.gov

ABSTRACT

Uncertainty about environmental mitigation needs at existing and proposed hydropower projects makes it difficult for stakeholders to minimize environmental impacts. Hydropower developers and operators desire tools to better anticipate mitigation requirements, while natural resource managers and regulators need tools to evaluate different mitigation scenarios and order effective mitigation. Here we sought to examine the feasibility of using a suite of multi-faceted explanatory variables within a spatially explicit modeling framework to fit predictive models for future environmental mitigation requirements at hydropower projects across the conterminous U.S. Using a database comprised of mitigation requirements from more than 300 hydropower project licenses, we were able to successfully fit models for nearly 50 types of environmental mitigation and to apply the predictive models to a set of more than 500 non-powered dams identified as having hydropower potential. The results demonstrate that mitigation requirements are functions of a range of factors, from biophysical to socio-political. Project developers can use these models to inform cost projections and design considerations, while regulators can use the models to more quickly identify likely environmental issues and potential solutions, hopefully resulting in more timely and more effective decisions on environmental mitigation.

Keywords hydropower, mitigation, modeling, prediction, environmental, sociopolitical

1. INTRODUCTION

Hydroelectric power is currently the largest of the renewable energy resources worldwide, contributing to electricity generation in 160 countries (Manzano-Agugliaro et al., 2013). The environmental impacts of hydropower are well established (Liermann et al., 2012; Nilsson et al., 2005; Poff et al., 1997; Poff et al., 2007), and are mitigated with mixed success (Trussart et al., 2002). In the United States (U.S.), the authority to issue 30-50 year licenses for the operation of non-federal hydropower facilities belongs to the U.S. Federal Energy Regulatory Commission (FERC). The passage of the Electric Consumers Protection Act of 1986 (ECPA) substantially changed FERC's consideration of environmental impacts with the requirement that equal consideration be given to the protection and enhancement of, and mitigation of damage to, wildlife, environmental quality, and recreational opportunity. Furthermore, a string of court rulings eroded FERC's singular authority to prescribe environmental mitigation by requiring FERC to include fishway prescriptions from the National Marine Fisheries Service or U.S. Fish and Wildlife Service, as well as minimum streamflow requirements included as part of state water quality certificates (Blumm and Nadol, 2001; Tarlock, 2012). The result was a significant increase in the number of mitigation requirements included in FERC licenses and a growing role of other federal and state agencies in the licensing process (Blumm and Nadol, 2001; Deshazo and Freeman, 2005; Kosnik, 2010).

FERC and the hydropower industry have suggested that this instable policy context and increased regulatory plurality have resulted in increased licensing time and increased uncertainty in mitigation requirements (FERC, 2001; U.S.-Congress, 2012). Original licenses for new projects and relicensing of existing projects provide a once in every 30 to 50 year opportunity to address environmental concerns at hydropower projects. FERC addresses potential environmental impacts by incorporating license conditions (mitigation requirements) where evidence shows project operations will impact environmental or recreational resources. With over 300 relicense applications anticipated between 2016 and 2026 (FERC, 2015), there is new urgency to integrate sustainability practices into future hydropower development by evaluating and balancing the environmental and social costs of hydropower projects with the variety of potential benefits hydropower projects provide.

Hydropower developers and owners desire some certainty and ability to better anticipate mitigation requirements. Similarly, resource managers and regulators must be able to evaluate likely mitigation scenarios and determine the relative effectiveness of mitigation implemented at similar projects. While each hydropower project is unique, Yu and Xu (2016) recommend development of common approaches and principles for designing ecological and social compensation mechanisms for hydropower development. The authors of this manuscript recently developed a database of environmental mitigation requirements in FERC licenses (Schramm et al., 2016) that presents new opportunities for analyzing past environmental mitigation requirements and predicting future mitigation requirements.

In this study we sought to examine the feasibility of using a suite of multidisciplinary explanatory variables to fit predictive models for environmental mitigation requirements at hydropower projects across the conterminous U.S. We developed a spatially explicit framework (applying niche modeling concepts common in landscape ecology) to predict nearly 50 types of environmental mitigation requirements using biological, facility, human, hydrologic, landscape, locational, and stream network characteristics. Our primary goal was to build statistical models to predict future mitigation requirements at hydropower project sites, while the secondary goal was to gain some understanding into potential key environmental and social drivers of these requirements that may warrant additional future research. As an example of how the models can be applied, we made predictions to a set of non-powered dams (NPDs) across the U.S. that were previously identified as having considerable energy potential.

2. MATERIALS AND METHODS

The conterminous United States (Fig. 1) is environmentally and culturally heterogeneous (Fig. 2), containing diverse physiographic regions ranging from mountains to inland and coastal plains, and encompassing examples of nearly every global climate. There is also considerable geographic variation in socio-political, economic, and cultural characteristics.

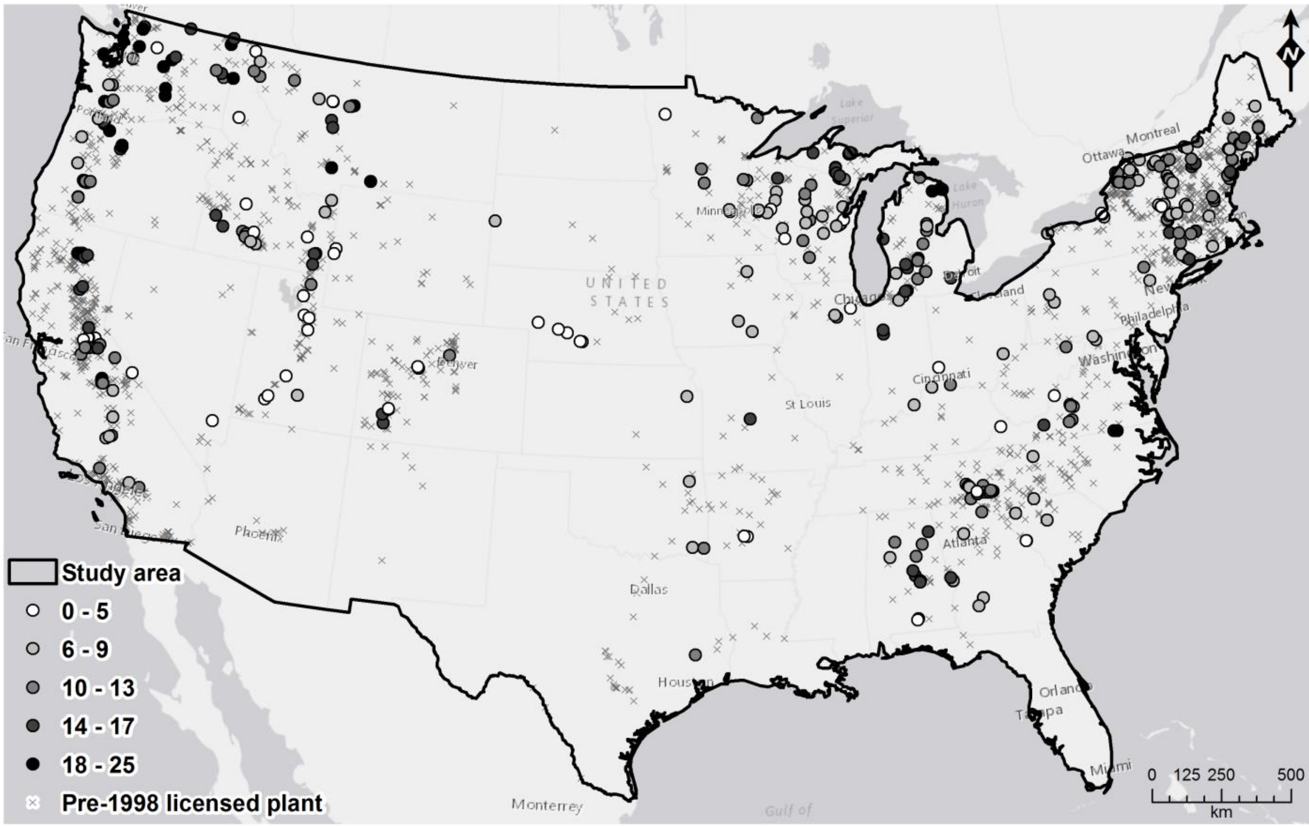


Fig. 1. Study area showing location of 463 hydropower plants licensed from 1998 through September 2015. Color of plant locations indicates number of mitigation requirements for mitigation categories selected for statistical modeling.

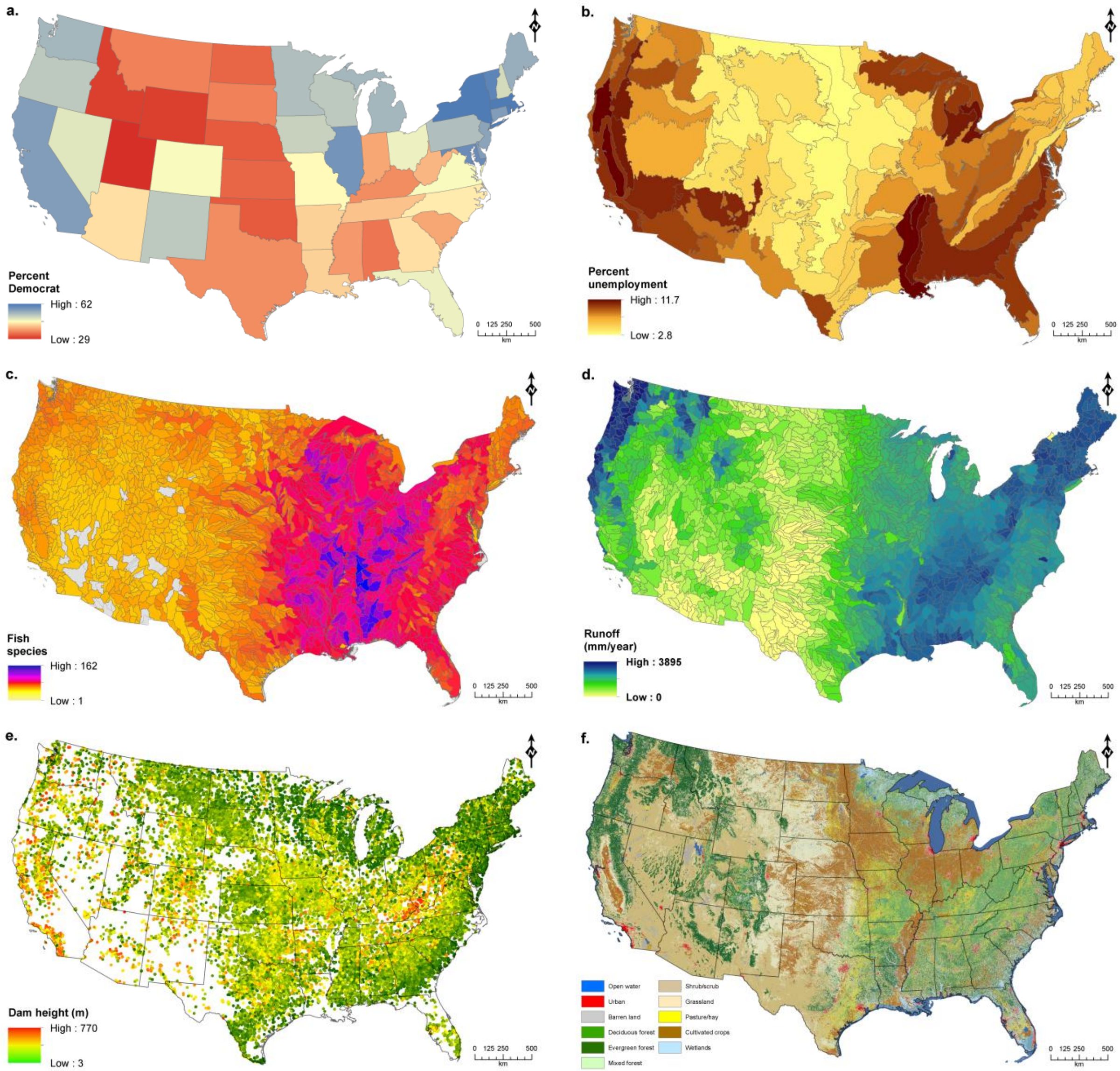


Fig. 2. a. Average percent democratic votes cast in U.S. presidential elections from 1996 to 2012. b. Percent unemployment from 2010 U.S. census, aggregated to physiographic region. c. Number of freshwater fish species per HUC8 watershed. d. Average annual runoff per HUC8 watershed. e. Dam locations symbolized by height. f. 2011 land cover.

2.1. Mitigation database and response variables

A database of environmental mitigation requirements was compiled for FERC licenses issued from 1998 through September 2015 (Schramm et al., 2016). Since our goal is prediction of future mitigation requirements, the manual review of licenses was limited to those issued from 1998 through 2015 with an assumption that more recently issued licenses would better reflect future mitigation requirements. The database includes Bernoulli distributed presence-absence mitigation data at 463 hydropower plants in the study area from 316 licenses. Six broad categories (Tier 1) of mitigation (biodiversity, fish passage, habitat, hydrology, recreation, and water quality) and 20 subcategories (Tier 2) were used to classify specific mitigation types in the hierarchical database. A full list of each of the mitigation types catalogued in the database and the percent of times each was required, including each of the 132 Tier 3 categories, is presented in Appendix A. Descriptions of each of the Tier 3 categories is provided in Appendix A of Schramm et al. (2016). Predictive models were built only if a mitigation type was required for at least 5% (Rickbeil et al., 2014) of the plants in the mitigation database. Models were not built for the very broad Tier 1 categories.

2.2. Explanatory variables

Given that hydropower project licensing is influenced by a suite of biophysical and socio-economic factors, the candidate predictor variables (Table 1) employed here were selected based on expert opinion and on previous research by Kosnik (2010) and Trussart (2002) as broad-scale measures of biological, facility, human, hydrologic, landscape, locational, and stream network characteristics thought to have some bearing on mitigation requirements. The models that each candidate predictor was included in are indicated in Table 1. We used expert opinion to identify candidate predictors for each of the six Tier 1 categories, and these six predictor sets were then used to build models for each Tier 2 and Tier 3 model nested within the Tier I categories. Given that our goal was prediction and not explanation, we did not delve into the exact causal role of each potential predictor. Instead, we selected predictors based on hypothesized quality of association between the predictor and the response, data quality, and data availability (Shmueli, 2010).

2.2.1. Biological

The presence or absence of important fish species can influence not only fish passage mitigation requirements but also other measures related to biological conservation (Cada, 1998; Fraley et al., 1989; Renofalt et al., 2010). We used conservation status in concert with expert opinion to compile a list of high profile migratory fish species supported by policy protections (McManamay et al., 2015). We then mapped distributions of each of these species using the NatureServe (2010) database of current

Table 1. Summary and description of input variables for the boosted regression tree models. Variables or units in bold and underlined indicate remaining predictor variables after collinearity analysis.

Variable	Description	Units	Spatial scale	Source	Models
		Biological			
bigPlyrSum	Major migratory fish species	Count	HUC8 watershed	NatureServe fish distributions, expert opinion	P, H, B, A, R
		Facility			
Height	Dam height	Feet	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
HY_MW	Generation capacity	Megawatts	Hydropower dam	ORNL NHAAP database	--
HY_MWh	Generation	Megawatt-hours	Hydropower plant	ORNL NHAAP database	--
Length	Dam length	Feet	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
Mode	Dam mode-of-operation	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
owner	Ownership type	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
PrmyPurps	Dam primary purpose	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
		Human			
birdG_xx	National Audubon Society chapters	Count, PA, PC	State	National Audubon Society	--
damR_xx	Dam removals	**Count**, PA, PC	State	American Rivers	P, H, W, B, A, R
education	Education attainment - percent bachelor's degree or higher	Percent	USEPA Level 3 Ecoregion	US Census	P, H, W, B, A, R
FishG_xx	TU and CCA chapters	**Count**, PA, **PC**	State	TU, CCA	P, H, B, A, R
hshldincm	Mean household income	US dollars	USEPA Level 3 Ecoregion	US Census	--
IssueYear	FERC hydropower project license issue year	Year	Hydropower plant	ORNL NHAAP database	P, H, W, B, A, R
LandG_xx	Land trusts	Count, PA, PC	State	Land Trust Alliance	--
politics	see note*	Difference	State	US Federal Election Commission	P, H, W, B, A, R
xx_POPDENS	2000 population density	Individuals/km2 (L and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A, R
q12_avg	Survey response on environmental impact of dams	Rank	Geographic region	MIT Energy Survey, 2008	P, H, W, B, A, R
q16_avg	Survey response on increasing or decreasing hydro power	Rank	Geographic region	MIT Energy Survey, 2008	P, H, W, B, A, R
SierG_xx	Sierra Club chapters	Count, **PA**, **PC**	State	Sierra Club	P, H, W, B, A, R
unemplymnt	Unemployment	Percent	USEPA Level 3 Ecoregion	US Census	P, H, W, B, A, R
usHouse	LCV US House of Rep. mean scorecard for 1998 to 2013	Percent	State	League of Conservation Voters	--
usSenate	LCV US Senate mean scorecard for 1998 to 2013	Percent	State	League of Conservation Voters	--
wshed_xx	Local watershed associations	**Count, PA, PC**	State	USEPA	P, H, W, B, A, R
		Hydrology			
ADRAIN	Total artificial drainage area	Square meters	NHD Plus V1 Catchment	USGS	H, W, A
BFI_MEAN	Mean base-flow index for GW discharge into streams	Percent	NHD Plus V1 Catchment	USGS	H, W, A
CNTC_MEAN	Baseflow residence time in the subsurface	Days	NHD Plus V1 Catchment	USGS	H, W, A
DITCHES	Estimated area subject to the practice of ditches	Square meters	NHD Plus V1 Catchment	USGS	--
FlowYr	Average annual flow	Cubic feet per second	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
IRRIG	Estimated area subject to the practice of irrigation	Square meters	NHD Plus V1 Catchment	USGS	--
KFACT	Soil erodibility factor	Dimensionless	NHD Plus V1 Catchment	USGS	H, W, A
MAVELU	Mean Annual Velocity (fps) at bottom of flowline	Cubic feet per second	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
MEAN_IEOF	Mean value for infiltration-excess overland flow	Percent	NHD Plus V1 Catchment	USGS	H, W, A
MEAN_RCHRG	Mean annual natural groundwater recharge	Millimeters	NHD Plus V1 Catchment	USGS	--
nidStorSum	Accumulated upstream storage	Acre-feet	Hydropower dam	National Anthropogenic Barriers Dataset	H, W, A
ResDay	Reservoir residence time	Days	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
SATOF_MEAN	Average value of saturation overland flow	Percent	NHD Plus V1 Catchment	USGS	H, W, A
SfArea	Reservoir surface area	Acres	Hydropower dam	National Inventory of Dams	P, H, W, B, A, R
Stor	Reservoir storage	Acre-feet	Hydropower dam	National Inventory of Dams	--
TILES	Estimated area of tile drains	Square meters	NHD Plus V1 Catchment	USGS	--
		Landscape			
xx_CROPS	Land cover classified as cultivated crops	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_NPDES	Number of NPDES sites	Count (**L** and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W
xx_PASTURE	Land cover classified as pasture/hay	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_ROADCRC	Road-stream crossings	Count (L and N)	NHD Plus V1 Catchment	USGS, National Fish Habitat Partnership	--
xx_URBANHC	Land cover classified as high intensity urban	Percent (L and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	--
xx_URBANL	Land cover classified as low intensity urban	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_URBANM	Land cover classified as medium intensity urban	Percent (**L** and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
CNPY_MEAN	Mean canopy cover	Percent	NHD Plus V1 Catchment	USGS	H, W, B, A
CROP_AREA	Total crop area for fertilizer/manure derived from land use	Square meters	NHD Plus V1 Catchment	USGS	H, W, B, A
d303_count	Impaired or threatened waters	Count	NHD Plus V1 Catchment	USEPA 303(d) list	H, W, B, A
IMPV_MEAN	Mean impervious surface	Percent	NHD Plus V1 Catchment	USGS	--
L_MINES	Number of mines or mineral processing plants	Count	NHD Plus V1 Catchment	National Fish Habitat Partnership	W
L_ROADLEN	Length of roads	Meters	NHD Plus V1 Catchment	National Fish Habitat Partnership	P, H, W, B, A
MAXELEVSMO	Maximum elevation	Meters	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
PPT30MEAN	30-year (1971-2000) average annual precipitation	Millimeters	NHD Plus V1 Catchment	USGS	P, H, W, B, A, R
SLOPE	Slope of stream reach	Unitless	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
SLP_PERC	Landscape slope	Percent	NHD Plus V1 Catchment	USGS	P, H, W, B, A, R
TMAX30_MEA	30-year (1971-2000) average annual maximum temperature	Celsius	NHD Plus V1 Catchment	USGS	--
		Location			
POINT_X	Longitude	Decimal degrees	Hydropower dam	National Anthropogenic Barriers Dataset	P, H, W, B, A, R
POINT_Y	Latitude	Decimal degrees	Hydropower dam	National Anthropogenic Barriers Dataset	P, H, W, B, A, R
		Stream network			
dist2Mouth	Stream network distance to network mouth	Meters	Entire downstream flow path	Calculated from NHD Plus V1 flowlines	P, H, W, B, A, R
DrArea	Drainage area upstream of dam	Square miles	Hydropower dam	National Inventory of Dams	--
dsDams	Downstream dams on flow path to network mouth	Count	Entire downstream flow path	Calculated from NHD Plus V1 and NABD	P, H, W, B, A, R
N_DAMSC	Number of dams within network catchment	Count	NHD Plus V1 Catchment	National Fish Habitat Partnership	P, H, W, B, A, R
SO	Strahler stream order	Strahler number	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R

PA = per area; PC = per capita; L = local catchment; N = entire network catchment; xx indicates variable derived for multiple units; P = fish passage; H = hydrology; W = water quality; B = biodiversity; A = habitat; R = recreation; *politics is the difference between mean percent democrat and republican from 1996 to 2012 presidential elections; LCV = League of Conservation Voters; TU = Trout Unlimited; CCA = Coastal Conservation Association.

distributions of freshwater fishes of the U.S. at the 8-digit hydrologic unit code (HUC8) scale to allow for analysis of interactions between these high profile species and hydropower project locations. The biological explanatory variable is a count of the number of key fish species per HUC8 (McManamay et al., 2015).

2.2.2. Facility characteristics

The Oak Ridge National Laboratory (ORNL) National Hydropower Asset Assessment Program (NHAAP) is an integrated energy, water, and ecosystem research effort for sustainable hydroelectricity generation and water management. The ORNL NHAAP database (http://nhaap.ornl.gov/) integrates data from multiple data sources and provides the most current, detailed, and spatially comprehensive information for analyzing and visualizing existing U.S. hydropower assets. We included hydropower facility characteristics from the NHAAP database thought to be important drivers of prescribed mitigation such as dam height, generation capacity, dam mode-of-operation, and geographic location (Kosnik, 2010).

2.2.3. Human dimensions

The convergence of different anthropogenic characteristics such as presence of environmental interest groups, political climate, population demographics, and regulatory tendencies can be impactful on mitigation requirements (Kosnik, 2010). Consistent with the interest group theory of regulation (Peltzman, 1976), Knittel (2006) concluded that electricity industry regulators respond to lobbying from interest groups. In research focused on explaining drivers of environmental mitigation requirements at hydropower projects, Kosnik (2010) found the largest influences on FERC's regulatory decisions to be congressional politics and regulatory tendencies. In an attempt to capture the socio-political and regulatory landscape, we included numerous anthropogenic predictors that serve as direct measures or proxies for local, state, or regional political tendencies, environmental awareness, regulatory trends, and public attitudes toward dams. Candidate predictors aimed at capturing political tendencies include presidential election voting averaged over time and congressional politics. Different aspects of environmental awareness were estimated at a state-scale using prevalence of non-profit organizations, including protection of birds and their habitats (using National Audubon Society chapters), fish and their habitats (using Trout Unlimited chapters for freshwater and Coastal Conservation Association chapters for marine), land conservation (using land trusts), water quality conservation (using local watershed associations), and general environmental awareness (using Sierra Club chapters). Regulatory trends were estimated using the issue year of the license. Prevalence of dam removals and citizen survey responses on energy and environmental impacts from dams were used as estimates of public attitudes toward dams.

2.2.4. Hydrology

Operation of a hydropower facility typically involves modifications to hydrologic regimes both upstream and downstream of dams, reservoirs, or river diversions (Fraley et al., 1989; Ligon et al., 1995; Poff et al., 1997). The magnitude of these flow disturbances can be minimized by discharge management, and there is increasing pressure from regulatory agencies to incorporate ecological flow requirements in licenses and operational plans for hydropower projects (Bunn and Arthington, 2002; Renofalt et al., 2010; Trussart et al., 2002). We included a suite of explanatory variables derived at the stream reach and watershed scale that describe different aspects of the hydrologic regime of a given area, including surface water, groundwater, and reservoir storage characteristics.

2.2.5. Landscape

Broad-scale landscape descriptors such as land cover, terrain, and climate can influence prescribed mitigation in all six of the Tier 1 mitigation categories, either directly or indirectly. Thus we included numerous land cover metrics derived at multiple scales (Tong and Chen, 2002; Wang et al., 2001), topographic variables such as slope and elevation (Moore et al., 1991), and the core climatic variables of average annual precipitation and air temperature (Grimm et al., 2008).

2.2.6. Location

In the U.S., there are tangible trends and patterns in environmental, economic, cultural, and social conditions from east to west and north to south. We included latitude and longitude to account for spatial effects and capture spatial patterns across the large study area that may be insufficiently represented in the other predictors (Fink et al., 2010; Oppel et al., 2012).

2.2.7. Stream network

Stream network position and the prevalence of upstream and downstream dams are important descriptors of network fragmentation/connectivity (Kuby et al., 2005). Where a hydropower project falls on the stream network in relation to other barriers and the network mouth can have a strong influence on the nature and magnitude of ordered mitigation (Fraley et al., 1989; Kosnik, 2010).

2.3. Statistical analyses

Model development was carried out in R version 3.2.2 (R-Core-Team). Boosted regression trees (a machine-learning technique) were used to develop the predictive models, as this method has been demonstrated to have high predictive performance with presence-absence response variables, allows for complex regression analyses of complex responses, and can handle continuous and categorical explanatory variables (Abram et al., 2015; Arganaraz et al., 2015; Elith et al., 2006; Elith et al., 2008). Before running the models, all predictor variables were assessed for collinearity using Pearson's correlation coefficients (r). When r values exceeded 0.7 (Dormann et al., 2013), the variable deemed more functionally applicable to hydropower mitigation (Arganaraz et al., 2015; Rickbeil et al., 2014) or that was derived at a higher spatial resolution was retained (Table 1). The data were split into training (80%) and validation (20%) data using the caret package in R, which creates random splits within each class so that the overall class distribution is preserved as well as possible (Kuhn, 2008).

Given the novelty of the mitigation database, we were unable to obtain an independent validation dataset as recommended by Araujo and Guisan (2006). The optimal number of trees was determined using 10-fold cross validation (CV), with the bag fraction set to 0.5 and the learning rate set to 0.001 to ensure that each model had at least 1,000 trees (Elith et al., 2008). The area under the receiver-operating characteristic curve (ROC) calculated on the validation dataset was used to assess predictive performance. We implemented the ROC interpretation presented by Hosmer et al. (2013) where an ROC value of 0.7-

0.8 is considered an acceptable prediction, 0.8-0.9 is excellent, and >0.9 is outstanding. For a model to be deemed acceptable, both the internal CV ROC and the validation ROC had to be ≥0.7. We generated partial dependence plots to examine the nature of the models and to interpret the effect of a variable on the response after accounting for the average effects of all other variables in the model (Elith et al., 2008). Spatial autocorrelation of model residuals was evaluated using Moran's *I* statistics (Dormann et al., 2007) calculated with the Spatial Autocorrelation tool in ArcGIS version 10.2.2 (ESRI).

2.4. Example model application at non-powered dams

While approximately 2,500 dams in the U.S. provide 78 gigawatts (GW) of conventional and 22 GW of pumped-storage hydropower, there are hundreds of NPDs originally built for other purposes that may be retrofitted for hydropower to produce an additional 12 GW of estimated renewable energy for the U.S (Hadjerioua et al., 2012). While many of the monetary costs and environmental impacts have already been incurred at these sites, our models can be used as a tool to assess potential environmental mitigation requirements that may arise during the hydropower licensing process. As an example of how the modeling can be applied, we made predictions for each of the acceptable models to 568 NPDs estimated by Hadjerioua et al. (2012) to have >1 megawatt (MW) in potential capacity. We used the optimal threshold function in the R package SDMtools (VanDerWal et al., 2012) to identify the value on the ROC curve that is closest to a perfect model fit, and then we applied that value as the predicted present/absent threshold when making predictions to the NPDs.

3. RESULTS

3.1. BRT models

Predictive models were built only if a mitigation type was required for at least 5% (Rickbeil et al., 2014) of the plants in the mitigation database, resulting in 57 Tier 3 mitigation types being modelled and all 20 of the Tier 2 mitigations being modelled (see Table 2 for modeling results). Eight of the 57 Tier 3 models were rejected due to either a CV ROC or validation ROC <0.7, leaving 49 Tier 3 models with at least an acceptable fit. All 20 of the Tier 2 models had an ROC ≥0.7. Significant spatial autocorrelation of model residuals was detected in 4 of 20 Tier 2 models and 11 of 49 Tier 3 models.

Table 2. Model results summary.

Tier 1	Tier 2	Tier 3	ModelID	Trees	CV ROC	V ROC	OT	MI	Influential Variable 1	Influential Variable 2	Influential Variable 3
Fish Passage	DS	NA (see Tier 2 category)	F101	5550	0.867	0.916	0.36	-0.165	POINT_Y (13)	dist2Mouth (10.7)	FlowYr (8)
	Fish Passage	DS Passage Plan Study Design	F101010	5300	0.892	0.829	0.30	-0.309	POINT_Y (11.1)	dist2Mouth (9.9)	MAXELEVSMO (9.1)
	US	NA (see Tier 2 category)	F102	5250	0.899	0.896	0.29	-0.436	MAXELEVSMO (14.5)	POINT_Y (13)	bigPlyrSum (8.3)
	Fish Passage	Eelway	F102017	4350	0.956	0.966	0.33	*-1.178*	MAXELEVSMO (36.5)	POINT_X (11.9)	bigPlyrSum (7.7)
		US passage study plan or design	F102023	5000	0.909	0.854	0.29	*-0.703*	MAXELEVSMO (12.3)	POINT_X (10.5)	POINT_Y (9.7)
	Passage	NA (see Tier 2 category)	F103	2850	0.780	0.856	0.27	*-0.467*	POINT_Y (9.3)	dsDams (8.1)	MAXELEVSMO (7.5)
	Planning	DS fish passage mon. sampling	F103029	3200	0.888	0.924	0.22	*-0.908*	MAXELEVSMO (18.2)	Height (10)	POINT_X (8.9)
		Fish passage & operations plan	F103031	1050	0.739	0.749	0.08	-0.246	wshed_PC (11.8)	L_ROADLEN (10.5)	dsDams (9.9)
		Fish stranding plan mon. evaluation	F103033	1100	0.712	0.605	--		--	--	--
		US fish passage mon. sampling	F103036	3050	0.891	0.865	0.18	*-0.504*	MAXELEVSMO (16.2)	Height (11.9)	POINT_Y (8.7)
	Entrainment	NA (see Tier 2 category)	F104	3450	0.849	0.756	0.29	-0.222	ResDay (9.8)	SierG_PC (7.8)	politics (7)
		Trash or bar rack	F104043	3700	0.917	0.833	0.22	0.147	POINT_X (10.4)	SierG_PC (9.6)	fishGroups (8.8)
Hydrology	Flow	NA (see Tier 2 category)	F205	4700	0.785	0.787	0.56	*-1.413*	PrmryPurps (7.7)	N_URBANLC (5.3)	Height (5.1)
	Mitigation	Tailrace flow mon. plan	F205045	5850	0.822	0.920	0.39	-0.076	POINT_X (8.5)	IssueYear (4.8)	politics (4.7)
		Tailrace flow or stage mon. equipment	F205048	2650	0.784	0.867	0.17	-0.191	N_CROPSC (9)	Length (7.3)	dist2Mouth (4.7)
		Tailrace ramping rate restriction	F205050	2400	0.790	0.834	0.19	-0.333	SfArea (8.3)	CNPY_MEAN (5.9)	nidStorSum (5.8)
		Bypass flow mon. plan	F205052	2900	0.802	0.853	0.20	*-0.522*	politics (6.9)	Length (6.7)	SfArea (5.6)
		Bypass flushing or flood flow	F205054	3750	0.890	0.951	0.15	0.187	POINT_X (24.6)	N_PASTUREC (6.8)	PPT30MEAN (6.2)
		Bypass flow or stage mon. equipment	F205055	1500	0.735	0.779	0.13	*-0.679*	SLP_PERC (18)	POINT_X (7)	bigPlyrSum (5.2)
		Bypass ramping rate restriction	F205057	2900	0.878	0.802	0.16	0.170	POINT_X (28.6)	BFI_MEAN (6.7)	CNTC_MEAN (6.5)
	Tailrace	NA (see Tier 2 category)	F206	3800	0.863	0.845	0.63	-0.198	POINT_X (10.4)	Mode (10.3)	SLP_PERC (7.2)
	Minimum	Run-of-river Tailrace	F206058	3700	0.904	0.911	0.37	-0.349	Mode (39.2)	Height (7.8)	POINT_X (4.8)
	Flow	Seasonal Tailrace	F206059	2700	0.850	0.846	0.20	0.087	Mode (24.6)	POINT_Y (9.2)	N_PASTUREC (5.7)
		Year-round Tailrace	F206061	1500	0.787	0.899	0.19	-0.312	Mode (22.7)	owner (9.2)	Length (4.3)
	Bypass	NA (see Tier 2 category)	F207	3250	0.808	0.771	0.46	*-0.676*	SfArea (16.3)	MAXELEVSMO (6.6)	MAVELU (4.7)
	Minimum	Seasonal Bypass	F207063	1450	0.678	0.668	--		--	--	--
	Flow	Year-round Bypass	F207065	1200	0.720	0.805	0.23	-0.339	SfArea (15.6)	Height (10.2)	Length (6.9)
	Sediment	NA (see Tier 2 category)	F208	4850	0.767	0.851	0.49	-0.364	IssueYear (6.1)	CNPY_MEAN (5.3)	unemplymnt (4.6)
		Sediment & erosion control plan or mon.	F208066	4100	0.778	0.838	0.47	-0.257	IssueYear (6.6)	CNPY_MEAN (5)	dist2Mouth (4.9)
	Recreation	NA (see Tier 2 category)	F209	1550	0.733	0.796	0.17	-0.127	POINT_X (12.3)	Height (6.8)	SierG_PA (6.8)
	Flow	Provide recreational flow releases	F209071	700	0.655	0.713	--		--	--	--
	Operations	NA (see Tier 2 category)	F210	3050	0.734	0.819	0.53	0.040	FlowYr (8.6)	q16_avg (4.5)	N_URBANLC (4.5)
		Flow mgmt. plan	F210073	3350	0.893	0.985	0.09	-0.399	Length (13)	wshedG_PA (11.4)	IssueYear (8.6)
		Operations compliance mon. plan	F210074	5150	0.807	0.913	0.41	0.146	politics (8.8)	PrmryPurps (4.8)	FlowYr (4.7)
		Provide flow or lake levels electronically	F210075	1750	0.795	0.917	0.14	-0.282	SierG_PA (18.8)	POINT_X (6.2)	Mode (6.1)
Water Quality	DS	NA (see Tier 2 category)	F311	4300	0.838	0.887	0.55	-0.240	ResDay (12.3)	SfArea (6.8)	wshedG_PA (4.6)
	Water	Benthic macroinvertebrate mon.	F311077	1500	0.724	0.938	0.12	-0.396	BFI_MEAN (16.9)	unemplymnt (8.3)	POINT_X (7.8)
	Quality	DO enhancement or mitigation plan	F311078	2200	0.832	0.676	--		--	--	--
		Water quality mon. plan	F311086	6000	0.852	0.873	0.50	-0.375	ResDay (7.7)	SfArea (6.9)	IssueYear (5.1)
	US	NA (see Tier 2 category)	F312	4000	0.831	0.860	0.23	-0.320	unemplymnt (7.6)	POINT_Y (6.8)	N_PASTUREC (6.8)
	Water	Fish tissue sampling & analysis	F312087	4500	0.965	0.823	0.39	*-1.704*	unemplymnt (13.4)	wshedG_PA (9.7)	dist2Mouth (8.8)
	Quality	Impoundment sediment analysis	F312088	4100	0.993	0.999	0.19	0.334	wshedG_PA (21)	dist2Mouth (13.3)	unemplymnt (11.8)
		Inflow water quality mon. plan	F312090	1650	0.831	0.904	0.11	0.106	wshedG_PA (12.1)	KFACT (9.3)	Length (6)
		Impoundment water quaity mon. plan	F312091	4100	0.828	0.805	0.22	-0.166	N_PASTUREC (10)	unemplymnt (6.5)	CNTC_MEAN (5.6)
Biodiversity	Terrestrial	NA (see Tier 2 category)	F413	4150	0.847	0.832	0.64	*-0.638*	POINT_X (17.8)	SfArea (8.7)	Height (6.4)
		Noxious weed & invasive plant mgmt.	F413094	6650	0.912	0.901	0.39	0.068	POINT_X (13.8)	IssueYear (10.4)	PPT30MEAN (6.2)
		Species conservation mgmt. mon.	F413095	5850	0.832	0.899	0.40	-0.265	damRmvls (8.9)	Length (8.5)	Mode (6.7)
		T&E species protection plan	F413096	3950	0.879	0.905	0.21	*0.965*	L_POPDENS (12.5)	SLP_PERC (6.8)	SfArea (5.8)
		Transmission related avian & bat protection	F413097	6250	0.936	0.941	0.19	-0.109	PPT30MEAN (10.3)	POINT_X (9.7)	dsDams (7.5)
		Wildlife terrestrial habitat mgmt.	F413098	4100	0.844	0.937	0.27	0.146	FlowYr (6.3)	SfArea (5.6)	SierG_PC (5.2)
	Aquatic	NA (see Tier 2 category)	F414	3500	0.791	0.859	0.35	-0.271	FlowYr (11.1)	CNPY_MEAN (6.2)	PPT30MEAN (5.7)
		Aquatic species conservation mgmt. mon.	F414100	3400	0.807	0.869	0.34	-0.336	FlowYr (7.9)	POINT_X (7.1)	dist2Mouth (5.4)
		Diadromous species mgmt. mon.	F414101	3000	0.871	0.901	0.26	0.124	POINT_Y (22.1)	FlowYr (11.2)	PPT30MEAN (8.4)
		Invasive aquatic species mgmt.	F414102	2800	0.800	0.881	0.19	*0.552*	FlowYr (15.3)	L_POPDENS (6.4)	POINT_Y (6.2)
Habitat	Fisheries	NA (see Tier 2 category)	F515	2650	0.776	0.730	0.27	-0.038	POINT_X (8.5)	PPT30MEAN (7.4)	Length (5.6)
		DS habitat enhancement	F515105	1200	0.687	0.680	--		--	--	--
		DS woody debris restoration or passage	F515106	2850	0.863	0.879	0.25	0.071	POINT_Y (7.3)	Length (6.2)	damRmvls (5.5)
	Riparian	NA (see Tier 2 category)	F516	2600	0.771	0.869	0.28	0.084	L_POPDENS (8.6)	BFI_MEAN (7.8)	POINT_X (7.4)
		Establish riparian buffers	F516108	3100	0.866	0.864	0.25	*0.673*	Mode (10)	IssueYear (9)	MAVELU (6.5)
		Riparian habitat mon. or planning	F516110	2300	0.793	0.912	0.22	-0.368	L_POPDENS (8.1)	SierG_PA (7.4)	PPT30MEAN (7)
	Reservoir	NA (see Tier 2 category)	F517	6100	0.858	0.905	0.40	-0.082	ResDay (6.4)	IssueYear (5.8)	SfArea (5.4)
		Noxious invasive aquatic plant mgmt.	F517111	6950	0.928	0.952	0.25	*-0.348*	fishGroups (9.2)	IssueYear (7.6)	Length (4.9)
		Shoreline mgmt. plan or program	F517112	4800	0.856	0.952	0.27	0.192	POINT_Y (10.2)	SfArea (8.4)	Height (7.5)
	Wetlands	NA (see Tier 2 category)	F518	3800	0.828	0.874	0.19	0.207	SierG_PA (10.5)	POINT_Y (10.4)	L_POPDENS (5.4)
		Wetland protection	F518116	3500	0.878	0.875	0.14	0.082	POINT_Y (12.4)	Mode (8.5)	PPT30MEAN (6.6)
Recreation	Resources	NA (see Tier 2 category)	F619	3200	0.741	0.744	0.65	0.089	SfArea (10.6)	FlowYr (7.6)	Length (6.9)
	and	Boating facilities	F619118	1200	0.625	0.660	--		--	--	--
	Mitigation	Canoe portage launch	F619119	5000	0.859	0.773	0.37	-0.161	POINT_X (10.2)	dsDams (8)	N_DAMSC (6.7)
		Fishing pier	F619120	1700	0.797	0.869	0.13	0.055	N_DAMSC (18.3)	SfArea (10)	POINT_Y (8.2)
		Interpretive education sign & displays	F619123	1900	0.720	0.731	0.20	-0.192	MAVELU (7.3)	PPT30MEAN (7.3)	dist2Mouth (6.7)
		Parking	F619125	3550	0.715	0.722	0.32	0.441	MAXELEVSMO (13)	FlowYr (8.2)	ResDay (6.5)
		Shoreline access	F619128	750	0.623	0.759	--		--	--	--
		Stocking recreational fish species	F619129	1250	0.756	0.796	0.09	0.069	FlowYr (14)	PPT30MEAN (9.8)	Height (9.7)
		Trail trailhead or camping areas	F619130	3200	0.781	0.601	--		--	--	--
		Other day use area improvements	F619132	4900	0.750	0.781	0.44	-0.430	ResDay (8.5)	IssueYear (7.3)	PPT30MEAN (6.3)
	Planning	NA (see Tier 2 category)	F620	2450	0.753	0.883	0.74	-0.001	PrmryPurps (15.8)	SfArea (10.6)	Length (9.6)
		Recreational mgmt. plan study or mon.	F620131	2450	0.753	0.883	0.74	-0.001	PrmryPurps (15.8)	SfArea (10.6)	Length (9.6)

See Table 1 for variable descriptions; if no influential variables are shown, model rejected due to poor fit; mgmt. = management; DS = downstream; US = upstream; T&E = threatened and endangered, mon. = monitoring; NA = not applicable; CV ROC = internal cross-validation ROC; V ROC = validation ROC; OT = optimal threshold; MI = Moran's Index; italics indicates spatial autocorrelation detected in training data; color scheme for influential variables corresponds to Table 1 color scheme.

3.2. Explanatory variables

The three variables with the highest relative influence in each model are presented in Table 2, and partial dependence plots for these variables are presented in Appendix B. Overall, we considered a variable important if its relative influence was ≥5% (Parisien et al., 2011). A summary of the important variables for the Tier 3 models (Fig. 3) shows that nearly all the categories of variables (i.e. biological, facility, human, hydrologic, landscape, locational, and stream network) were influential within each Tier 1 category.

Fish passage (n=7)		
Variable	**F**	**Inf**
POINT_X	7	0.63
MAXELEVSMO	6	0.92
POINT_Y	4	0.67
bigPlyrSum	4	0.41
Height	3	0.58
dist2Mouth	3	0.57
dsDams	2	0.74
fishG_PC	2	0.39
FlowYr	2	0.33
wshed_PC	1	1.00
SierG_PC	1	0.92
L_ROADLEN	1	0.89
fishGroups	1	0.84
politics	1	0.74
PrmryPurps	1	0.62
SfArea	1	0.50
SLOPE	1	0.43
MAVELU	1	0.43

Hydrology (n=15)		
Variable	**F**	**Inf**
Mode	5	0.79
POINT_X	5	0.74
Length	4	0.81
SfArea	3	0.94
nidStorSum	3	0.64
Height	3	0.45
politics	2	1.00
IssueYear	2	0.83
CNPY_MEAN	2	0.73
N_PASTUREC	2	0.25
unemplymnt	2	0.22
SierG_PA	1	1.00
N_CROPSC	1	1.00
SLP_PERC	1	1.00
wshedG_PA	1	0.88
fishG_PC	1	0.63
owner	1	0.40
POINT_Y	1	0.37
FlowYr	1	0.33
bigPlyrSum	1	0.29
dist2Mouth	1	0.29
PPT30MEAN	1	0.25
BFI_MEAN	1	0.24
CNTC_MEAN	1	0.23
MAVELU	1	0.21
SierG_PC	1	0.19

Water quality (n=7)		
Variable	**F**	**Inf**
unemplymnt	4	0.68
wshedG_PA	3	0.91
BFI_MEAN	3	0.61
dist2Mouth	3	0.57
SierG_PC	3	0.41
SfArea	2	0.87
IssueYear	2	0.56
POINT_X	2	0.44
N_PASTUREC	1	1.00
N_DAMSC	1	1.00
ResDay	1	1.00
KFACT	1	0.76
POINT_Y	1	0.67
CNTC_MEAN	1	0.55
Length	1	0.50
damRmvls	1	0.44
Height	1	0.35

Biodiversity (n=8)		
Variable	**F**	**Inf**
FlowYr	4	0.88
PPT30MEAN	4	0.66
POINT_X	3	0.95
SfArea	3	0.57
L_POPDENS	2	0.71
POINT_Y	2	0.70
SLP_PERC	2	0.55
damRmvls	1	1.00
Length	1	0.95
SierG_PC	1	0.82
Mode	1	0.75
IssueYear	1	0.75
dsDams	1	0.73
PrmryPurps	1	0.70
dist2Mouth	1	0.69
Height	1	0.38

Habitat (n=6)		
Variable	**F**	**Inf**
POINT_Y	3	1.00
Mode	3	0.74
PPT30MEAN	3	0.64
IssueYear	2	0.86
fishG_PC	2	0.59
L_POPDENS	1	1.00
fishGroups	1	1.00
SierG_PA	1	0.91
Length	1	0.85
SfArea	1	0.82
damRmvls	1	0.75
Height	1	0.73
MAVELU	1	0.65
education	1	0.59
wshedG_PA	1	0.42
PrmryPurps	1	0.41

Recreation (n=10)		
Variable	**F**	**Inf**
FlowYr	7	0.72
PPT30MEAN	6	0.72
ResDay	5	0.59
MAXELEVSMO	4	0.80
Height	4	0.66
Length	4	0.66
dist2Mouth	4	0.65
SfArea	4	0.56
L_POPDENS	3	0.73
dsDams	3	0.54
MAVELU	2	0.84
N_DAMSC	2	0.83
IssueYear	2	0.77
SLOPE	2	0.72
POINT_X	2	0.70
POINT_Y	2	0.54
PrmryPurps	1	1.00
unemplymnt	1	0.48
Mode	1	0.47

Fig. 3. Explanatory variables with relative influence ≥5 for Tier 3 models, broken down by Tier 1 category. Relative influence normalized to 0 to 1 scale for each model; Inf = mean relative influence for variable across all models in which relative influence ≥5; F = frequency, or number of times variable had Inf ≥5; color scheme corresponds to Table 1.

Across all Tier 3 models (Fig. 4), the most important variables were longitude (location), reservoir surface area (hydrology), average annual flow (hydrology), precipitation (landscape), and latitude (location). Stream network, facility, human, and biological variables were also important but exceeded the ≥5% relative influence threshold less frequently.

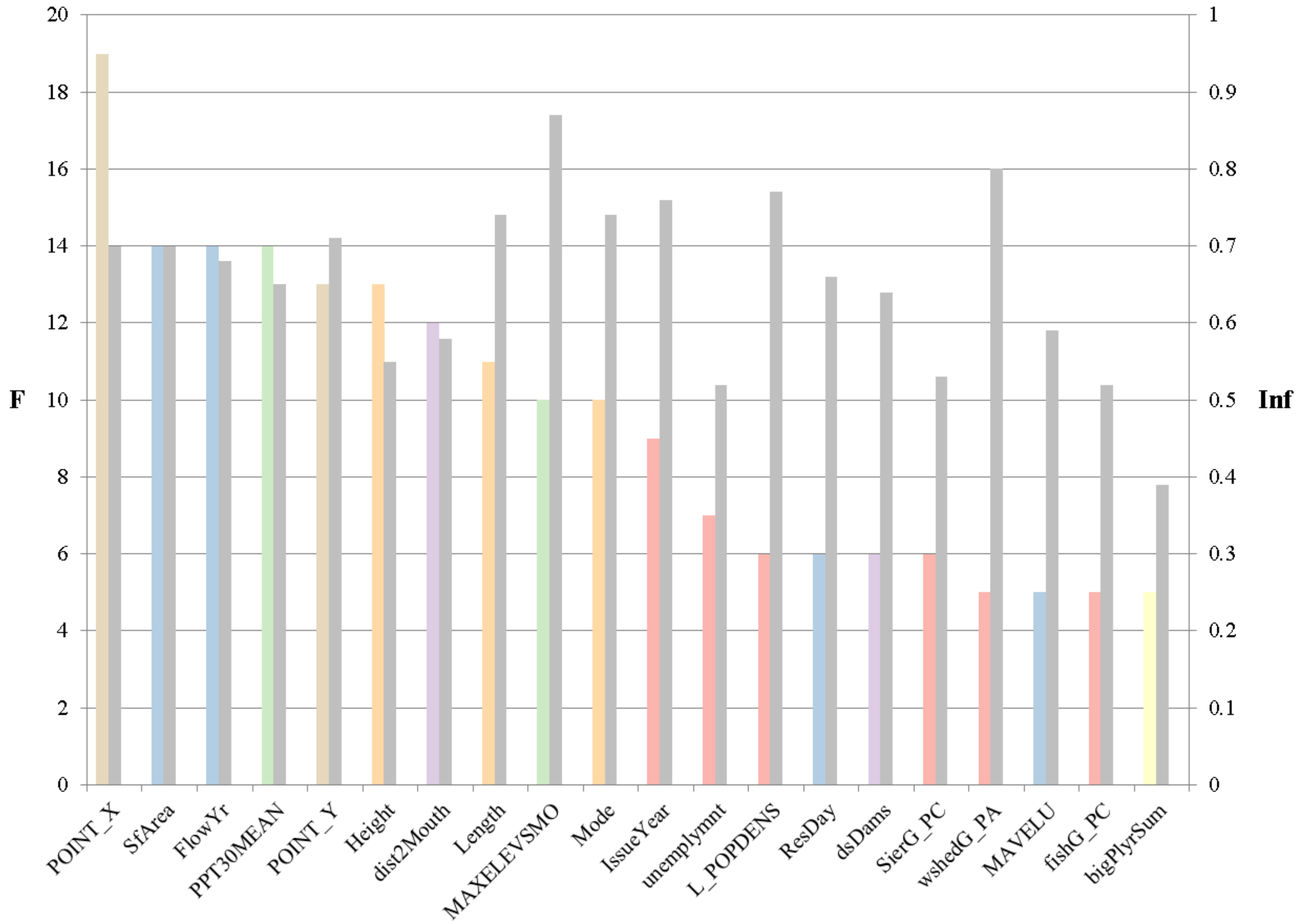


Fig. 4. The 20 most frequently occurring important variables across all Tier 3 models, sorted in descending order from left to right by frequency of occurrence. Colored bars present frequency, while grey bars present the normalized average relative influence for the variable across all of the models in which it was important.

To identify potential key environmental and social drivers of mitigation that may warrant additional future research, we examined important variables across all of our Tier 3 models based on frequency of importance and average relative influence. We grouped important variables into the potential future research areas of socio-political conditions, regional trends, network/landscape position, hydrology/site design, regulatory tendencies, and fisheries (Table 3).

Table 3. The 20 most frequently occurring important variables across all Tier 3 models, with potential future research areas that correspond to each variable. F= frequency; Inf = normalized average relative influence.

Variable	Category	F	Inf	Future research area
POINT_X	Location	19	0.70	Regional trends
SfArea	Hydrology	14	0.70	Hydrology/site design
FlowYr	Hydrology	14	0.68	Hydrology/site design
PPT30MEAN	Landscape	14	0.65	Hydrology/site design
POINT_Y	Location	13	0.71	Regional trends
Height	Facility	13	0.55	Hydrology/site design
dist2Mouth	Stream network	12	0.58	Network/landscape position
Length	Facility	11	0.74	Hydrology/site design
MAXELEVSMO	Landscape	10	0.87	Network/landscape position
Mode	Facility	10	0.74	Hydrology/site design
IssueYear	Human	9	0.76	Regulatory tendencies
unemplymnt	Human	7	0.52	Socio-political conditions
L_POPDENS	Human	6	0.77	Socio-political conditions
ResDay	Hydrology	6	0.66	Hydrology/site design
dsDams	Stream network	6	0.64	Network/landscape position
SierG_PC	Human	6	0.53	Socio-political conditions
wshedG_PA	Human	5	0.80	Socio-political conditions
MAVELU	Hydrology	5	0.59	Hydrology/site design
fishG_PC	Human	5	0.52	Socio-political conditions
bigPlyrSum	Biological	5	0.39	Fisheries

3.3. Predictions to NPDs

We made predictions to 568 NPDs with >1MW potential capacity for each of the 49 acceptable Tier 3 models (Fig. 5). The optimal present/absent threshold for each model is presented in Table 2. The number of predicted mitigation requirements ranged from 9 to 34.

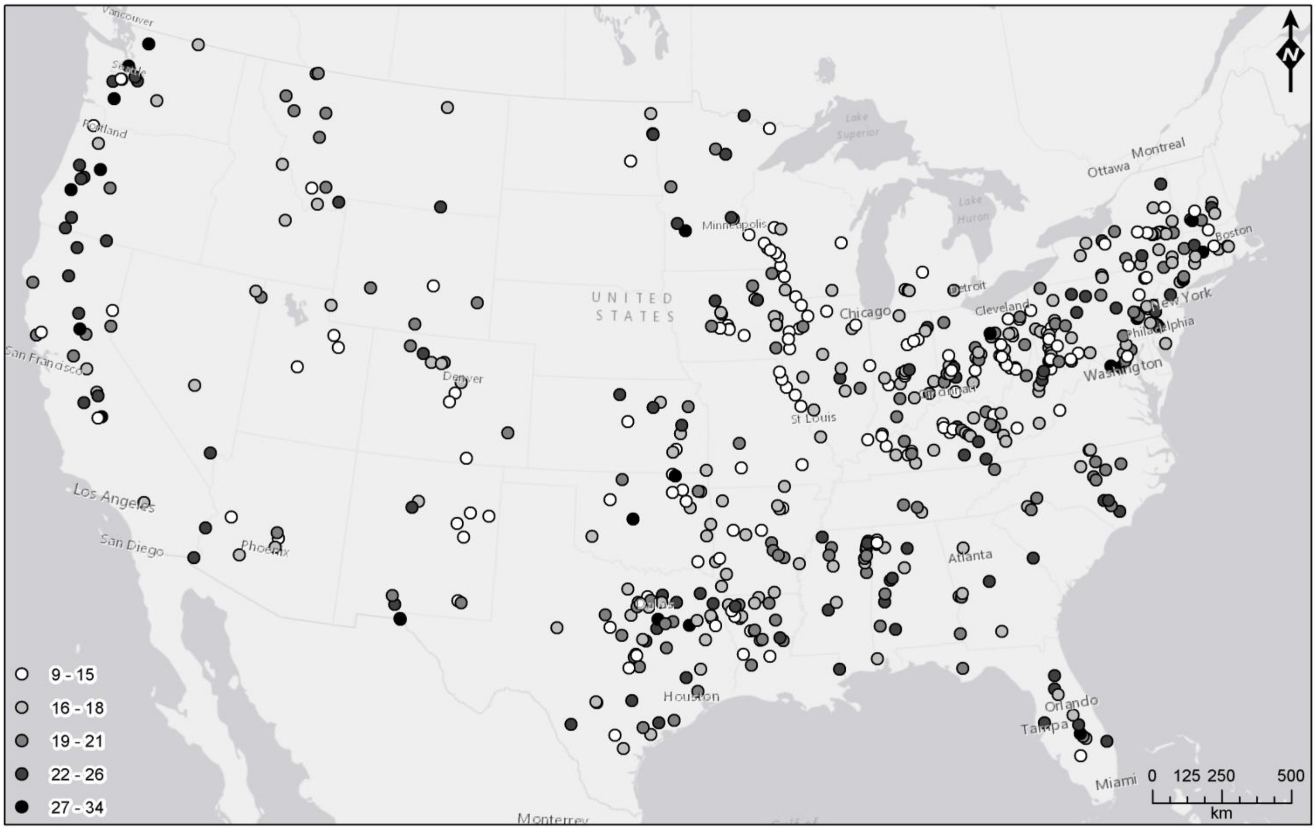


Fig. 5. Number of predicted Tier 3 mitigation requirements at NPDs with >1MW energy potential.

4. DISCUSSION

The spatial modeling approach developed here integrates GIS techniques, novel data, machine-learning algorithms, and niche modeling concepts common in landscape ecology (see Guisan and Thuiller, 2005) to predict environmental mitigation requirements at hydropower project sites. Given the multifaceted, complex nature of demonstrated (Kosnik, 2010) and hypothesized (FERC, 2001) drivers of environmental mitigation requirements, we were uncertain of their predictability. However, we have demonstrated that a broad-scale, multidisciplinary geographical predictor dataset can effectively predict many environmental mitigation requirements across an environmentally and culturally heterogeneous study area.

We summarized and evaluated the influence of the important (relative influence ≥5%) explanatory variables at several different levels of aggregation (Table 2, Fig. 3, and Fig. 4). Since nearly all the categories (e.g. biological, facility, etc.) of variables were influential within each Tier 1 category (Fig. 3) and every Tier 3 model had at least two variable categories represented in the top 3 influential variables (Table 2), it appears that the multi-faceted nature of the predictor dataset we compiled was a key to our modeling success.

Based on our analysis of the top predictor variables across all of the Tier 3 models (Fig. 4), the most common important predictors include metrics of project location (latitude and longitude), project size (annual flow, reservoir size, dam height, and dam length), stream network position (distance along the stream network to network mouth), and climate (precipitation). Elevation above mean sea level, statewide prevalence of local watershed associations, local population density, license issue year, dam length, and dam mode-of-operation were the predictors with the highest average relative influence (Fig. 4) among the most important predictors. Given that the study area is large, environmentally and culturally heterogeneous, and comprised of many diverse physiographic regions, we anticipated latitude and longitude would be valuable predictors that capture regional trends across the U.S. For example, the inclusion of latitude and longitude in models predicting fish passage are related to the fact that most mitigation for passage occurs in the US northeast and northwest. We expected variables related to project size, facility characteristics, and hydrology to be important, given that larger projects are likely to have a higher impact to the environmental and social landscape than smaller projects. Elevation is a proxy for head and a measure of landscape position, and is a very powerful descriptor of landscape context. Stream network position, such as distance to river mouth, can explain presence of diadromous fish species, network connectivity, and existing hydrologic alteration, all of which can heavily influence decisions on mitigation requirements. It is well known in the U.S. that environmental stakeholder groups can be influential in ordered mitigation, so it was not surprising that anthropogenic variables, such as the prevalence of environmental groups, were important. Previous research (Kosnik, 2010) has shown that regulatory trends can influence hydropower mitigation requirements, and the license issue year proved to be an important variable in several models.

Examination of partial dependence plots to assess the direction of variable influence (Appendix B) seems to show that, while there appears to be some consistent direction of influence for important predictors, particularly in the fish passage and water quality models, there are as many examples of contrasting direction of influence within the six broad mitigation categories. This underscores the complexity of the interplay of the nature and magnitude of a given mitigation requirement with the environmental, economic, political, cultural, and social conditions that coalesce at a project and also underscores the need for further investigation into the causality of different drivers of mitigation.

While it is impractical to research causality for all specific mitigation requirements given the sheer number of different types, we identified several potential future research areas (Table 3) that warrant further investigation. One approach to prioritizing future research into mitigation requirement causality would be to delve further into the socio-political and environmental concerns of non-governmental organizations and environmental resource agencies regarding hydropower development, and how those concerns are manifest in prescribed mitigation. These stakeholder groups have a powerful voice and are important to engage early and throughout the project development process if hydropower's contribution to the U.S. renewable energy portfolio is going to be optimized (Fu et al., 2014). A high-level review of The Nature Conservancy's Hydropower by Design strategy (The Nature Conservancy, 2015) and American Rivers Hydropower Reform Coalition platform (Hydropower Reform Coalition, 2016) reveals a common theme of maximizing hydropower sustainability through 1) careful selection of dam location within river networks to optimize both hydropower and conservation objectives, 2) implementing cumulative watershed-scale mitigation strategies, 3) reducing uncertainty and risk associated with project development by directing dam development away from environmentally and socially sensitive areas, and 4) improved outcomes for ecosystem services. Future research into the interplay between socio-political demographics, stream connectivity, ecosystem services, and watershed-scale mitigation approaches and their influence on project siting and ultimate success or failure could serve to catalyze future sustainable hydropower development in the 21st century (Crook et al., 2015; Fu et al., 2014; Karjalainen and Jarvikoski, 2010; Yu et al., 2016).

Another future direction of this research space is the inclusion of cost estimates for different mitigations, which could inform a cost-based approach for identifying priority mitigation types for future investigation of causality. Cost data would also provide a useful constraint for model predictions. Hydropower projects included in the mitigation database (Fig. 1) have a maximum number of 25 mitigation requirements (of the 49 that we modeled), while the model predictions to NPDs included as many as 34 mitigation requirements. Incorporating cost data would allow for additional realism to be integrated into the predictions by sequentially predicting mitigation types from most to least costly with a control on cost.

Our results should be interpreted with caution given that several models showed significant spatial autocorrelation. Since one of our goals was spatial prediction beyond the spatial extent of our dataset, we did not implement methods for accounting for spatial autocorrelation because previously developed methods do not allow for prediction beyond the dataset (Dormann et al., 2007; Rickbeil et al., 2014). We recognize that our models did not use an independent validation dataset, but rather a split of our original dataset. Since there is no comparable dataset available, we argue that our data split combined with tenfold internal cross-validation allowed for reliable evaluation of model performance to be made (Rickbeil et al., 2014).

The BRT models could potentially be improved by improving some of the more coarse resolution predictors – such as those derived at the state-scale – to represent a more refined local scale. A disconnect may exist between the spatial scale at which mitigation requirements are ordered and the scale at which some of our explanatory variables are derived. This disparity of scales and varying resolution of predictors can affect the apparent importance of a predictor variable (Brewer et al., 2007). Schramm et al. (2016)described several possible limitations to the development of the mitigation database, which was limited to a review of mitigation prescribed explicitly in FERC licenses issued from 1998 to 2015. More specifically, some of the reviewed licenses were for relicensing of existing projects and thus may not include previously required mitigation under the original license. Also, FERC encourages the use of settlement agreements (legal agreements developed between hydropower developers, agencies, and other stakeholders on project operations and environmental conditions) that may include mitigation not included in the final license.

5. CONCLUSION

We demonstrated in this study an approach including specific statistical models that can be used by developers and regulators alike to identify and anticipate likely environmental mitigation at existing and proposed hydropower projects in the U.S. The results demonstrate that mitigation requirements in existing licenses have been a result of a range of factors from biological and hydrological to political and cultural. That such a range of variable types is needed to predict mitigation requirements explains much of the difficulty and uncertainty that surrounds the development of effective environmental mitigation during the licensing process in the U.S. Further research is needed to establish robust links between specific explanatory variables, mitigation requirements, and mitigation strategies. However, use of these models by developers can reduce uncertainty with regards to cost projections and inform decisions about project design. Regulators will be able to use the models to more quickly identify likely environmental issues and potential solutions, hopefully resulting in more timely and more effective decisions on environmental mitigation.

6. ACKNOWLEDGEMENTS

We would like to thank R. McManamay for statistical advice, suggestions for predictor variables, and valuable comments on this manuscript. We also thank S.C. Kao for assistance with the NHAAP database and FERC licenses. This study was funded by the U.S. Department of Energy (DOE) Energy Efficiency and Renewable Energy Office, Wind and Water Power Technologies Program through Oak Ridge

National Laboratory, which is managed by UT-Battelle, LLC, for the DOE under contract DE-AC05-00OR22725.

7. REFERENCES

Abram NK, Meijaard E, Wells JA, Ancrenaz M, Pellier A-S, Runting RK, et al. Mapping perceptions of species' threats and population trends to inform conservation efforts: the Bornean orangutan case study. Diversity and Distributions 2015; 21: 487-499.

Araujo M, Guisan A. Five (or so) challenges for species distribution modelling. Journal of Biogeography 2006; 33: 1677-1688.

Arganaraz J, Pizarro G, Zak M, Landi M, Bellis L. Human and biophysical drivers of fires in Semiarid Chaco mountains of Central Argentina. Science of the Total Environment 2015; 520: 1-12.

Blumm MC, Nadol VA. The decline of the hydropower czar and the rise of agency pluralism in hydroelectric licensing. Columbia Journal of Environmental Law 2001; 26.

Brewer SK, Rabeni CF, Sowa SP, Annis G. Natural landscape and stream segment attributes influencing the distribution and relative abundance of riverine smallmouth bass in Missouri. North American Journal of Fisheries Management 2007; 27: 326-341.

Bunn S, Arthington A. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environmental Management 2002; 30: 492-507.

Cada GF. Fish passage mitigation at hydroelectric power projects in the United States. Fish Migration and Fish Bypasses 1998: 208-219.

Crook D, Lowe W, Allendorf F, Eros T, Finn D, Gillanders B, et al. Human effects on ecological connectivity in aquatic ecosystems: Integrating scientific approaches to support management and mitigation. Science of the Total Environment 2015; 534: 52-64.

Deshazo JR, Freeman J. Public Agencies as Lobbyists. Columbia Law Review 2005; 105.

Dormann C, Elith J, Bacher S, Buchmann C, Carl G, Carre G, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 2013; 36: 27-46.

Dormann C, McPherson J, Araujo M, Bivand R, Bolliger J, Carl G, et al. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 2007; 30: 609-628.

Elith J, Graham C, Anderson R, Dudik M, Ferrier S, Guisan A, et al. Novel methods improve prediction of species' distributions from occurrence data. Ecography 2006; 29: 129-151.

Elith J, Leathwick J, Hastie T. A working guide to boosted regression trees. Journal of Animal Ecology 2008; 77: 802-813.

FERC. Report on hydroelectric licensing policies, procedures, and regulations, http://www.ferc.gov/legal/maj-ord-reg/land-docs/ortc_final.pdf (accessed 1.15.15), 2001.

FERC. workload, http://www.ferc.gov/industries/hydropower/gen-info/licensing/workload.xls (accessed 1.15.15), 2015.

Fink D, Hochachka W, Zuckerberg B, Winkler D, Shaby B, Munson M, et al. Spatiotemporal exploratory models for broad-scale survey data. Ecological Applications 2010; 20: 2131-2147.

Fraley J, Marotz B, Decker-Hess J, Beattie W, Zubik R. Mitigation, compensation, and future protection for fish populations affected by hydropower development in the upper Columbia system, Montana, U.S.A. River Research and Applications 1989; 3: 3-18.

Fu B, Wang Y, Xu P, Yan K, Li M. Value of ecosystem hydropower service and its impact on the payment for ecosystem services. Science of the Total Environment 2014; 472: 338-346.

Grimm N, Foster D, Groffman P, Grove J, Hopkinson C, Nadelhoffer K, et al. The changing landscape: ecosystem responses to urbanization and pollution across climatic and societal gradients. Frontiers in Ecology and the Environment 2008; 6: 264-272.

Guisan A, Thuiller W. Predicting species distribution: offering more than simple habitat models. Ecology Letters 2005; 8: 993-1009.

Hadjerioua B, Wei Y, Kao S-C. An Assessment of Energy Potential at Non-Powered Dams in the United States. ORNL/TM-2011/341. Oak Ridge National Laboratory, 2012.

Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression. Hoboken, NJ, U.S.A.: John Wiley & Sons, 2013.

Hydropower Reform Coalition. Platform, http://hydroreform.org/sites/default/files/HRC%20Platform_2.pdf (accessed 2.15.16), 2016.

Karjalainen TP, Jarvikoski TB. Negotiating river ecosystems: Impact assessment and conflict mediation in the cases of hydro-power construction. Environmental Impact Assessment Review 2010; 30: 319-327.

Knittel CR. The Adoption of State Electricity Regulation: The Role of Interest Groups. Journal of Industrial Economics 2006; 54: 201-222.

Kosnik L. Balancing Environmental Protection and Energy Production in the Federal Hydropower Licensing Process Land Economics 2010; 86: 444-466.

Kuby MJ, Fagan WF, ReVelle CS, Graf WL. A multiobjective optimization model for dam removal: an example trading off salmon passage with hydropower and water storage in the Willamette basin. Advances in Water Resources 2005; 28: 845-855.

Kuhn M. Building Predictive Models in R Using the caret Package. Journal of Statistical Software 2008; 28: 1-26.

Liermann C, Nilsson C, Robertson J, Ng R. Implications of Dam Obstruction for Global Freshwater Fish Diversity. Bioscience 2012; 62: 539-548.

Ligon FK, Dietrich WE, Trush WJ. Downstream ecological effects of dams. Bioscience 1995; 45: 183-192.

Manzano-Agugliaro F, Alcayde A, Montoya F, Zapata-Sierra A, Gil C. Scientific production of renewable energies worldwide: An overview. Renewable & Sustainable Energy Reviews 2013; 18: 134-143.

McManamay R, Samu N, Kao S, Bevelhimer M, Hetrick S. A Multi-scale Spatial Approach to Address Environmental Effects of Small Hydropower Development. Environmental Management 2015; 55: 217-243.

Moore ID, Grayson RB, Ladson AR. Digital Terrain Modeling - A review of hydrological, geomorphological, and biological applications. Hydrological Processes 1991; 5: 3-30.

NatureServe. Digital Distribution Maps of the Freshwater Fishes in the Conterminous United States, Arlington, VA. U.S.A., 2010.

Nilsson C, Reidy C, Dynesius M, Revenga C. Fragmentation and flow regulation of the world's large river systems. Science 2005; 308: 405-408.

Oppel S, Meirinho A, Ramirez I, Gardner B, O'Connell A, Miller P, et al. Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. Biological Conservation 2012; 156: 94-104.

Parisien M, Parks S, Krawchuk M, Flannigan M, Bowman L, Moritz M. Scale-dependent controls on the area burned in the boreal forest of Canada, 1980-2005. Ecological Applications 2011; 21: 789-805.

Peltzman S. Toward a more general theory of regulation. Journal of Law and Economics 1976; 19: 211-240.

Poff N, Allan J, Bain M, Karr J, Prestegaard K, Richter B, et al. The natural flow regime. Bioscience 1997; 47: 769-784.

Poff N, Olden J, Merritt D, Pepin D. Homogenization of regional river dynamics by dams and global biodiversity implications. Proceedings of the National Academy of Sciences of the United States of America 2007; 104: 5732-5737.

R-Core-Team. R: a language and environment for statistical computing, Vienna, Austria. Version 3.2.2. Available at https://www.r-project.org/. R Foundation for Statistical Computing, 2015.

Renofalt B, Jansson R, Nilsson C. Effects of hydropower generation and opportunities for environmental flow management in Swedish riverine ecosystems. Freshwater Biology 2010; 55: 49-67.

Rickbeil G, Coops N, Drever M, Nelson T. Assessing coastal species distribution models through the integration of terrestrial, oceanic and atmospheric data. Journal of Biogeography 2014; 41: 1614-1625.

Schramm MP, Bevelhimer MS, DeRolph CR. A synthesis of environmental and recreational mitigation requirements at hydropower projects in the United States. Environmental Science & Policy 2016; 61: 87-96.

Shmueli G. To Explain or to Predict? Statistical Science 2010; 25: 289-310.

Tarlock D. Hydro law and the future of hydroelectric power generation in the United States. Vanderbilt Law Review 2012; 65: 1723-1767.

The Nature Conservancy. The Power of Rivers, http://www.nature.org/media/freshwater/power-of-rivers-report.pdf (accessed 2.15.16), 2015.

Tong S, Chen W. Modeling the relationship between land use and surface water quality. Journal of Environmental Management 2002; 66: 377-393.

Trussart S, Messier D, Roquet V, Aki S. Hydropower projects: a review of most effective mitigation measures. Energy Policy 2002; 30: 1251-1259.

U.S.-Congress. House Commitee on Natural Resources. Mandatory conditioning requirements on hydropower: How federal resources agencies are driving up electricity costs and decreasing the original green energy, 2012.

VanDerWal J, Falconi L, Januchowski S, Shoo L, Storlie C. Package 'SDMTools', http://cran.r-project.org/web/packages/SDMTools/index.html (accessed August 2015), 2012.

Wang L, Lyons J, Kanehl P. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 2001; 28: 255-266.

Yu B, Xu L. Review of ecological compensation in hydropower development. Renewable & Sustainable Energy Reviews 2016; 55: 729-738.

Yu B, Xu L, Yang Z. Ecological compensation for inundated habitats in hydropower developments based on carbon stock balance Journal of Cleaner Production 2016; 114: 334-342.

Appendix A

Mitigation categories in the hierarchical database and the percent of times each was required. Bold indicates model was fit for mitigation category.

Tier 1	Tier 2	Tier 3	ModelID	Percent required
Fish Passage	--	--	F1	48.8%
	Downstream Fish Passage	--	**F101**	**24.9%**
		Surface Collector	F101001	1.9%
		Trap and Transport	F101002	1.4%
		Modification of Spill or Gate Operation	F101003	3.3%
		Sluiceway	F101004	0.4%
		Bypass Facility	F101005	2.1%
		Conduit	F101006	1.2%
		Fish Friendly Turbine	F101007	0.2%
		Generation Shut Down	F101008	2.1%
		Flashboard Removal or Modification	F101009	0.2%
		Downstream Passage Plan Study Design	**F101010**	**15.5%**
		Modify spill or gate design	F101011	2.7%
		Modify sluiceway	F101012	1.6%
		Modify bypass facility	F101013	0.2%
		Modify intake	F101014	0.2%
	Upstream Fish Passage	--	**F102**	**19.4%**
		Adult fishway	F102015	0.4%
		Conduit	F102016	0.2%
		Eelway	**F102017**	**7.4%**
		Fish Ladder	F102018	3.3%
		Lock or elevator	F102019	2.3%
		Modify spill or gate operation	F102020	0.2%
		Tailrace exclusion device	F102021	2.5%
		Trap and transport	F102022	4.3%
		Upstream passage study plan or design	**F102023**	**12.2%**
		Modify adult fishway	F102024	0.2%
		Modify fish ladder	F102025	0.8%
		Modify lock or lift	F102026	0.2%
		Modify trap and transport	F102027	0.6%
	Passage Planning	--	**F103**	**26.2%**
		Design plan entrainment avoidance system	F103028	1.6%
		Downstream fish passage monitoring sampling	**F103029**	**13.6%**
		Entrainment or turbine mortality monitoring	F103030	3.7%
		Fish passage and operations plan	**F103031**	**7.4%**
		Fish passage feasibility assessment	F103032	3.1%
		Fish stranding plan monitoring evaluation	**F103033**	**7.2%**
		Fisheries disease management	F103034	0.6%
		Hatchery operations and management	F103035	1.9%
		Upstream fish passage monitoring sampling	**F103036**	**10.7%**
	Entrainment	--	**F104**	**23.9%**
		Barrier or guidance net	F104037	1.6%
		Fish screen	F104038	4.3%
		Gatewell exclusion screen	F104039	0.4%
		Perforated plate	F104040	0.2%
		Solid panel and bar rack	F104041	0.4%
		Strobe light	F104042	0.2%
		Trash or bar rack	**F104043**	**17.1%**
Hydrology	--	--	F2	95.1%
	Flow Mitigation	--	**F205**	**61.9%**
		Tailrace adaptive flow management	F205044	1.6%
		Tailrace flow monitoring plan	**F205045**	**34.0%**
		Tailrace flow studies	F205046	3.5%
		Tailrace flushing or flood flows	F205047	1.9%
		Tailrace flow or stage monitoring equipment	**F205048**	**14.2%**
		Tailrace flow control device	F205049	2.7%
		Tailrace ramping rate restriction	**F205050**	**11.1%**
		Bypass adaptive flow management	F205051	1.9%
		Bypass flow monitoring plan	**F205052**	**12.8%**
		Bypass flow study	F205053	2.5%

Appendix A
Mitigation categories in the hierarchical database and the percent of times each was required. Bold indicates model was fit for mitigation category.

Tier 1	Tier 2	Tier 3	ModelID	Percent required
		Bypass flushing or flood flow	**F205054**	**5.6%**
		Bypass flow or stage monitoring equipment	**F205055**	**7.2%**
		Bypass flow control device	F205056	0.8%
		Bypass ramping rate restriction	**F205057**	**6.2%**
	Tailrace Minimum Flow	--	**F206**	**64.5%**
		Run-of-river Tailrace	**F206058**	**39.0%**
		Seasonal Tailrace	**F206059**	**13.6%**
		Seasonal and type of year Tailrace	F206060	1.6%
		Year-round Tailrace	**F206061**	**10.3%**
	Bypass Minimum Flow	--	**F207**	**41.9%**
		Seasonal Bypass	**F207063**	**17.1%**
		Seasonal and type of year Bypass	F207064	4.5%
		Year-round Bypass	**F207065**	**20.2%**
	Sediment	--	**F208**	**42.9%**
		Sediment and erosion control plan or monitoring	**F208066**	**41.6%**
		Dredging	F208067	0.2%
		Install or operate gate to flush sediment	F208068	0.8%
		Sediment flushing flows	F208069	0.8%
	Recreation Flow	--	**F209**	**13.2%**
		Maintain recreational lake levels	F209070	3.3%
		Provide recreational flow releases or structures	**F209071**	**9.7%**
		Recreational flow studies	F209072	4.1%
	Operations	--	**F210**	**54.8%**
		Flow management plan	**F210073**	**6.6%**
		Operations compliance monitoring plan	**F210074**	**40.6%**
		Provide flow or lake levels electronically	**F210075**	**10.7%**
Water Quality	--	--	F3	53.7%
	Downstream Water Quality	--	**F311**	**54.0%**
		Adaptive water quality management	F311076	3.7%
		Benthic macroinvertebrate monitoring	**F311077**	**5.4%**
		DO enhancement or mitigation plan	**F311078**	**5.4%**
		Establish or fund water quality stations and stream gages	F311079	3.3%
		Forebay aeration	F311080	0.2%
		Operational changes	F311081	2.7%
		Powerhouse aeration	F311082	2.1%
		Tailrace structures for aeration	F311083	0.2%
		Temperature regulating device or structure	F311084	0.6%
		Temperature regulation or mitigation plan	F311085	0.4%
		Water quality monitoring plan	**F311086**	**50.3%**
	Upstream Water Quality	--	**F312**	**24.5%**
		Fish tissue sampling and analysis	**F312087**	**8.2%**
		Impoundment sediment analysis	**F312088**	**6.4%**
		Macroinvertebrate monitoring	F312089	0.6%
		Inflow water quality monitoring plan	**F312090**	**8.9%**
		Impoundment water quaity monitoring plan	**F312091**	**17.3%**
Biodiversity	--	--	F4	71.4%
	Terrestrial	--	**F413**	**66.6%**
		Acquistion easements conservation or important habitat	F413092	4.1%
		Install upgrade monitor wildlife crossings	F413093	4.1%
		Noxious terrestrial weed and invasive plant management	**F413094**	**25.6%**
		Species conservation management monitoring	**F413095**	**42.9%**
		Threatened and endangered species protection plan	**F413096**	**10.9%**
		Transmission related avaian and bat protection	**F413097**	**15.5%**
		Wildlife terrestrial habitat management	**F413098**	**27.0%**
	Aquatic	--	**F414**	**35.5%**
		Adaptive fishery management	F414099	3.9%
		Aquatic species conservation management monitoring	**F414100**	**25.6%**
		Diadromous species management monitoring	**F414101**	**7.4%**
		Invasive aquatic species management (fish and molluscs)	**F414102**	**9.3%**
		Stocking fish species of concern	F414103	4.5%

Appendix A

Mitigation categories in the hierarchical database and the percent of times each was required. Bold indicates model was fit for mitigation category.

Tier 1	Tier 2	Tier 3	ModelID	Percent required
Habitat	--	--	F5	57.1%
	Fisheries	--	**F515**	**26.8%**
		Downstream gravel and sediment restoration	F515104	4.5%
		Downstream habitat enhancement	**F515105**	**8.7%**
		Downstream woody debris restoration or passage	**F515106**	**15.9%**
		Reservoir fishery habitat enhancement	F515133	3.3%
	Riparian	--	**F516**	**20.4%**
		Dust control and abatement	F516107	0.6%
		Establish riparian buffers	**F516108**	**7.2%**
		Riparian habitat enhancement	F516109	4.7%
		Riparian habitat monitoring or planning	**F516110**	**12.0%**
	Reservoir	--	**F517**	**31.8%**
		Noxious invasive aquatic plant management	**F517111**	**21.0%**
		Shoreline management plan or program	**F517112**	**16.1%**
	Wetlands	--	**F518**	**11.5%**
		Wetland enhancement	F518113	0.8%
		Wetland mitigation	F518114	4.9%
		Wetland monitoring	F518115	2.9%
		Wetland protection	**F518116**	**6.8%**
Recreation	--	--	F6	82.3%
	Resources and Mitigation	--	**F619**	**66.2%**
		Appoint historic cultural resource coordinator	F619117	0.4%
		Boating facilities	**F619118**	**23.1%**
		Canoe portage launch	**F619119**	**24.1%**
		Fishing pier	**F619120**	**8.9%**
		Floating debris removal	F619121	1.0%
		Install fish attracting structure for recreational fishing	F619122	2.7%
		Interpretive education sign and displays	**F619123**	**15.5%**
		Navigational aids and improvements	F619124	1.0%
		Parking	**F619125**	**26.4%**
		Protection of specific historic cultural resource sites	F619126	3.3%
		Public outreach education programs	F619127	0.6%
		Shoreline access	**F619128**	**17.9%**
		Stocking recreational fish species	**F619129**	**6.6%**
		Trail trailhead or camping areas	**F619130**	**14.2%**
		Other day use area improvements	**F619132**	**35.3%**
	Planning	--	**F620**	**72.8%**
		Recreational management plan study or monitoring	**F620131**	**72.8%**

Appendix B

Partial dependence plots for the three variables with the highest relative influence for each Tier 2 and Tier 3 model with an internal CV ROC and independent ROC ≥0.7

NOTES: The ModelID for each model is shown in the upper-right hand corner of each set of three partial dependence plots; see Table 1 for variable descriptions; see Table 2 for details on mitigation types; ticks across the top of each plot show the distribution of deciles for each predictor variable.

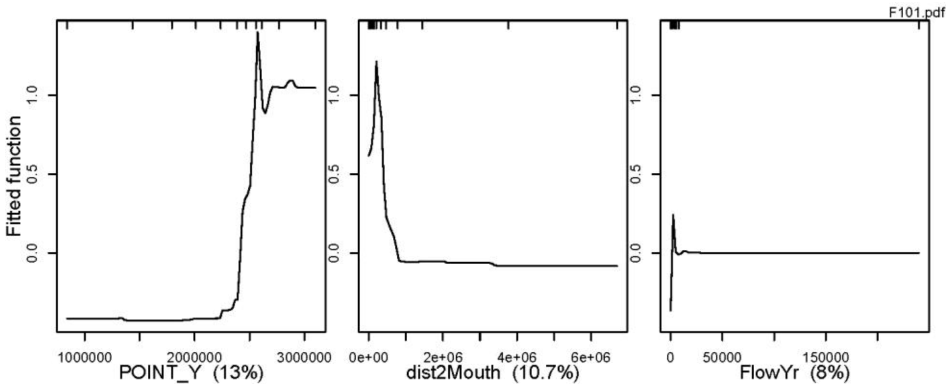
Definitions for categorical variables

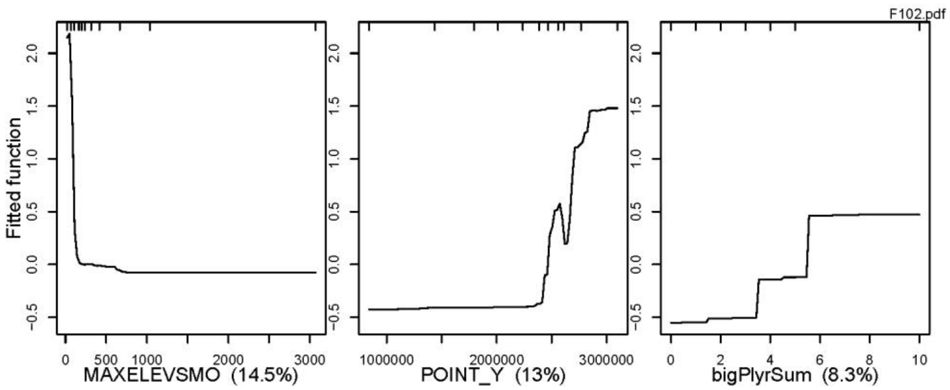
Mode	Value
Canal/Conduit	C
Intermediate Peaking	I
Peaking	P
Pumped Storage	S
Reregulating	R
Run-of-river	O
Run-of-river/Peaking	A
Run-of-river/Upstream Peaking	B

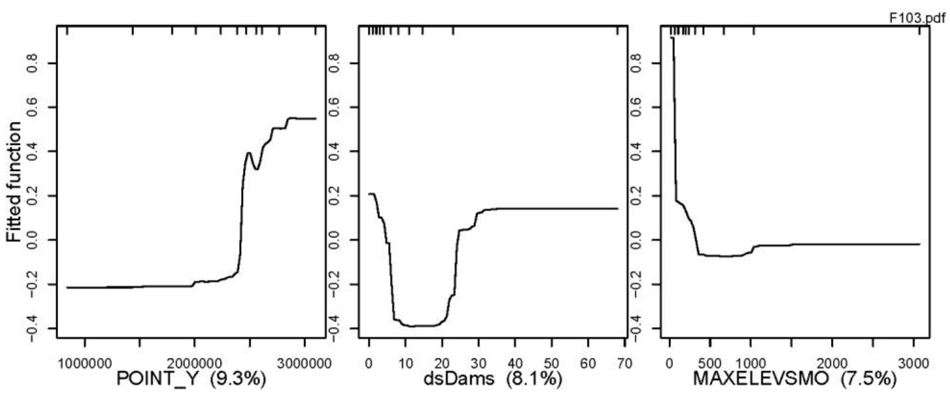
owner	Value
Cooperative	C
Private	P
Public	U
Wholesale Power Marketer	W

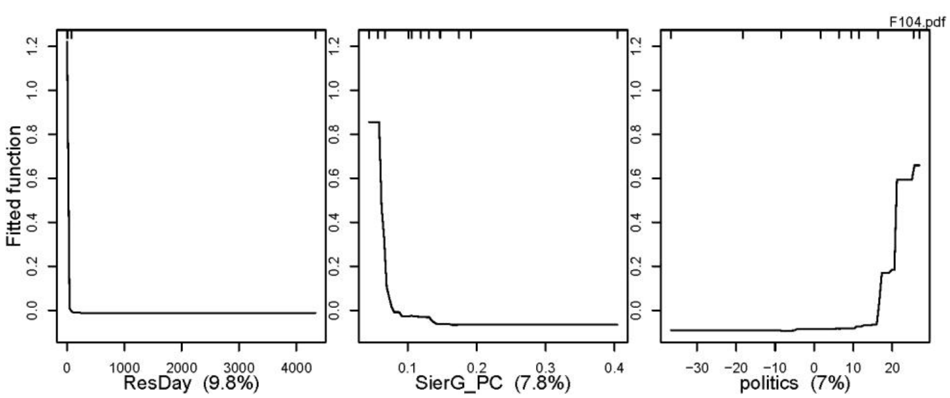
PrmryPurps	Value
Flood control and stormwater management	C
Fish and wildlife pond	F
Hydroelectric	H
Irrigation	I
Navigation	N
Other	O
Recreation	R
Water supply	S

Tier 2 Fish Passage

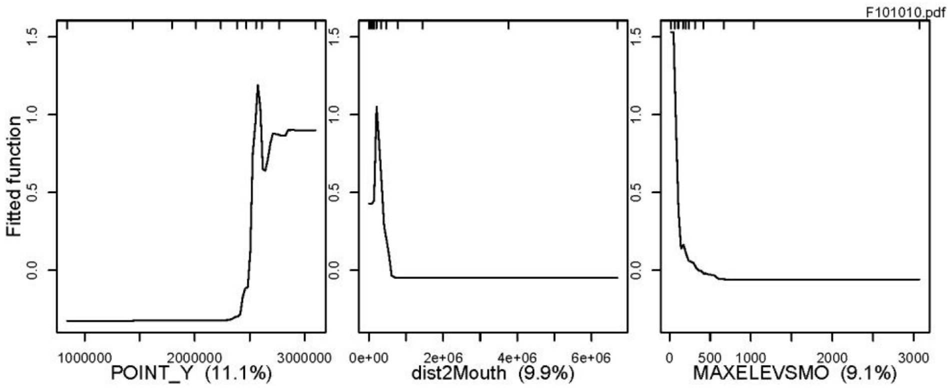


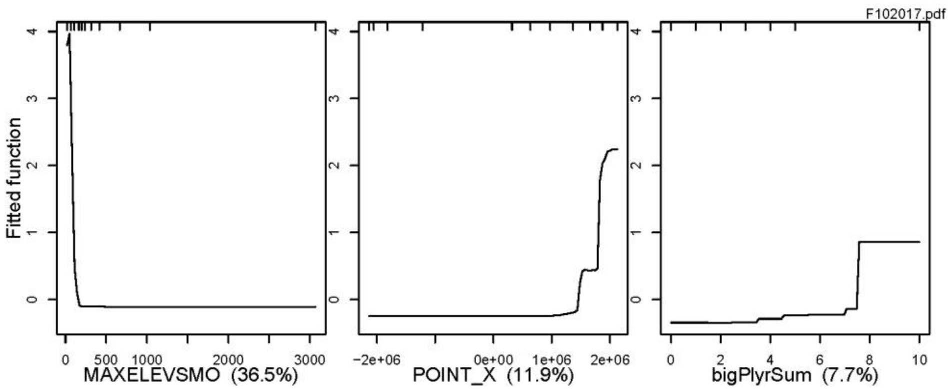


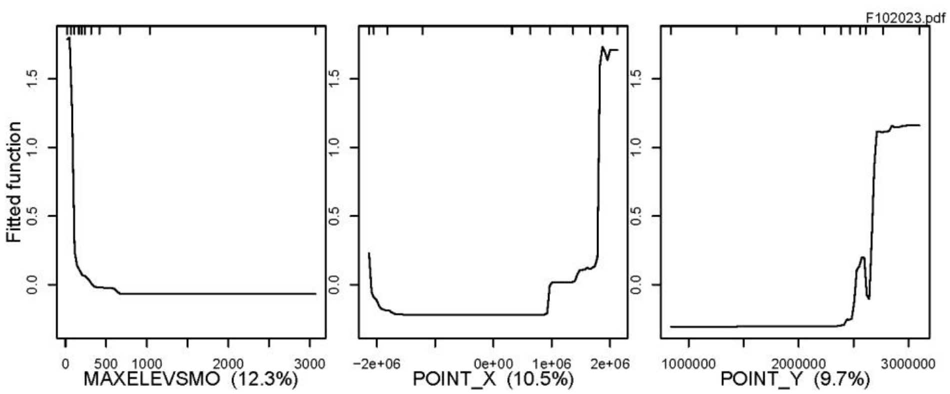


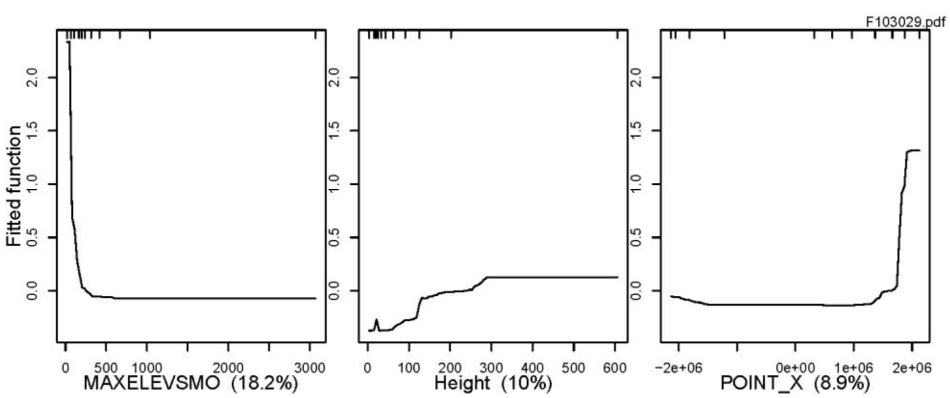


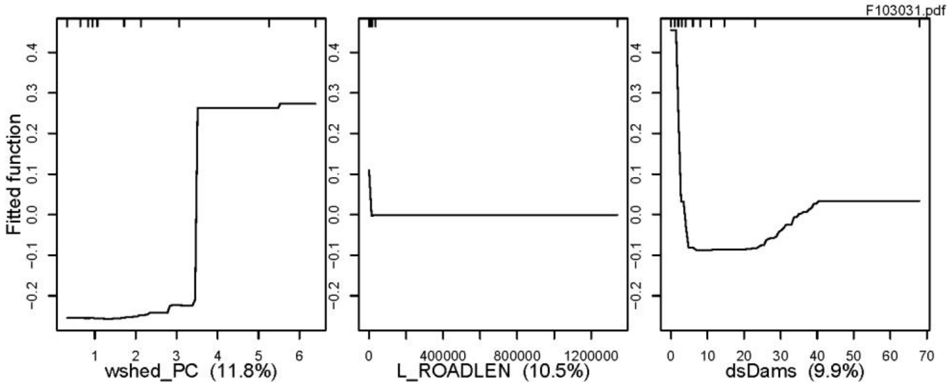
Tier 3 Fish Passage

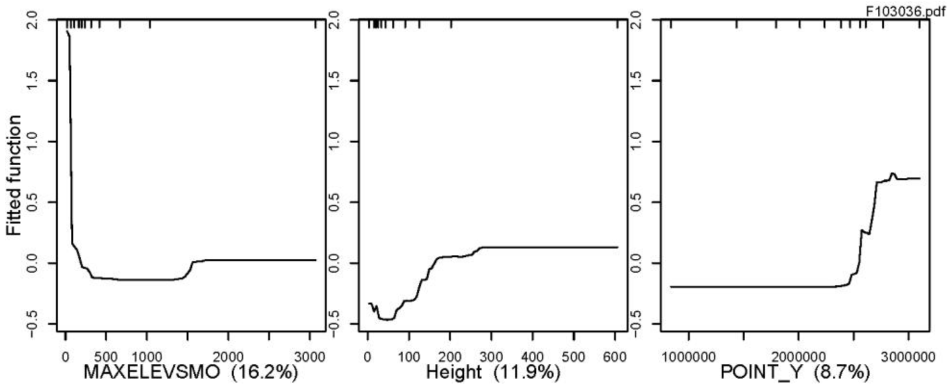


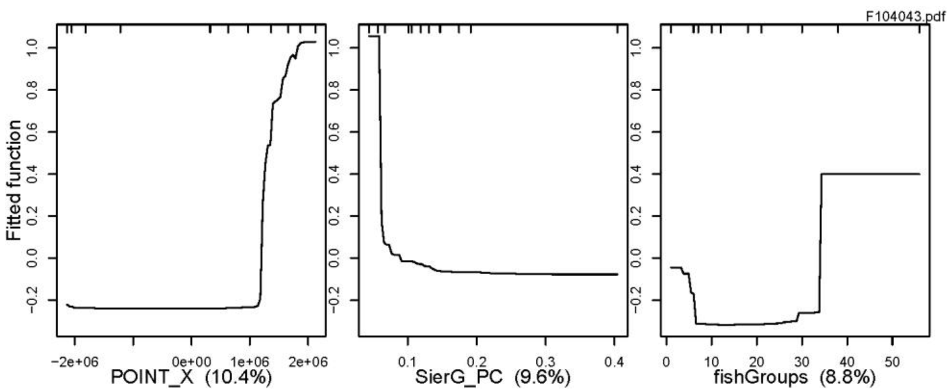




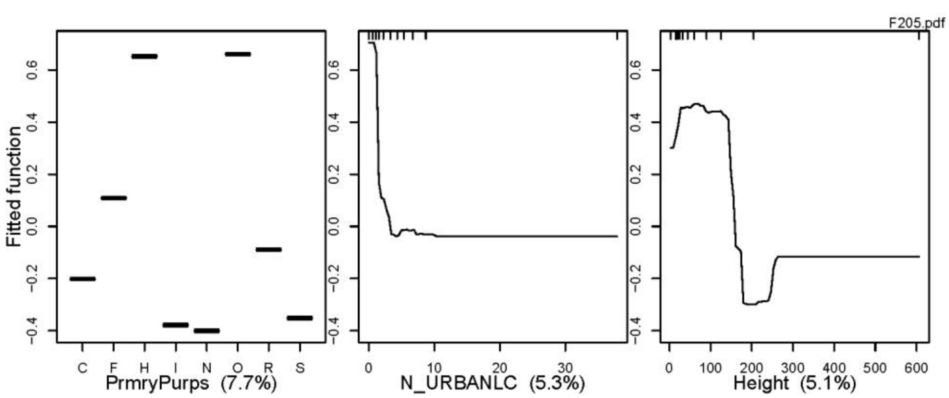


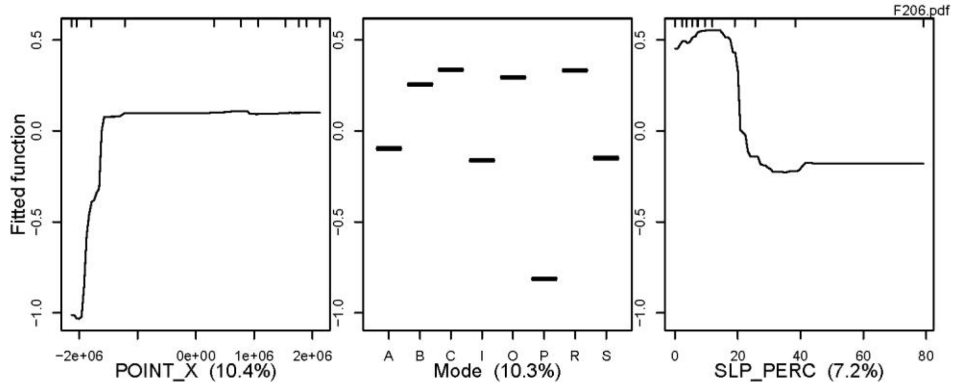


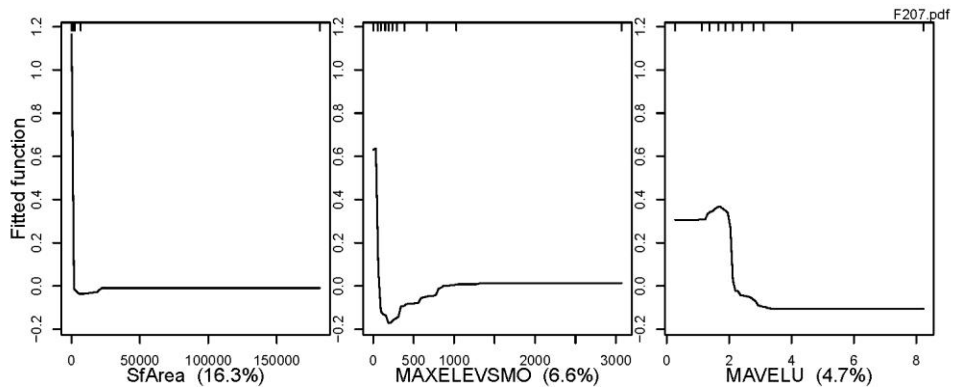


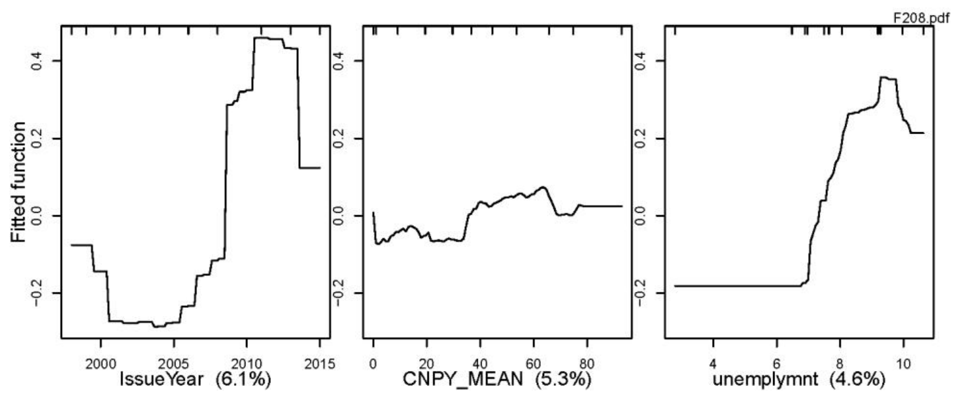


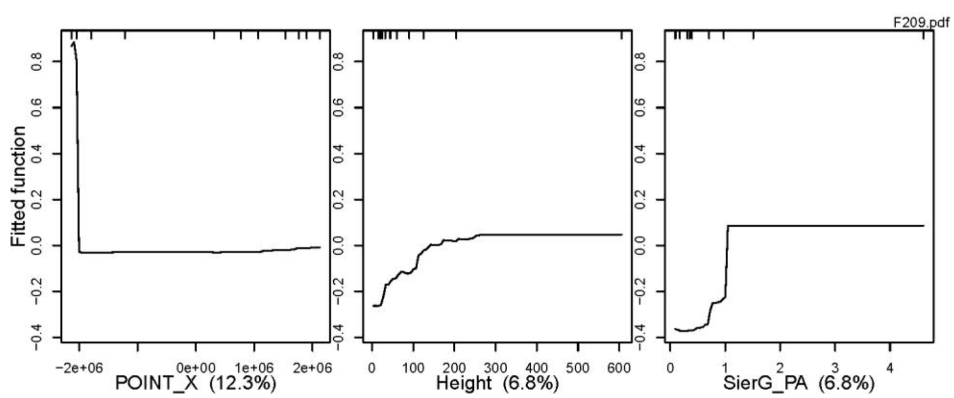
Tier 2 Hydrology

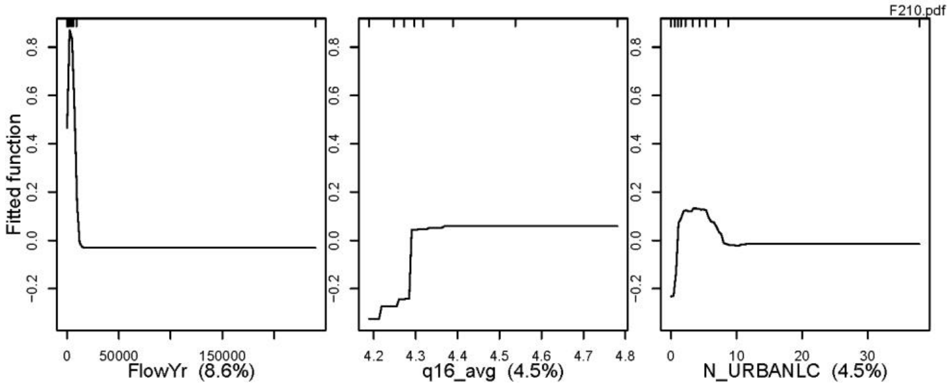



F206.pdf
Fitted function
POINT_X (10.4%)
Mode (10.3%)
SLP_PERC (7.2%)

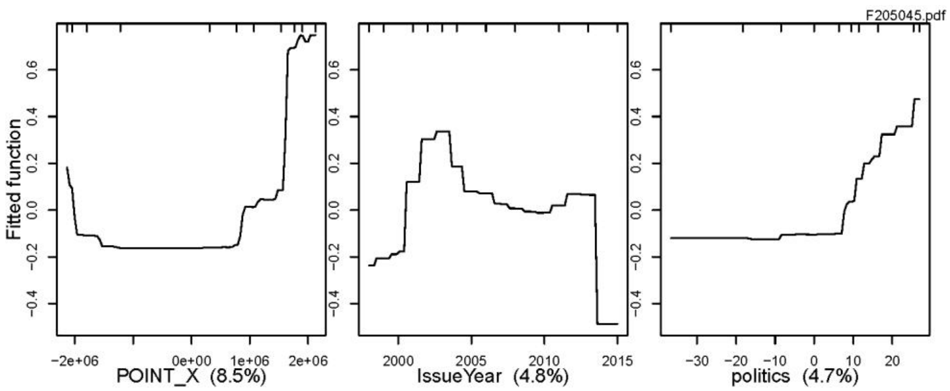

F207.pdf
Fitted function
SfArea (16.3%)
MAXELEVSMO (6.6%)
MAVELU (4.7%)

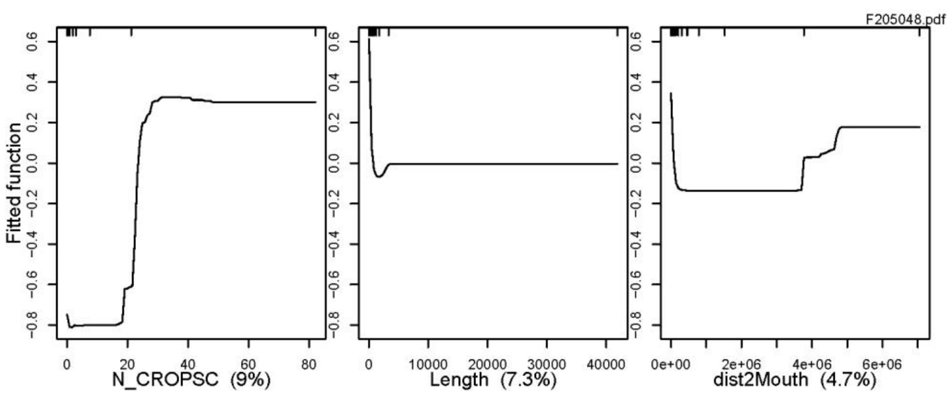

F208.pdf
Fitted function
IssueYear (6.1%)
CNPY_MEAN (5.3%)
unemplymnt (4.6%)

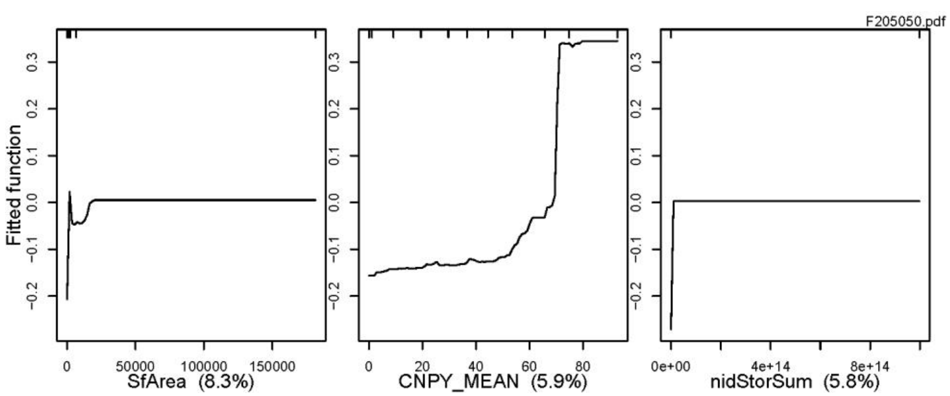

F209.pdf
Fitted function
POINT_X (12.3%)
Height (6.8%)
SierG_PA (6.8%)

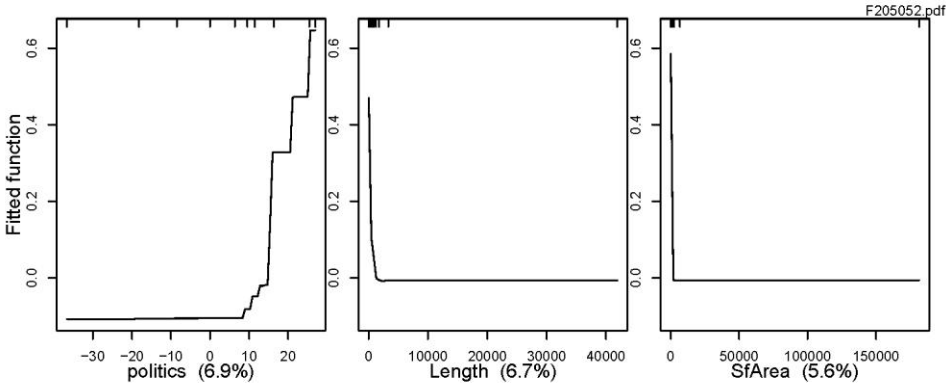


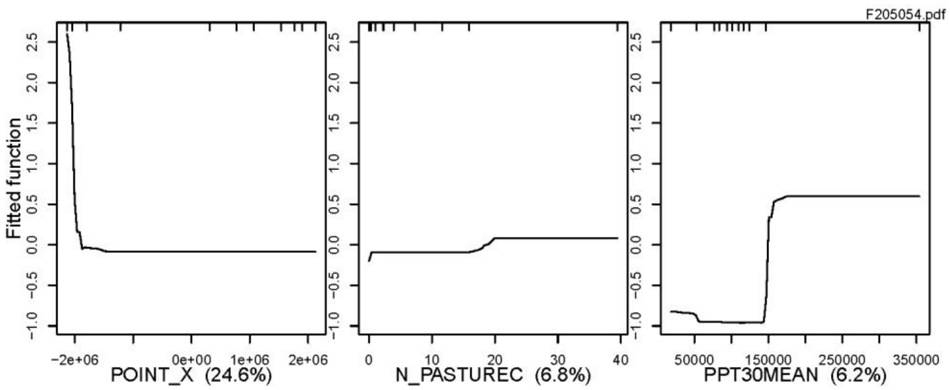
Tier 3 Hydrology

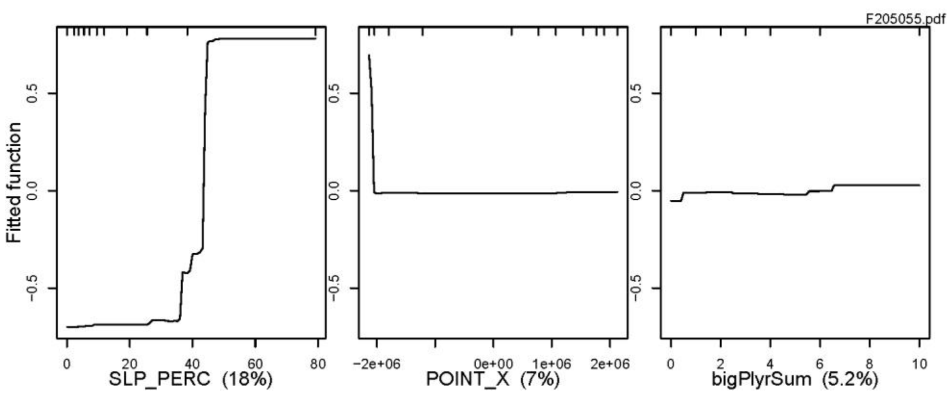


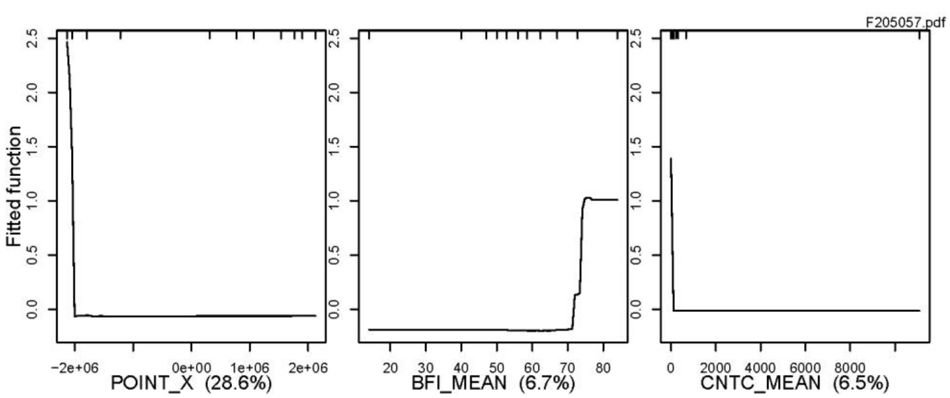


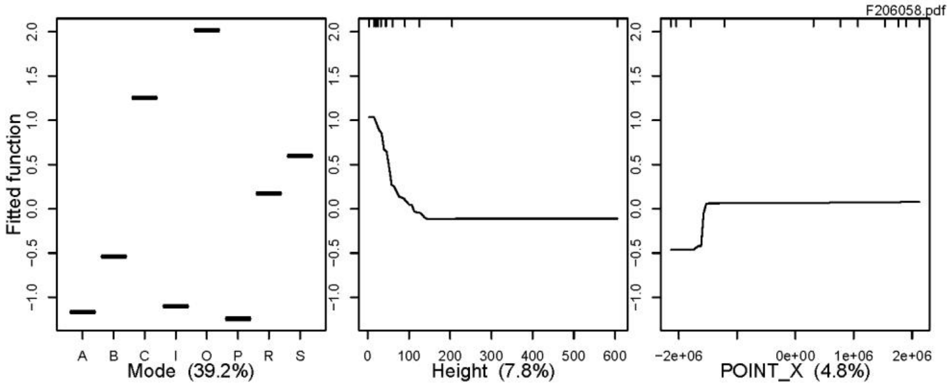


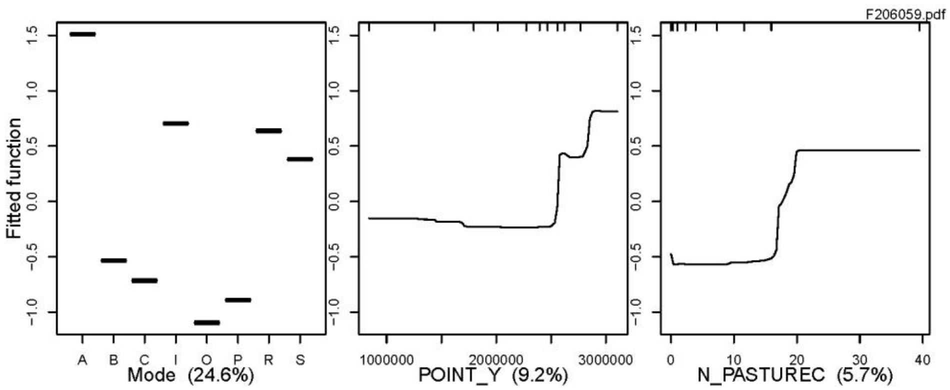

F205052.pdf
Fitted function
politics (6.9%)
Length (6.7%)
SfArea (5.6%)

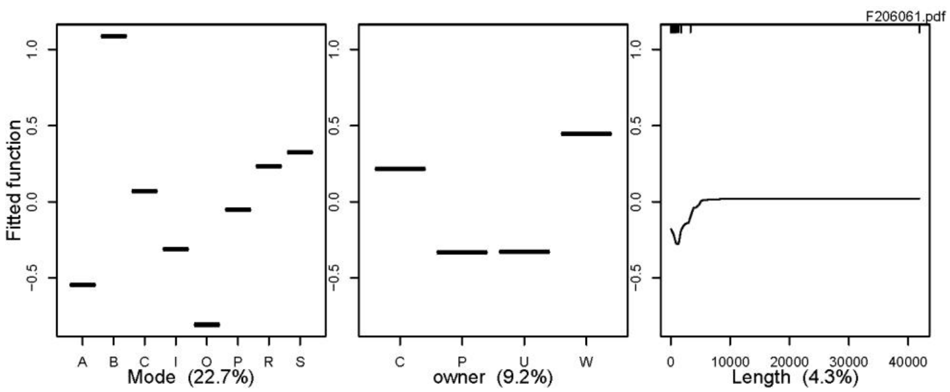

F205054.pdf
Fitted function
POINT_X (24.6%)
N_PASTUREC (6.8%)
PPT30MEAN (6.2%)

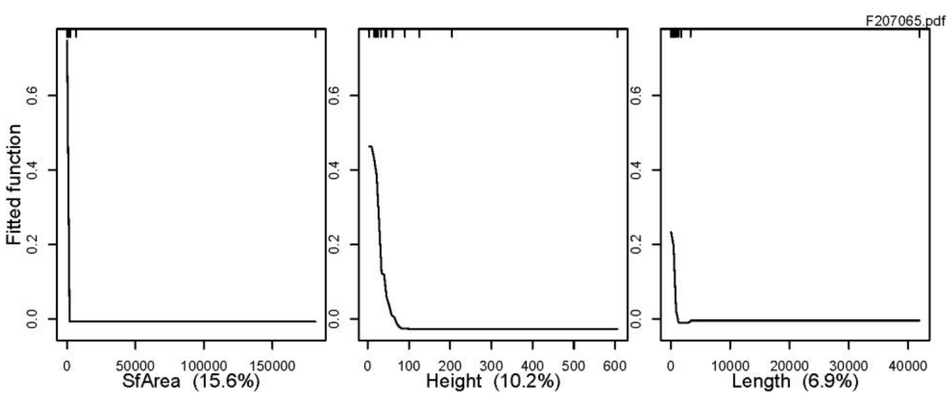

F205055.pdf
Fitted function
SLP_PERC (18%)
POINT_X (7%)
bigPlyrSum (5.2%)


F205057.pdf
Fitted function
POINT_X (28.6%)
BFI_MEAN (6.7%)
CNTC_MEAN (6.5%)

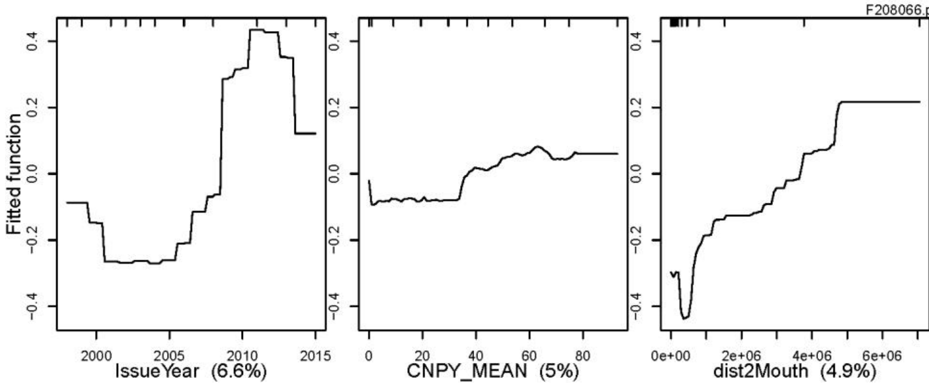

F206058.pdf
Fitted function
Mode (39.2%)
Height (7.8%)
POINT_X (4.8%)


F206059.pdf
Fitted function
Mode (24.6%)
POINT_Y (9.2%)
N_PASTUREC (5.7%)

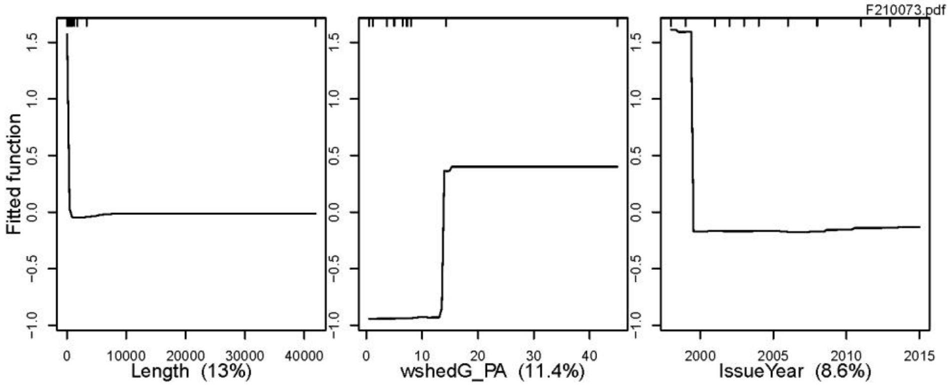

F206061.pdf
Fitted function
Mode (22.7%)
owner (9.2%)
Length (4.3%)


F207065.pdf
Fitted function
SfArea (15.6%)
Height (10.2%)
Length (6.9%)

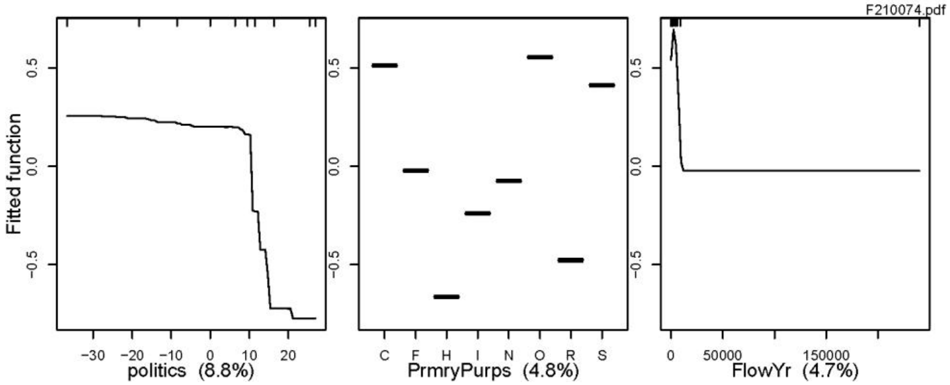
F208066.pdf



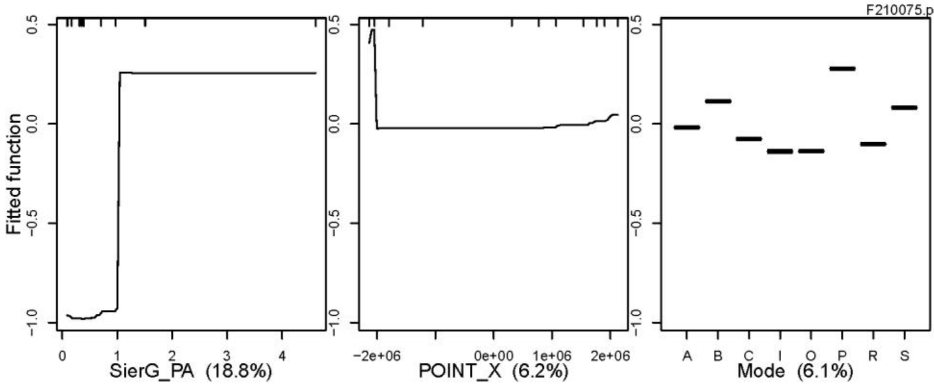
F210073.pdf



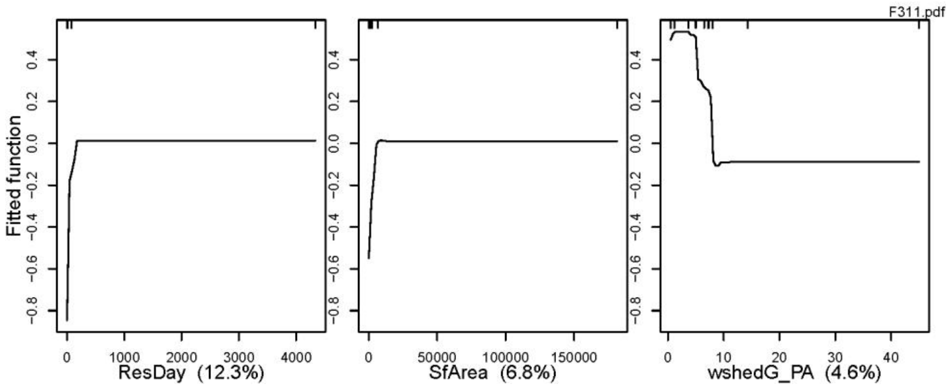
F210074.pdf

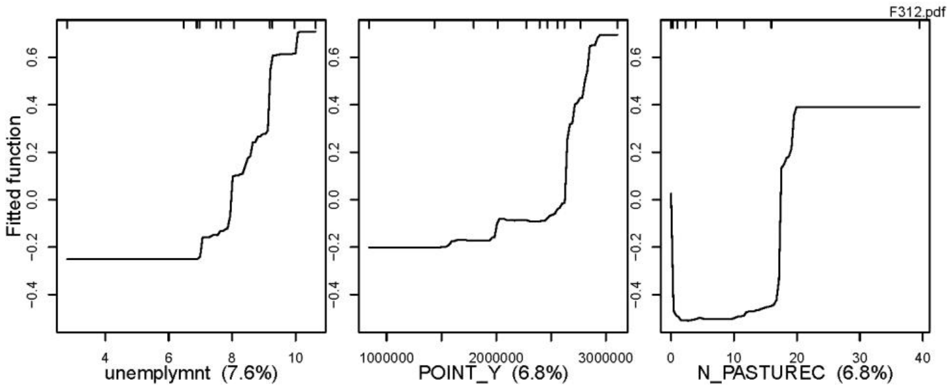


F210075.pdf

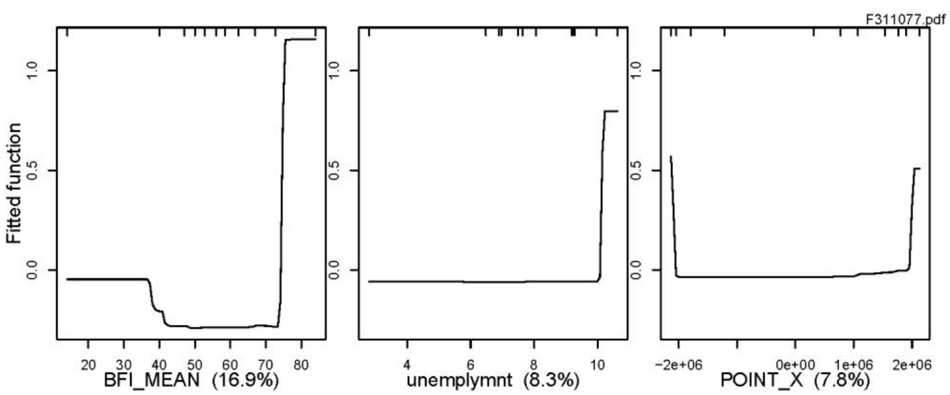


Tier 2 Water Quality

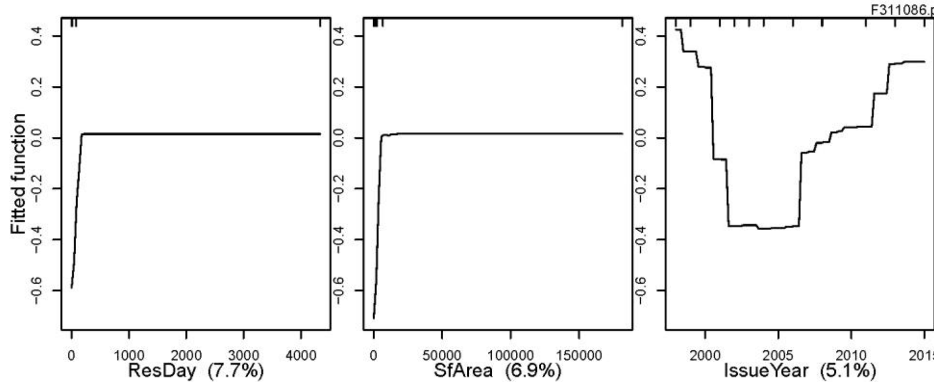




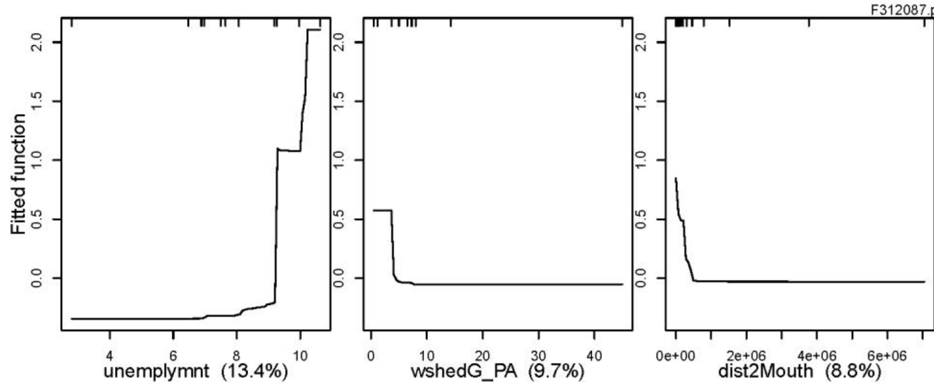
Tier 3 Water Quality



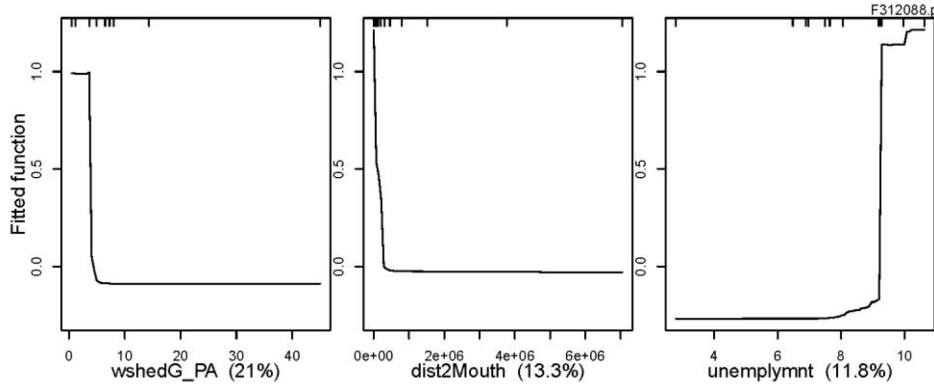
F311086.pdf



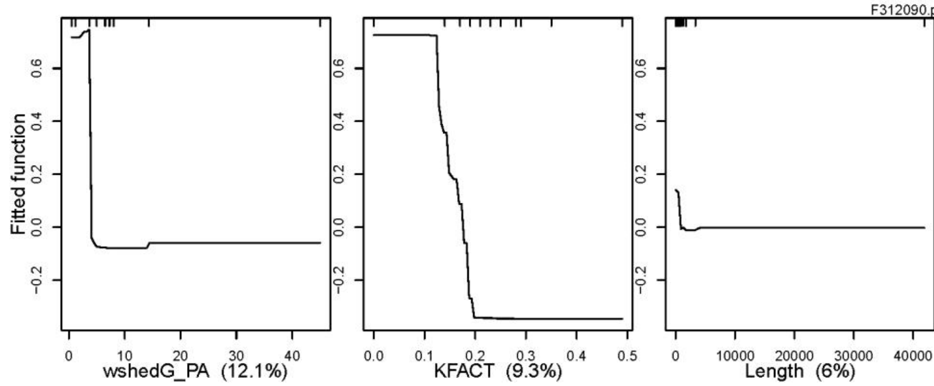
F312087.pdf

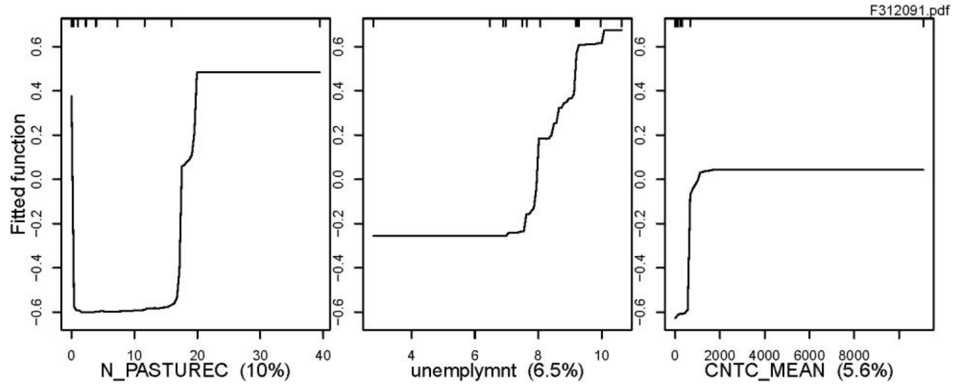


F312088.pdf

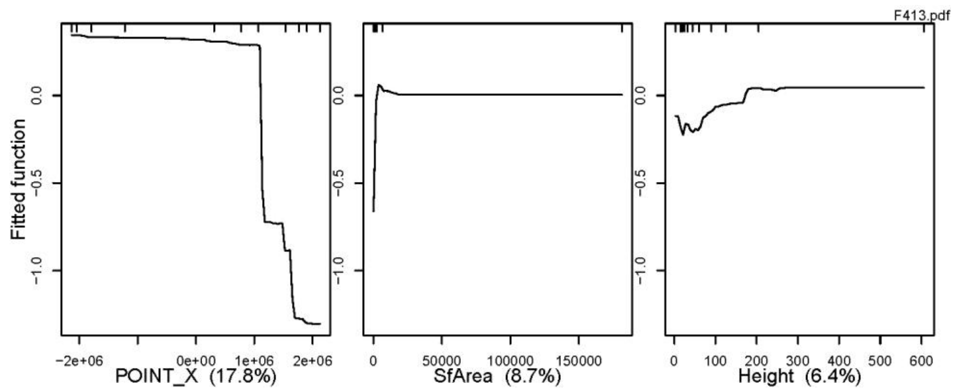


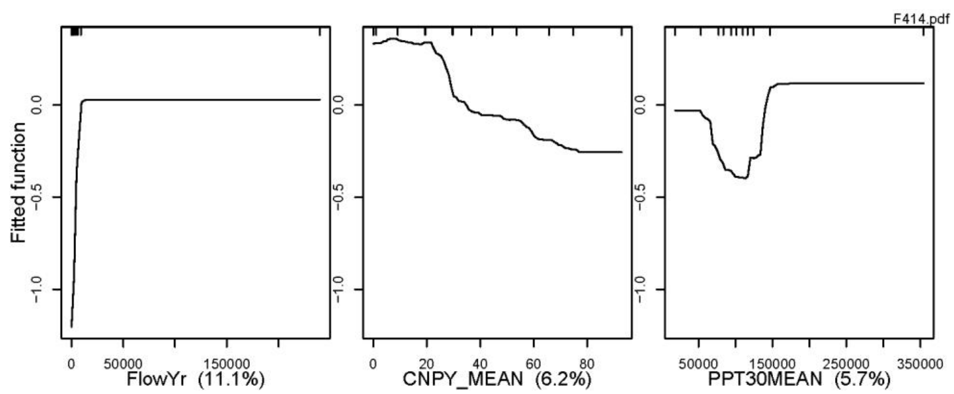
F312090.pdf



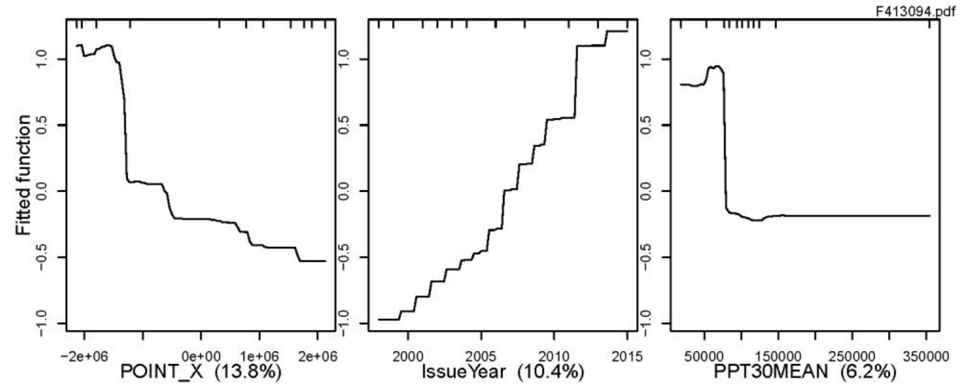


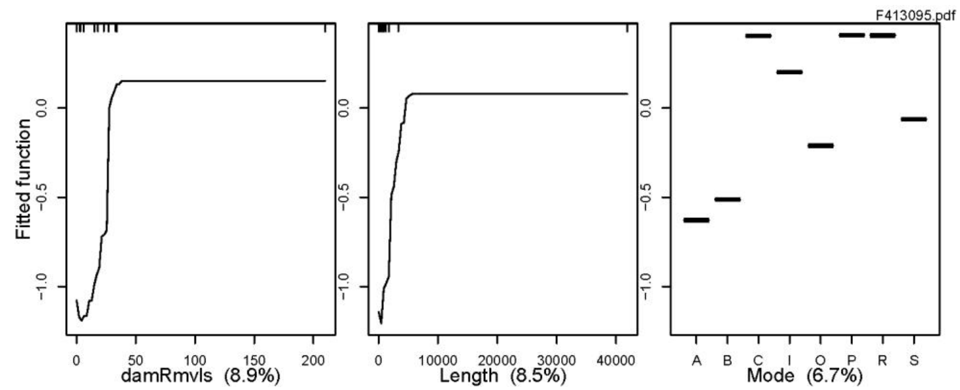
Tier 2 Biodiversity

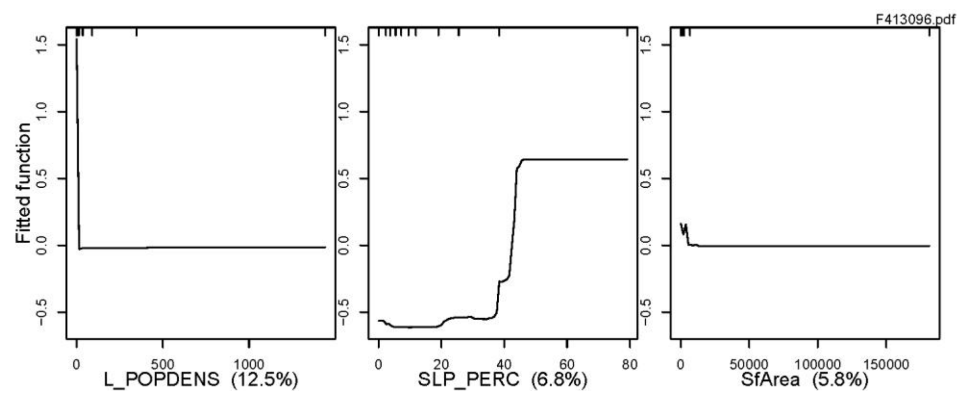


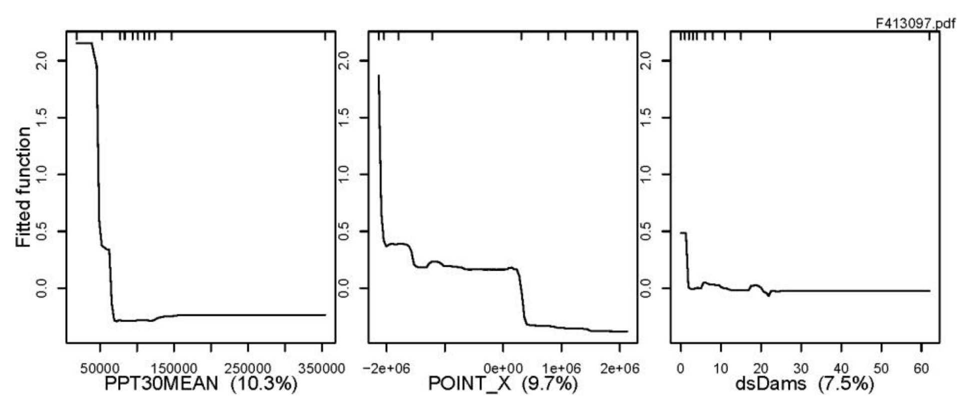


Tier 3 Biodiversity

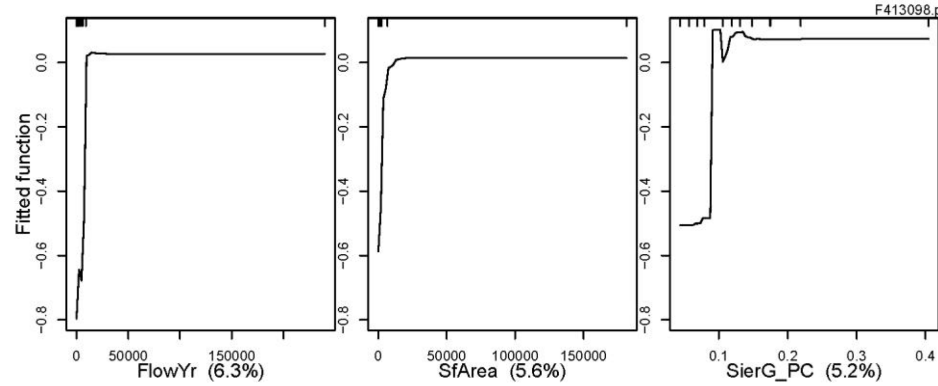




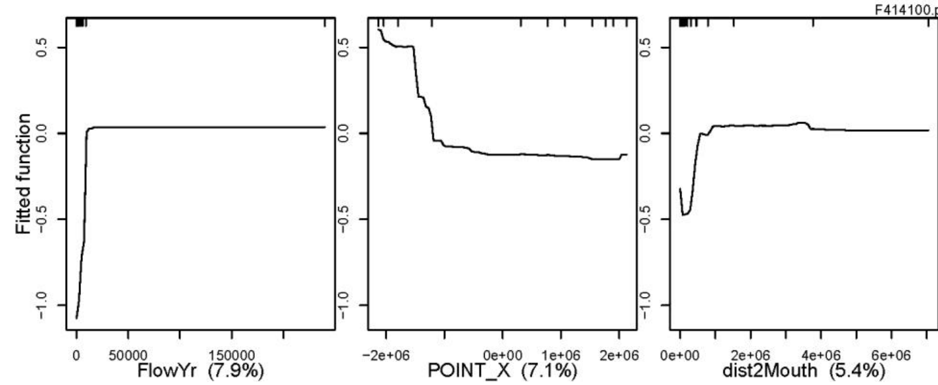




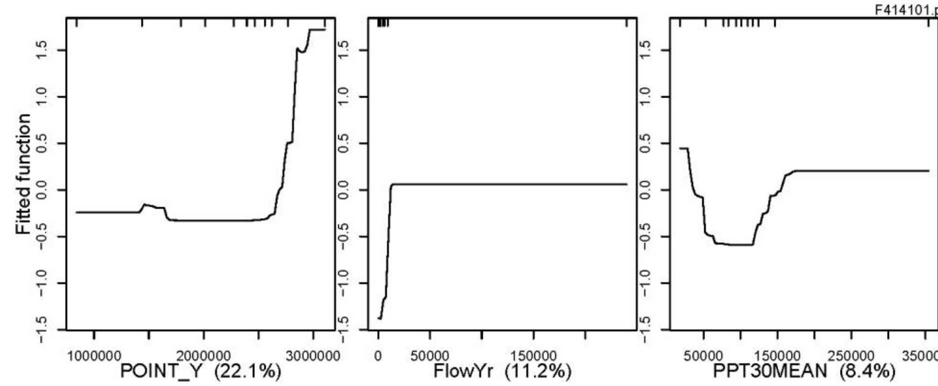
F413098.pdf



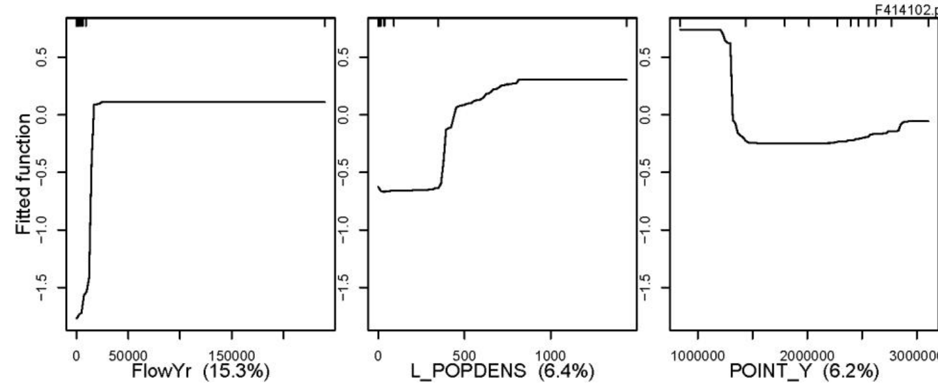
F414100.pdf



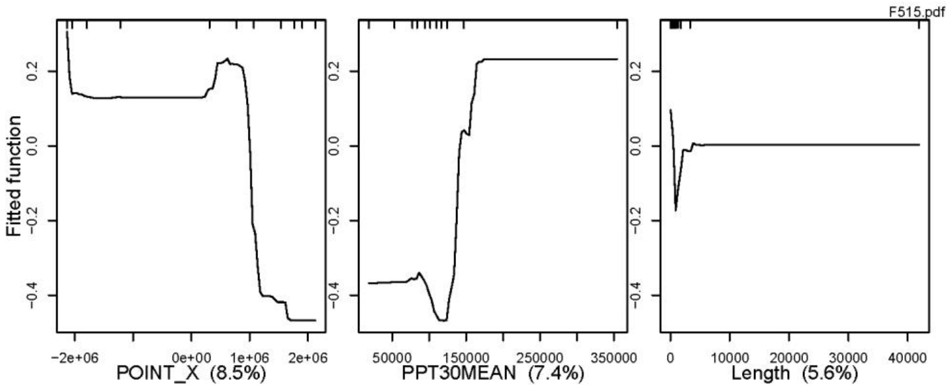
F414101.pdf

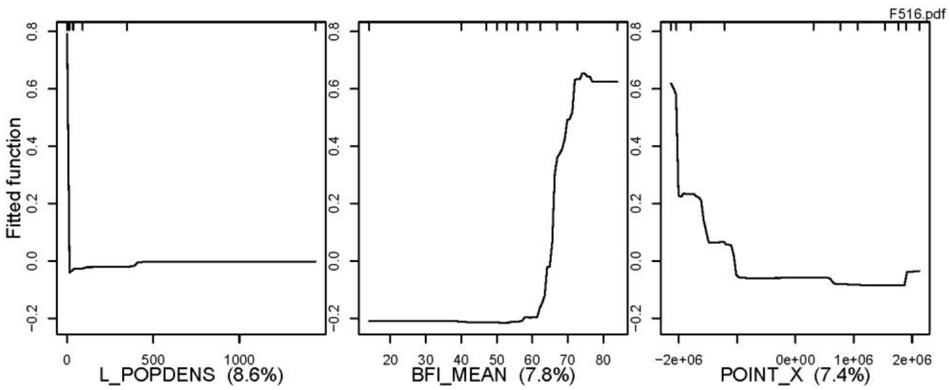


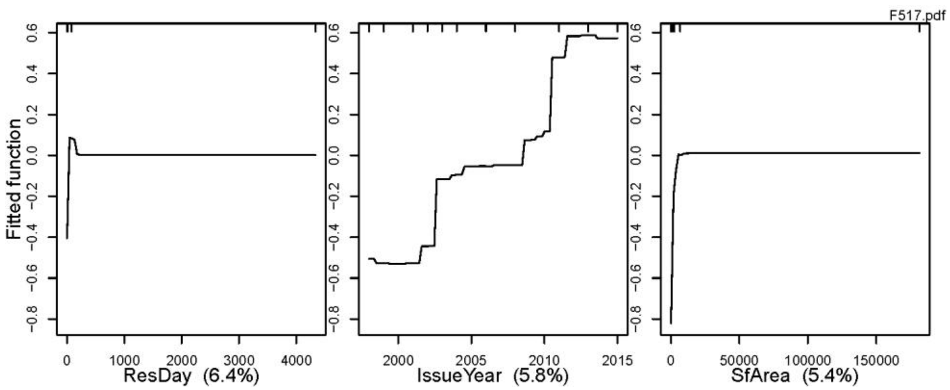
F414102.pdf

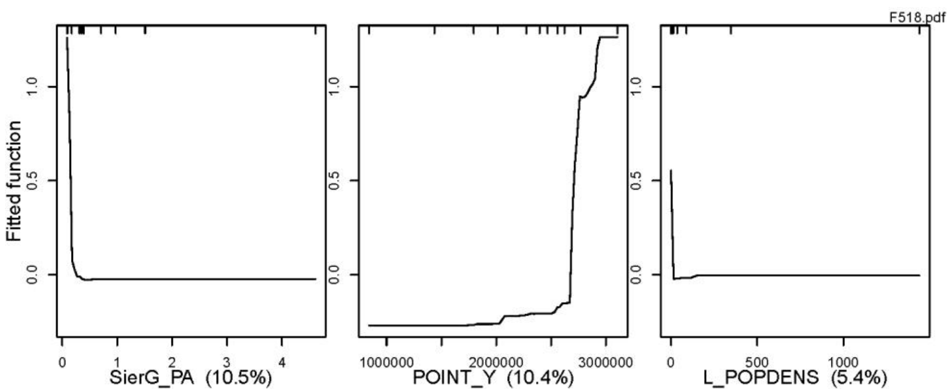


Tier 2 Habitat

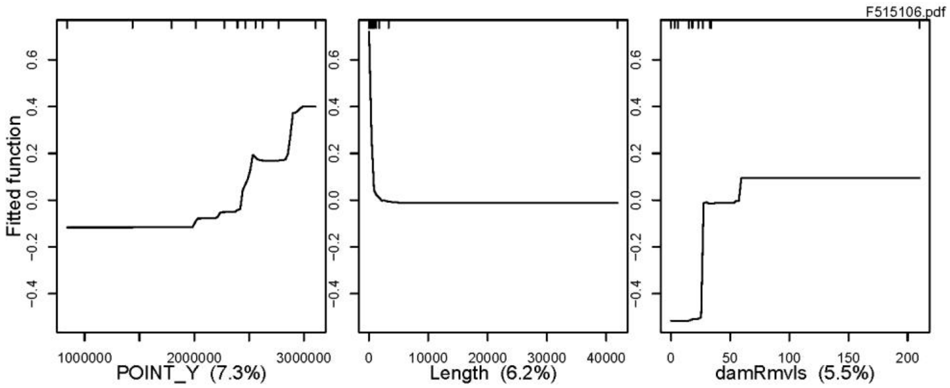


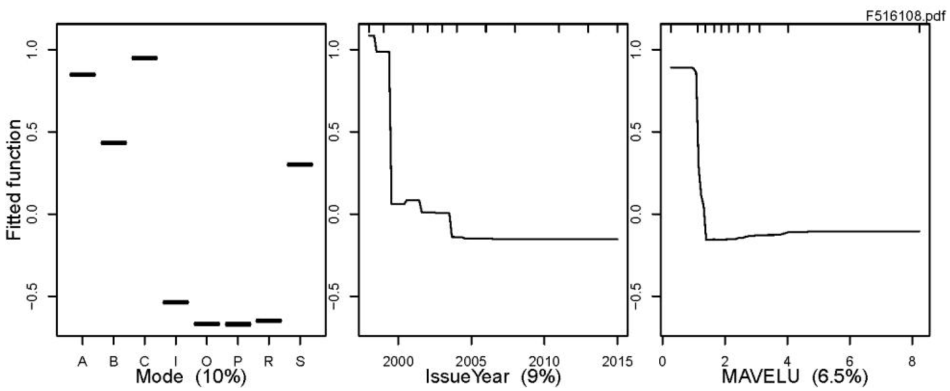


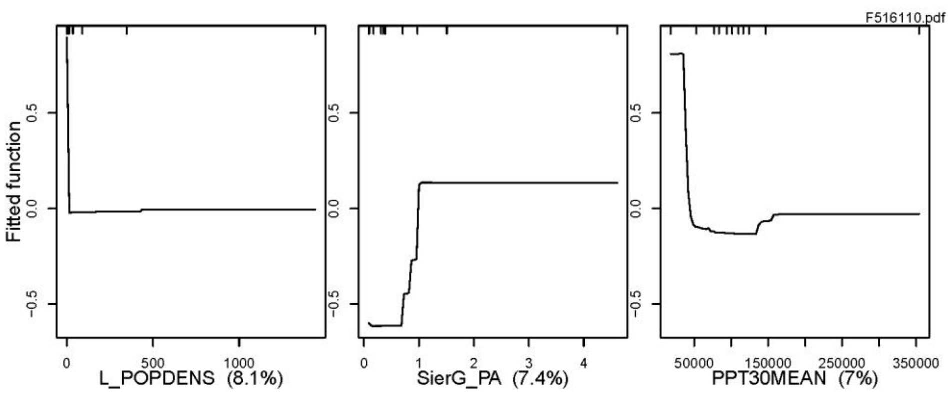


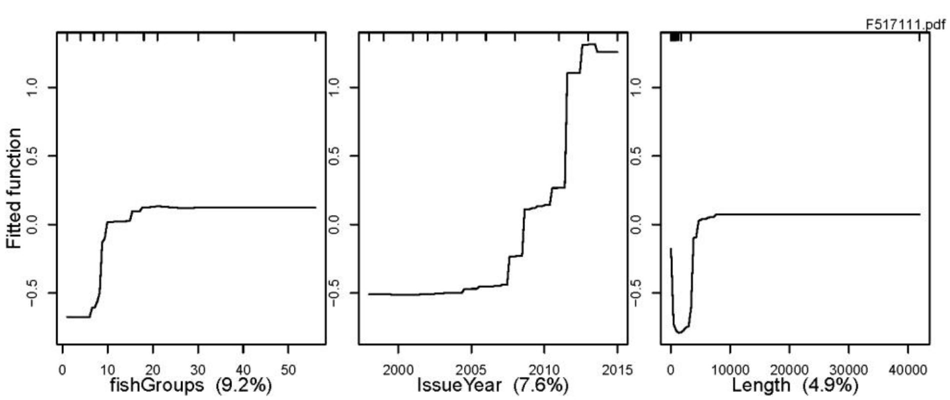


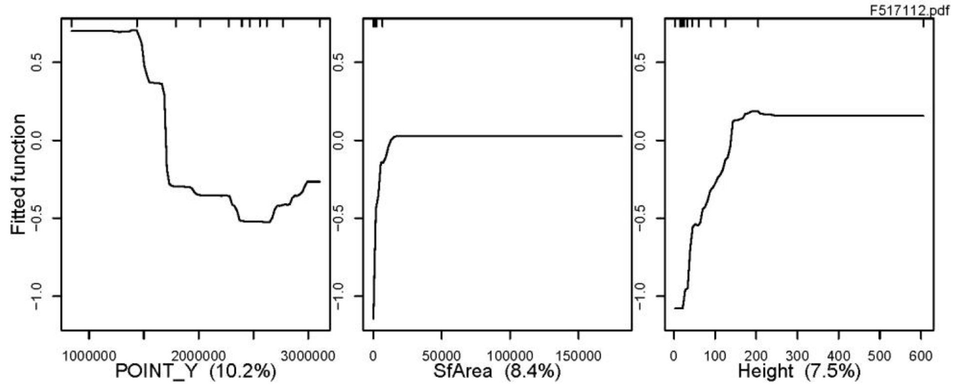
Tier 3 Habitat

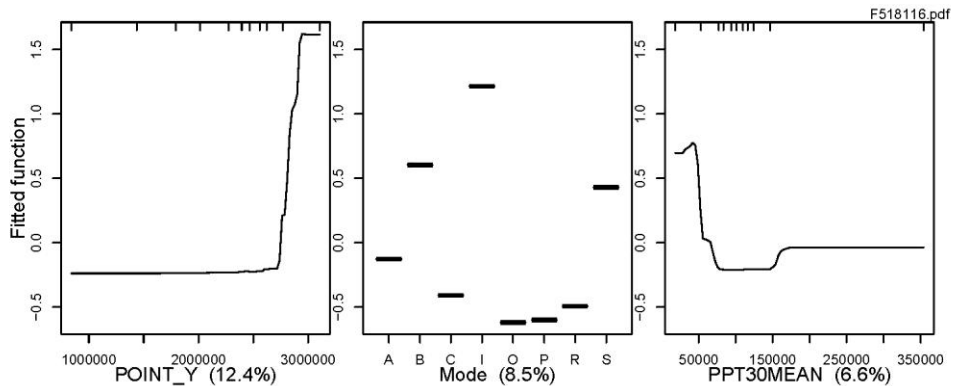




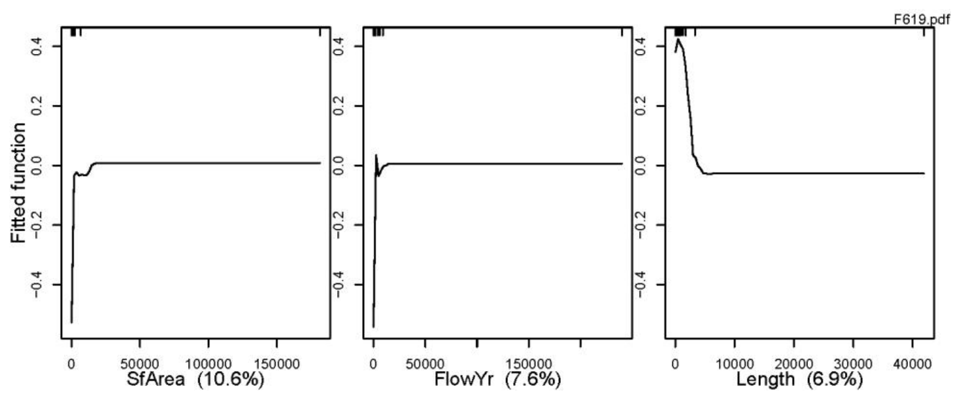


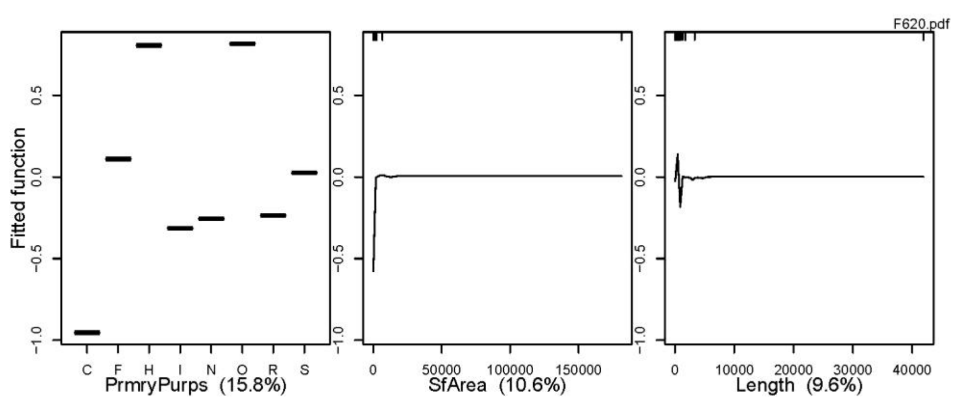




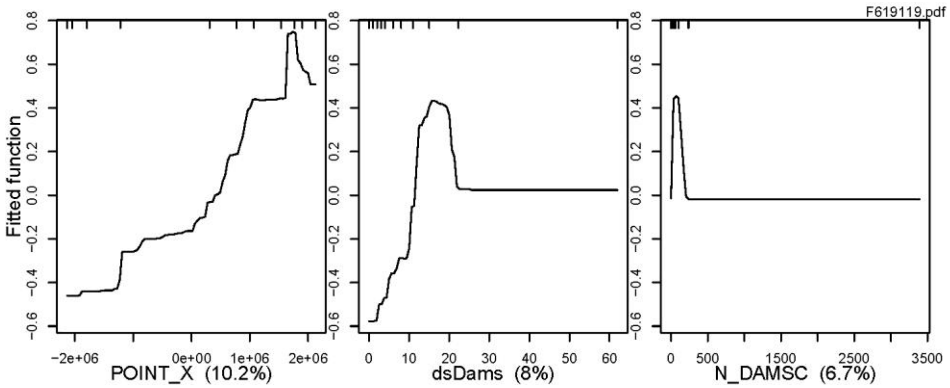


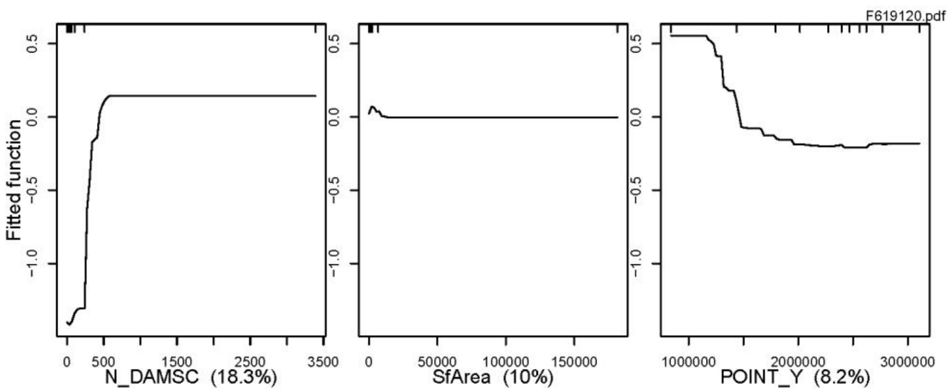
Tier 2 Recreation

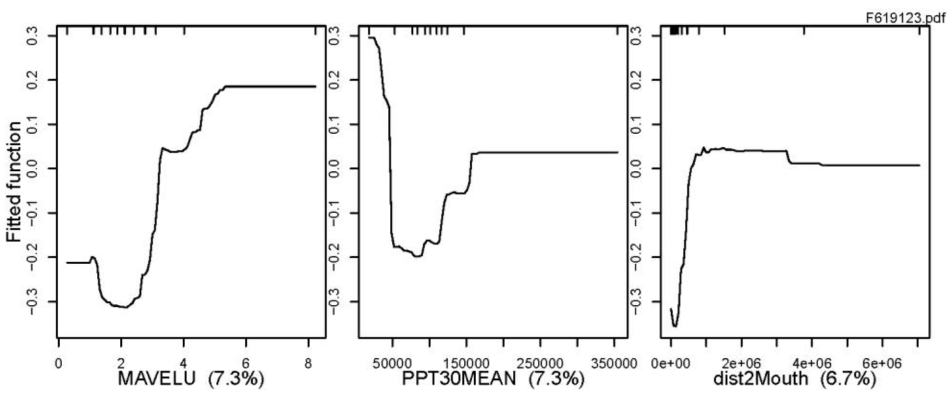


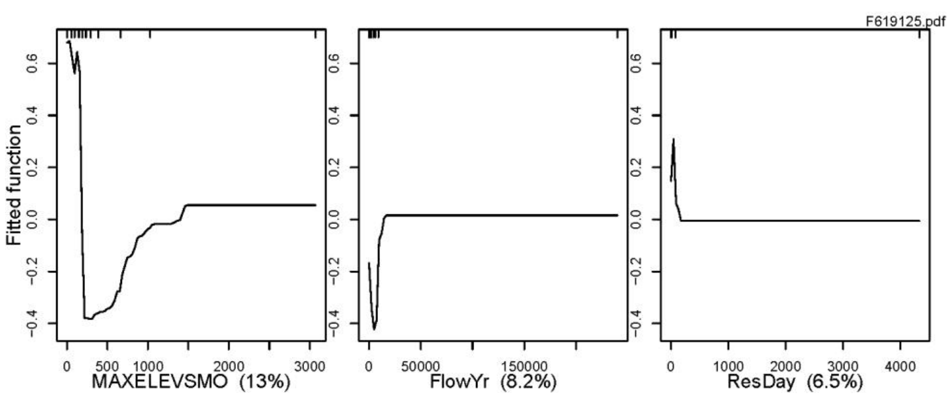


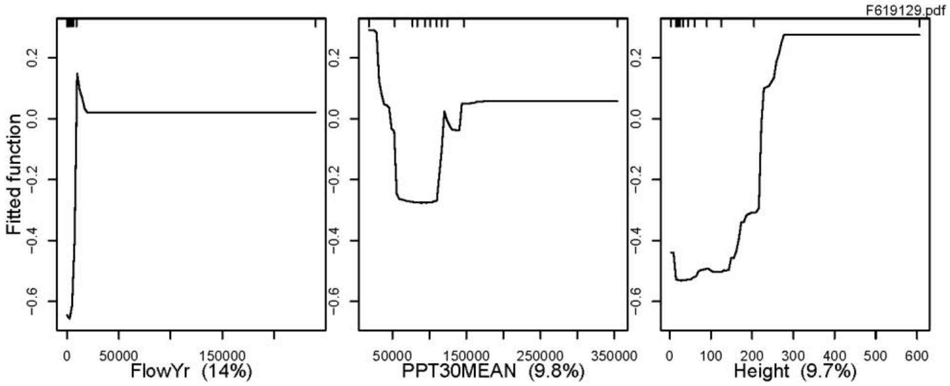
Tier 3 Recreation

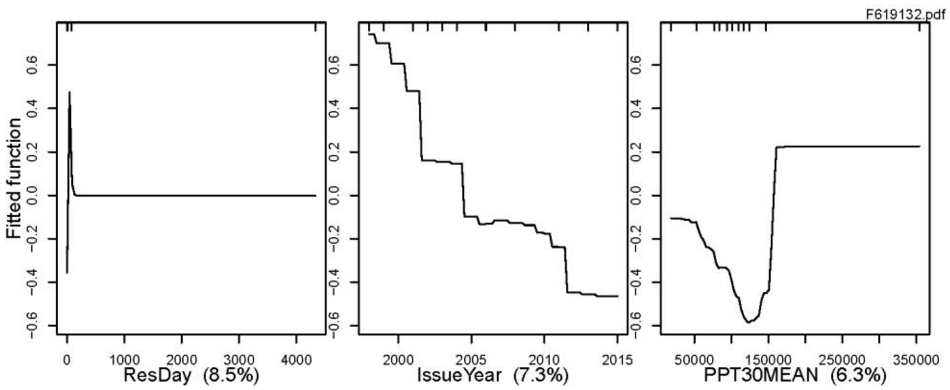









F619129.pdf
Fitted function
FlowYr (14%)
PPT30MEAN (9.8%)
Height (9.7%)


F619132.pdf
Fitted function
ResDay (8.5%)
IssueYear (7.3%)
PPT30MEAN (6.3%)

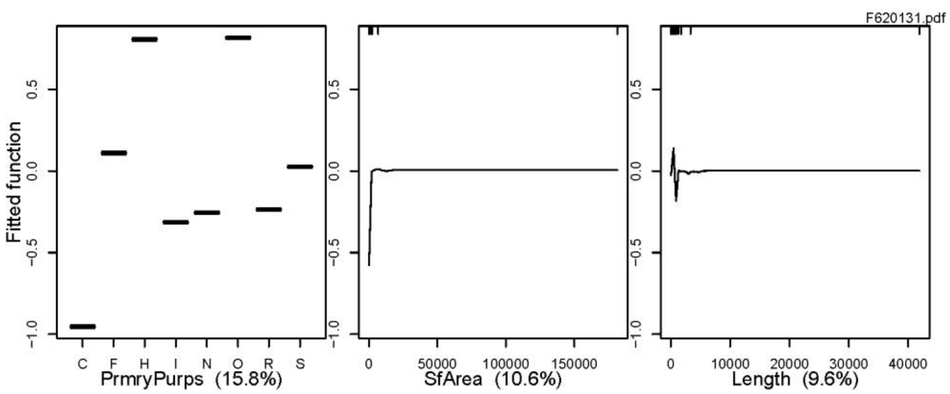

F620131.pdf
Fitted function
C F H I N O R S
PrmryPurps (15.8%)
SfArea (10.6%)
Length (9.6%)

Table 1. Summary and description of input variables for the boosted regression tree models. Variables or units in bold and underlined indicate remaining predictor variables after collinearity analysis.

Variable	Description	Units	Spatial scale	Source	Models
			Biological		
bigPlyrSum	Major migratory fish species	Count	HUC8 watershed	NatureServe fish distributions, expert opinion	P, H, B, A, R
			Facility		
Height	Dam height	Feet	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
HY_MW	Generation capacity	Megawatts	Hydropower dam	ORNL NHAAP database	--
HY_MWh	Generation	Megawatt-hours	Hydropower plant	ORNL NHAAP database	--
Length	Dam length	Feet	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
Mode	Dam mode-of-operation	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
owner	Ownership type	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
PrmyPurps	Dam primary purpose	Categorical	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
			Human		
birdG_xx	National Audubon Society chapters	Count, PA, PC	State	National Audubon Society	--
damR_xx	Dam removals	**Count**, PA, PC	State	American Rivers	P, H, W, B, A, R
education	Education attainment - percent bachelor's degree or higher	Percent	USEPA Level 3 Ecoregion	US Census	P, H, W, B, A, R
FishG_xx	TU and CCA chapters	**Count**, PA, **PC**	State	TU, CCA	P, H, B, A, R
hshldincm	Mean household income	US dollars	USEPA Level 3 Ecoregion	US Census	--
IssueYear	FERC hydropower project license issue year	Year	Hydropower plant	ORNL NHAAP database	P, H, W, B, A, R
LandG_xx	Land trusts	Count, PA, PC	State	Land Trust Alliance	--
politics	see note*	Difference	State	US Federal Election Commission	P, H, W, B, A, R
xx_POPDENS	2000 population density	Individuals/km2 (L and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A, R
q12_avg	Survey response on environmental impact of dams	Rank	Geographic region	MIT Energy Survey, 2008	P, H, W, B, A, R
q16_avg	Survey response on increasing or decreasing hydro power	Rank	Geographic region	MIT Energy Survey, 2008	P, H, W, B, A, R
SierG_xx	Sierra Club chapters	Count, **PA**, **PC**	State	Sierra Club	P, H, W, B, A, R
unemplymnt	Unemployment	Percent	USEPA Level 3 Ecoregion	US Census	P, H, W, B, A, R
usHouse	LCV US House of Rep. mean scorecard for 1998 to 2013	Percent	State	League of Conservation Voters	--
usSenate	LCV US Senate mean scorecard for 1998 to 2013	Percent	State	League of Conservation Voters	--
wshed_xx	Local watershed associations	**Count, PA, PC**	State	USEPA	P, H, W, B, A, R
			Hydrology		
ADRAIN	Total artificial drainage area	Square meters	NHD Plus V1 Catchment	USGS	H, W, A
BFI_MEAN	Mean base-flow index for GW discharge into streams	Percent	NHD Plus V1 Catchment	USGS	H, W, A
CNTC_MEAN	Baseflow residence time in the subsurface	Days	NHD Plus V1 Catchment	USGS	H, W, A
DITCHES	Estimated area subject to the practice of ditches	Square meters	NHD Plus V1 Catchment	USGS	--
FlowYr	Average annual flow	Cubic feet per second	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
IRRIG	Estimated area subject to the practice of irrigation	Square meters	NHD Plus V1 Catchment	USGS	--
KFACT	Soil erodibility factor	Dimensionless	NHD Plus V1 Catchment	USGS	H, W, A
MAVELU	Mean Annual Velocity (fps) at bottom of flowline	Cubic feet per second	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
MEAN_IEOF	Mean value for infiltration-excess overland flow	Percent	NHD Plus V1 Catchment	USGS	H, W, A
MEAN_RCHRG	Mean annual natural groundwater recharge	Millimeters	NHD Plus V1 Catchment	USGS	--
nidStorSum	Accumulated upstream storage	Acre-feet	Hydropower dam	National Anthropogenic Barriers Dataset	H, W, A
ResDay	Reservoir residence time	Days	Hydropower dam	ORNL NHAAP database	P, H, W, B, A, R
SATOF_MEAN	Average value of saturation overland flow	Percent	NHD Plus V1 Catchment	USGS	H, W, A
SfArea	Reservoir surface area	Acres	Hydropower dam	National Inventory of Dams	P, H, W, B, A, R
Stor	Reservoir storage	Acre-feet	Hydropower dam	National Inventory of Dams	--
TILES	Estimated area of tile drains	Square meters	NHD Plus V1 Catchment	USGS	--
			Landscape		
xx_CROPS	Land cover classified as cultivated crops	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_NPDES	Number of NPDES sites	Count (**L** and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W
xx_PASTURE	Land cover classified as pasture/hay	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_ROADCRC	Road-stream crossings	Count (L and N)	NHD Plus V1 Catchment	USGS, National Fish Habitat Partnership	--
xx_URBANHC	Land cover classified as high intensity urban	Percent (L and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	--
xx_URBANL	Land cover classified as low intensity urban	Percent (**L and N**)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
xx_URBANM	Land cover classified as medium intensity urban	Percent (**L** and N)	NHD Plus V1 Catchment	National Fish Habitat Partnership	H, W, B, A
CNPY_MEAN	Mean canopy cover	Percent	NHD Plus V1 Catchment	USGS	H, W, B, A
CROP_AREA	Total crop area for fertilizer/manure derived from land use	Square meters	NHD Plus V1 Catchment	USGS	H, W, B, A
d303_count	Impaired or threatened waters	Count	NHD Plus V1 Catchment	USEPA 303(d) list	H, W, B, A
IMPV_MEAN	Mean impervious surface	Percent	NHD Plus V1 Catchment	USGS	--
L_MINES	Number of mines or mineral processing plants	Count	NHD Plus V1 Catchment	National Fish Habitat Partnership	W
L_ROADLEN	Length of roads	Meters	NHD Plus V1 Catchment	National Fish Habitat Partnership	P, H, W, B, A
MAXELEVSMO	Maximum elevation	Meters	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
PPT30MEAN	30-year (1971-2000) average annual precipitation	Millimeters	NHD Plus V1 Catchment	USGS	P, H, W, B, A, R
SLOPE	Slope of stream reach	Unitless	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R
SLP_PERC	Landscape slope	Percent	NHD Plus V1 Catchment	USGS	P, H, W, B, A, R
TMAX30_MEA	30-year (1971-2000) average annual maximum temperature	Celsius	NHD Plus V1 Catchment	USGS	--
			Location		
POINT_X	Longitude	Decimal degrees	Hydropower dam	National Anthropogenic Barriers Dataset	P, H, W, B, A, R
POINT_Y	Latitude	Decimal degrees	Hydropower dam	National Anthropogenic Barriers Dataset	P, H, W, B, A, R
			Stream network		
dist2Mouth	Stream network distance to network mouth	Meters	Entire downstream flow path	Calculated from NHD Plus V1 flowlines	P, H, W, B, A, R
DrArea	Drainage area upstream of dam	Square miles	Hydropower dam	National Inventory of Dams	--
dsDams	Downstream dams on flow path to network mouth	Count	Entire downstream flow path	Calculated from NHD Plus V1 and NABD	P, H, W, B, A, R
N_DAMSC	Number of dams within network catchment	Count	NHD Plus V1 Catchment	National Fish Habitat Partnership	P, H, W, B, A, R
SO	Strahler stream order	Strahler number	NHD Plus V1 Flowline	NHD Plus V1	P, H, W, B, A, R

PA = per area; PC = per capita; L = local catchment; N = entire network catchment; xx indicates variable derived for multiple units; P = fish passage; H = hydrology; W = water quality; B = biodiversity; A = habitat; R = recreation; *politics is the difference between mean percent democrat and republican from 1996 to 2012 presidential elections; LCV = League of Conservation Voters; TU = Trout Unlimited; CCA = Coastal Conservation Association.

Table
Click here to download Table: Table 2.docx

Table 1. Model results summary.

Tier 1	Tier 2	Tier 3	ModelID	Trees	CV ROC	V ROC	OT	MI	Influential Variable 1	Influential Variable 2	Influential Variable 3
Fish Passage	DS	NA (see Tier 2 category)	F101	5550	0.867	0.916	0.36	-0.165	POINT_Y (13)	dist2Mouth (10.7)	FlowYr (8)
	Fish Passage	DS Passage Plan Study Design	F101010	5300	0.892	0.829	0.30	-0.309	POINT_Y (11.1)	dist2Mouth (9.9)	MAXELEVSMO (9.1)
	US	NA (see Tier 2 category)	F102	5250	0.899	0.896	0.29	-0.436	MAXELEVSMO (14.5)	POINT_Y (13)	bigPlyrSum (8.3)
	Fish Passage	Eelway	F102017	4350	0.956	0.966	0.33	*-1.178*	MAXELEVSMO (36.5)	POINT_X (11.9)	bigPlyrSum (7.7)
		US passage study plan or design	F102023	5000	0.909	0.854	0.29	*-0.703*	MAXELEVSMO (12.3)	POINT_X (10.5)	POINT_Y (9.7)
	Passage	NA (see Tier 2 category)	F103	2850	0.780	0.856	0.27	*-0.467*	POINT_Y (9.3)	dsDams (8.1)	MAXELEVSMO (7.5)
	Planning	DS fish passage mon. sampling	F103029	3200	0.888	0.924	0.22	*-0.908*	MAXELEVSMO (18.2)	Height (10)	POINT_X (8.9)
		Fish passage & operations plan	F103031	1050	0.739	0.749	0.08	-0.246	wshed_PC (11.8)	L_ROADLEN (10.5)	dsDams (9.9)
		Fish stranding plan mon. evaluation	F103033	1100	0.712	0.605	--		--	--	--
		US fish passage mon. sampling	F103036	3050	0.891	0.865	0.18	*-0.504*	MAXELEVSMO (16.2)	Height (11.9)	POINT_Y (8.7)
	Entrainment	NA (see Tier 2 category)	F104	3450	0.849	0.756	0.29	-0.222	ResDay (9.8)	SierG_PC (7.8)	politics (7)
		Trash or bar rack	F104043	3700	0.917	0.833	0.22	0.147	POINT_X (10.4)	SierG_PC (9.6)	fishGroups (8.8)
Hydrology	Flow	NA (see Tier 2 category)	F205	4700	0.785	0.787	0.56	*-1.413*	PrmryPurps (7.7)	N_URBANLC (5.3)	Height (5.1)
	Mitigation	Tailrace flow mon. plan	F205045	5850	0.822	0.920	0.39	-0.076	POINT_X (8.5)	IssueYear (4.8)	politics (4.7)
		Tailrace flow or stage mon. equipment	F205048	2650	0.784	0.867	0.17	-0.191	N_CROPSC (9)	Length (7.3)	dist2Mouth (4.7)
		Tailrace ramping rate restriction	F205050	2400	0.790	0.834	0.19	-0.333	SfArea (8.3)	CNPY_MEAN (5.9)	nidStorSum (5.8)
		Bypass flow mon. plan	F205052	2900	0.802	0.853	0.20	*-0.522*	politics (6.9)	Length (6.7)	SfArea (5.6)
		Bypass flushing or flood flow	F205054	3750	0.890	0.951	0.15	0.187	POINT_X (24.6)	N_PASTUREC (6.8)	PPT30MEAN (6.2)
		Bypass flow or stage mon. equipment	F205055	1500	0.735	0.779	0.13	*-0.679*	SLP_PERC (18)	POINT_X (7)	bigPlyrSum (5.2)
		Bypass ramping rate restriction	F205057	2900	0.878	0.802	0.16	0.170	POINT_X (28.6)	BFI_MEAN (6.7)	CNTC_MEAN (6.5)
	Tailrace	NA (see Tier 2 category)	F206	3800	0.863	0.845	0.63	-0.198	POINT_X (10.4)	Mode (10.3)	SLP_PERC (7.2)
	Minimum	Run-of-river Tailrace	F206058	3700	0.904	0.911	0.37	-0.349	Mode (39.2)	Height (7.8)	POINT_X (4.8)
	Flow	Seasonal Tailrace	F206059	2700	0.850	0.846	0.20	0.087	Mode (24.6)	POINT_Y (9.2)	N_PASTUREC (5.7)
		Year-round Tailrace	F206061	1500	0.787	0.899	0.19	-0.312	Mode (22.7)	owner (9.2)	Length (4.3)
	Bypass	NA (see Tier 2 category)	F207	3250	0.808	0.771	0.46	*-0.676*	SfArea (16.3)	MAXELEVSMO (6.6)	MAVELU (4.7)
	Minimum	Seasonal Bypass	F207063	1450	0.678	0.668	--		--	--	--
	Flow	Year-round Bypass	F207065	1200	0.720	0.805	0.23	-0.339	SfArea (15.6)	Height (10.2)	Length (6.9)
	Sediment	NA (see Tier 2 category)	F208	4850	0.767	0.851	0.49	-0.364	IssueYear (6.1)	CNPY_MEAN (5.3)	unemplymnt (4.6)
		Sediment & erosion control plan or mon.	F208066	4100	0.778	0.838	0.47	-0.257	IssueYear (6.6)	CNPY_MEAN (5)	dist2Mouth (4.9)
	Recreation	NA (see Tier 2 category)	F209	1550	0.733	0.796	0.17	-0.127	POINT_X (12.3)	Height (6.8)	SierG_PA (6.8)
	Flow	Provide recreational flow releases	F209071	700	0.655	0.713	--		--	--	--
	Operations	NA (see Tier 2 category)	F210	3050	0.734	0.819	0.53	0.040	FlowYr (8.6)	q16_avg (4.5)	N_URBANLC (4.5)
		Flow mgmt. plan	F210073	3350	0.893	0.985	0.09	-0.399	Length (13)	wshedG_PA (11.4)	IssueYear (8.6)
		Operations compliance mon. plan	F210074	5150	0.807	0.913	0.41	0.146	politics (8.8)	PrmryPurps (4.8)	FlowYr (4.7)
		Provide flow or lake levels electronically	F210075	1750	0.795	0.917	0.14	-0.282	SierG_PA (18.8)	POINT_X (6.2)	Mode (6.1)
Water Quality	DS	NA (see Tier 2 category)	F311	4300	0.838	0.887	0.55	-0.240	ResDay (12.3)	SfArea (6.8)	wshedG_PA (4.6)
	Water	Benthic macroinvertebrate mon.	F311077	1500	0.724	0.938	0.12	-0.396	BFI_MEAN (16.9)	unemplymnt (8.3)	POINT_X (7.8)
	Quality	DO enhancement or mitigation plan	F311078	2200	0.832	0.676	--		--	--	--
		Water quality mon. plan	F311086	6000	0.852	0.873	0.50	-0.375	ResDay (7.7)	SfArea (6.9)	IssueYear (5.1)
	US	NA (see Tier 2 category)	F312	4000	0.831	0.860	0.23	-0.320	unemplymnt (7.6)	POINT_Y (6.8)	N_PASTUREC (6.8)
	Water	Fish tissue sampling & analysis	F312087	4500	0.965	0.823	0.39	*-1.704*	unemplymnt (13.4)	wshedG_PA (9.7)	dist2Mouth (8.8)
	Quality	Impoundment sediment analysis	F312088	4100	0.993	0.999	0.19	0.334	wshedG_PA (21)	dist2Mouth (13.3)	unemplymnt (11.8)
		Inflow water quality mon. plan	F312090	1650	0.831	0.904	0.11	0.106	wshedG_PA (12.1)	KFACT (9.3)	Length (6)
		Impoundment water quaity mon. plan	F312091	4100	0.828	0.805	0.22	-0.166	N_PASTUREC (10)	unemplymnt (6.5)	CNTC_MEAN (5.6)
Biodiversity	Terrestrial	NA (see Tier 2 category)	F413	4150	0.847	0.832	0.64	*-0.638*	POINT_X (17.8)	SfArea (8.7)	Height (6.4)
		Noxious weed & invasive plant mgmt.	F413094	6650	0.912	0.901	0.39	0.068	POINT_X (13.8)	IssueYear (10.4)	PPT30MEAN (6.2)
		Species conservation mgmt. mon.	F413095	5850	0.832	0.899	0.40	-0.265	damRmvls (8.9)	Length (8.5)	Mode (6.7)
		T&E species protection plan	F413096	3950	0.879	0.905	0.21	*0.965*	L_POPDENS (12.5)	SLP_PERC (6.8)	SfArea (5.8)
		Transmission related avian & bat protection	F413097	6250	0.936	0.941	0.19	-0.109	PPT30MEAN (10.3)	POINT_X (9.7)	dsDams (7.5)
		Wildlife terrestrial habitat mgmt.	F413098	4100	0.844	0.937	0.27	0.146	FlowYr (6.3)	SfArea (5.6)	SierG_PC (5.2)
	Aquatic	NA (see Tier 2 category)	F414	3500	0.791	0.859	0.35	-0.271	FlowYr (11.1)	CNPY_MEAN (6.2)	PPT30MEAN (5.7)
		Aquatic species conservation mgmt. mon.	F414100	3400	0.807	0.869	0.34	-0.336	FlowYr (7.9)	POINT_X (7.1)	dist2Mouth (5.4)
		Diadromous species mgmt. mon.	F414101	3000	0.871	0.901	0.26	0.124	POINT_Y (22.1)	FlowYr (11.2)	PPT30MEAN (8.4)
		Invasive aquatic species mgmt.	F414102	2800	0.800	0.881	0.19	*0.552*	FlowYr (15.3)	L_POPDENS (6.4)	POINT_Y (6.2)
Habitat	Fisheries	NA (see Tier 2 category)	F515	2650	0.776	0.730	0.27	-0.038	POINT_X (8.5)	PPT30MEAN (7.4)	Length (5.6)
		DS habitat enhancement	F515105	1200	0.687	0.680	--		--	--	--
		DS woody debris restoration or passage	F515106	2850	0.863	0.879	0.25	0.071	POINT_Y (7.3)	Length (6.2)	damRmvls (5.5)
	Riparian	NA (see Tier 2 category)	F516	2600	0.771	0.869	0.28	0.084	L_POPDENS (8.6)	BFI_MEAN (7.8)	POINT_X (7.4)
		Establish riparian buffers	F516108	3100	0.866	0.864	0.25	*0.673*	Mode (10)	IssueYear (9)	MAVELU (6.5)
		Riparian habitat mon. or planning	F516110	2300	0.793	0.912	0.22	-0.368	L_POPDENS (8.1)	SierG_PA (7.4)	PPT30MEAN (7)
	Reservoir	NA (see Tier 2 category)	F517	6100	0.858	0.905	0.40	-0.082	ResDay (6.4)	IssueYear (5.8)	SfArea (5.4)
		Noxious invasive aquatic plant mgmt.	F517111	6950	0.928	0.952	0.25	*-0.348*	fishGroups (9.2)	IssueYear (7.6)	Length (4.9)
		Shoreline mgmt. plan or program	F517112	4800	0.856	0.952	0.27	0.192	POINT_Y (10.2)	SfArea (8.4)	Height (7.5)
	Wetlands	NA (see Tier 2 category)	F518	3800	0.828	0.874	0.19	0.207	SierG_PA (10.5)	POINT_Y (10.4)	L_POPDENS (5.4)
		Wetland protection	F518116	3500	0.878	0.875	0.14	0.082	POINT_Y (12.4)	Mode (8.5)	PPT30MEAN (6.6)
Recreation	Resources	NA (see Tier 2 category)	F619	3200	0.741	0.744	0.65	0.089	SfArea (10.6)	FlowYr (7.6)	Length (6.9)
	and	Boating facilities	F619118	1200	0.625	0.660	--		--	--	--
	Mitigation	Canoe portage launch	F619119	5000	0.859	0.773	0.37	-0.161	POINT_X (10.2)	dsDams (8)	N_DAMSC (6.7)
		Fishing pier	F619120	1700	0.797	0.869	0.13	0.055	N_DAMSC (18.3)	SfArea (10)	POINT_Y (8.2)
		Interpretive education sign & displays	F619123	1900	0.720	0.731	0.20	-0.192	MAVELU (7.3)	PPT30MEAN (7.3)	dist2Mouth (6.7)
		Parking	F619125	3550	0.715	0.722	0.32	0.441	MAXELEVSMO (13)	FlowYr (8.2)	ResDay (6.5)
		Shoreline access	F619128	750	0.623	0.759	--		--	--	--
		Stocking recreational fish species	F619129	1250	0.756	0.796	0.09	0.069	FlowYr (14)	PPT30MEAN (9.8)	Height (9.7)
		Trail trailhead or camping areas	F619130	3200	0.781	0.601	--		--	--	--
		Other day use area improvements	F619132	4900	0.750	0.781	0.44	-0.430	ResDay (8.5)	IssueYear (7.3)	PPT30MEAN (6.3)
	Planning	NA (see Tier 2 category)	F620	2450	0.753	0.883	0.74	-0.001	PrmryPurps (15.8)	SfArea (10.6)	Length (9.6)
		Recreational mgmt. plan study or mon.	F620131	2450	0.753	0.883	0.74	-0.001	PrmryPurps (15.8)	SfArea (10.6)	Length (9.6)

See Table 1 for variable descriptions; if no influential variables are shown, model rejected due to poor fit; mgmt. = management; DS = downstream; US = upstream; T&E = threatened and endangered, mon. = monitoring; NA = not applicable; CV ROC = internal cross-validation ROC; V ROC = validation ROC; OT = optimal threshold; MI = Moran's Index; italics indicates spatial autocorrelation detected in training data; color scheme for influential variables corresponds to Table 1 color scheme.

Table 3. The 20 most frequently occurring important variables across all Tier 3 models, with potential future research areas that correspond to each variable. F= frequency; Inf = normalized average relative influence.

Variable	Category	F	Inf	Future research area
POINT_X	Location	19	0.70	Regional trends
SfArea	Hydrology	14	0.70	Hydrology/site design
FlowYr	Hydrology	14	0.68	Hydrology/site design
PPT30MEAN	Landscape	14	0.65	Hydrology/site design
POINT_Y	Location	13	0.71	Regional trends
Height	Facility	13	0.55	Hydrology/site design
dist2Mouth	Stream network	12	0.58	Network/landscape position
Length	Facility	11	0.74	Hydrology/site design
MAXELEVSMO	Landscape	10	0.87	Network/landscape position
Mode	Facility	10	0.74	Hydrology/site design
IssueYear	Human	9	0.76	Regulatory tendencies
unemplymnt	Human	7	0.52	Socio-political conditions
L_POPDENS	Human	6	0.77	Socio-political conditions
ResDay	Hydrology	6	0.66	Hydrology/site design
dsDams	Stream network	6	0.64	Network/landscape position
SierG_PC	Human	6	0.53	Socio-political conditions
wshedG_PA	Human	5	0.80	Socio-political conditions
MAVELU	Hydrology	5	0.59	Hydrology/site design
fishG_PC	Human	5	0.52	Socio-political conditions
bigPlyrSum	Biological	5	0.39	Fisheries