8-2015

Conservation Planning in a Changing World

Austin Walker Milt
University of Tennessee - Knoxville, austin.w.milt@gmail.com

To the Graduate Council:

I am submitting herewith a dissertation written by Austin Walker Milt entitled "Conservation Planning in a Changing World." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Ecology and Evolutionary Biology.

Paul R. Armsworth, Major Professor

We have read this dissertation and recommend its acceptance:

Christopher Clark, Louis Gross, Nathan Sanders

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Conservation Planning in a Changing World

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Austin Walker Milt
August 2015

for

Clara Bean

who reminds me that our children are worthy of the cause

Acknowledgements

First and foremost thank you to my major adviser Paul Armsworth for being my closest collaborator, my ultimate source of academic wisdom and strategy and the source most of “my” good ideas, especially when I was stuck, uninspired, or otherwise fed up with academic life. Your reputation for being the most highly skilled, well prepared, lightening fast, constructively critical, and forward thinking academic is, as far as I can tell, completely accurate and is only marred by your willingness to listen and be unquestioningly kind when needed.

I am grateful to my committee Drs. Chris Clark, Lou Gross, and Nate Sanders for constructive feedback and continual support. Your different perspectives have more than once given me the pause I need to become a better scientist.

To my family and close friends, my thanks and love are immeasurable. Mom and Dad: thank you for fostering in me a love of nature, creating an environment in which that love could grow, and providing the support necessary for me to make a difference in the world. Eric Bilodeau and Katelyn Brown-Gomez: thank you for being family as close as blood, as role models for humor, humility, kindness, skill, endurance, and more. You’re in The Circle. Ryan and Adam: thank you for being inspirations in your own right, through testing the waters so I didn’t always have to, and for taking complementary tracks in life. Dylan Magers and Giles Duncan: thank you for having been my two most dependable sources of entertainment, especially when I needed some down time.

Many thanks to the graduate students and postdocs I have had the pleasure to interact with. Thanks especially to the Armsworthians Alison Boyer, Stacie Armsworth, Christine Dumoulin, Rachel Fovargue, Sean Hoban, Gwen Iacona, Eric Larson, Diane Le Bouille, Nate Sutton, and Shelby Ward for listening to half-baked ideas and bad presentations and for reading incoherent babble and turning it all into mostly baked, mostly coherent products. You have been awesome colleagues and my closest friends and confidants here at UT.

My thanks to the SCALE-ITeers for social (and sometimes professional) stimulation outside of EEB. Thanks especially to Harry Richards for including me in SCALE-IT, for ensuring I got the best out of it, and for continuing to be an invaluable source of advice and mentorship long after my SCALE-IT days were officially done.

To colleagues at the Nature Conservancy, including Scott Bearer, Tamara Gagnolet, Nels Johnson and Thomas Minney: thank you for providing insight, direction, and grounding and for being largely responsible for the track I took in my dissertation. Thank you also to

Sally Palmer, Joey Wisby, and Silas Mathes and others at TNC TN for providing datasets and collaboration for my first chapter. Thank you especially to Tamara Gagnolet for being my (second) closest collaborator and growing friend over the past five years. Countless telephone conferences, webinars, document drafts, and meetings with you and the gas industry have done more for my professional growth than I can measure.

Thank you to other colleagues and collaborators over the years, especially Eddy Grey and Steve Copeland at Triana Energy, LLC for feedback and priceless data, and to Daniel Mahr at Cadmus Group for nerding out over Python and doing the brunt of the work on EnSitu.

I have been fortunate to be well funded by diverse sources during my dissertation. From 2010-2012 I was fully funded by the NSF IGERT, SCALE-IT, and received partial funding 2012-2013. I received partial funding from EEB at UTK and the Colcom Foundation in 2012-2013. I was funded by both the Richard King Mellon Foundation and National Institute for Mathematical and Biological Synthesis between 2013 and 2015.

Abstract

As a science and practice dedicated to preventing, stopping, and reversing negative effects on nature, conservation is constantly faced with new challenges. Combine this fact with the rise of large, freely available datasets and computational power, and the result is a need to advance the methods and conceptual approach to conservation planning. In my dissertation I present novel methods and address research questions that aim to keep conservation science and practice relevant and effective in a changing world. This picture of continual change is illustrated in Chapter 1, in which I explore how the ongoing collection of observations of rare species changes spatial conservation priorities. I find that even after a century of data collection, new records do and will continue to significantly affect spatial priorities. I then moved to consider a new threat: the environmental impacts from shale gas surface infrastructure. I focus on how those environmental impacts may be partially abated by changing the locations of infrastructure. In Chapter 2 I assess the relative performance of simple guidelines for placing well pads, access roads, and gathering pipelines for shale gas development. I find that while targeted guidelines can be effective, none are universally so. In Chapter 3, I examine the site-level tradeoffs between reducing environmental impacts and increased construction costs for shale gas surface infrastructure. I find notable heterogeneity among sites in both the degree to which impacts can be reduced and the relative cost of doing so. Finally in Chapter 4, I evaluate the cost effectiveness of different regulations for reducing aggregate impacts from surface infrastructure across sites and find large gains from trade when implementing a cap and trade system. Overall, my dissertation facilitates a transition of knowledge for conservation planning to be able to better adapt to and cope with the changing world.

Table of Contents

List of Tables

Main Tables

Appendix Tables

List of Figures

Main Figures

Appendix Figures

Introduction

Background and Motivation

The world is changing. Change is not new, but the nature of this change is. Some aspects of this change include new threats to conservation priorities, an increasingly global scope for problems and decisions, the integration of conservation practice and human development (Sutherland *et al.* 2009; Rands *et al.* 2010), the increasing availability of large datasets (Hampton *et al.* 2013), and increasing computational power. These changes point to a need for conservation planning as a science and practice to keep pace by updating.

Conservation resources are limited relative to biodiversity needs (Waldron *et al.* 2013), which increases the need for new science to inform conservation practice now, not in decades when it may be too late. Sunderland *et al.* (2009) argue that the well-accepted gap between conservation science and practice can be partially ameliorated by increasing the accessibility and relevance of science to practitioners. One way this has been done in the past decades has been to incorporate limited budgets into conservation planning (e.g a few from the past year: Bode *et al.* 2015; Lentini & Wintle 2015; Boyd, Epanchin-Niell & Siikamäki 2015). However, conservation can only benefit by further increasing our commitment to immediately applicable science. Further, it is increasingly recognized that in order to be relevant, science needs to reflect the realities of conservation practice (Salafsky *et al.* 2002). This means designing science that matches the scales, process, data, and assumptions of conservation practitioners, sometimes at a cost of theoretically optimal outcomes.

Increasingly common are the availability and use of large datasets for science and practice (Hampton *et al.* 2013), the format and methods of use of which are important. For instance, the use of range maps or species distribution models may lead to different planning decisions from the direct use of presence only point occurrence data (Wilson *et al.* 2005; Rondinini *et al.* 2006). In conservation practice it is common to use point occurrences in isolation or with modeled distributions. Similarly, the choice of planning method may significantly affect decisions (Rondinini *et al.* 2006; Lentini & Wintle 2015).

We have an increasing need and ability to make decisions at larger spatial and temporal scales. However, larger scale conservation planning does not obviate the need to think about smaller scales at high resolutions. The focus of studies at different scales should be proportional to the needs of decision makers operating at those scales.

One aspect of the changing world to be addressed is how the ongoing collection of biodiversity data is influencing spatial conservation priorities. Spatial prioritization is prominent in the science (Brooks *et al.* 2006; Pressey *et al.* 2007; Watson *et al.* 2011; Coll *et al.* 2012; Rainho & Palmeirim 2013) and practice (Redford *et al.* 2003; Didier *et al.* 2009; Henson *et al.* 2009; Groves *et al.* 2012) of conservation. This is due to the long history of spatial conservation planning methods (Diamond 1975; Margules & Pressey 2000; Watson *et al.* 2011), to the recognition that conservation funding is limited (Ando 1998; Myers *et al.* 2000; Bottrill *et al.* 2008) and to the spatial nature of many conservation decisions (Pressey *et al.* 2007). Although several studies exist that look at the influence of data characteristics on spatial conservation priorities (Freitag & van Jaarsveld 1998; Polasky & Solow 2001; Gladstone & Davis 2003; Gaston & Rodrigues 2003; Grantham *et al.* 2008; De Ornellas, Milner-Gulland & Nicholson 2011), they tend to use snapshots in time or focus on data weaknesses rather than on the actual use of those datasets over long time periods. Doing this allows authors to cover a range of data collection scenarios in order to generalize across many situations that might be encountered. Equally, however, an argument can be made for the analysis of longer-term datasets that maintain the autocorrelated structure of the data over time and reflect how datasets would be used in practice.

Another aspect of the changing world is the integration of conservation and development. This change to both activities has some advantages. Development will continue in places of high conservation value, so it is in the interest of conservation stakeholders to work with developers. Also, since conservation is partially anthropocentric, it makes sense to integrate land use decisions to incorporate multiple competing objectives, since only by simultaneous planning can tradeoffs be minimized (Naidoo *et al.* 2008).

One especially noteworthy example of a context that can benefit from the integration of conservation and development is shale gas in Appalachia. Shale gas development has been an increasing source of environmental and human health concerns in recent years. In the United States, shale gas production has increased steadily over the past decades and now makes up ~40% of gas production (*Annual Energy Outlook 2014* 2014). Concerns have been raised about the environmental (Gillen & Kiviat 2012; Kiviat 2013; Olmstead *et al.* 2013; Jones *et al.* 2014) and human health (Perry 2012) effects of shale energy production, leading to careful consideration of how to protect society and nature from those effects (Howarth, Ingraffea & Engelder 2011; Hays *et al.* 2015) and at times outright bans on development.

Terrestrial impacts resulting from the spatial locations of gas surface infrastructure are an understudied and important issue. Impacts occur at all stages of the development process, from pre-production through post-production (Burton *et al.* 2014). Here I focus exclusively on pre-production activities and specifically the construction of well pads, access roads, and gathering pipelines, the impacts of which may play out at different rates and spatial extents. The magnitudes and types of impacts change from stage to stage, but it is clear that site construction incurs the most direct and thus quantifiable land use changes. Below, I describe some of the impacts resulting from shale gas surface infrastructure.

In much of the dissertation, I examine how development for shale gas, a new kind of threat to species and habitats, can be reconfigured to reduce its potential environmental impacts while still allowing development to proceed. In doing so, my research builds on a rich history of conservation biology that focuses on mitigating and moderating threats from particular industries. For example, my research is thematically similar to past work on bycatch reduction measures in fisheries (Crowder & Murawski 1998), especially with regards to the protection of loggerhead sea turtle populations (Crouse, Crowder & Caswell 1987; Crowder *et al.* 1994; Lewison, Freeman & Crowder 2004), sea birds, and marine mammals (Cox *et al.* 2007). Another prominent theme in the fisheries literature comes from observations of the destruction of seafloors by trawling gear spurred further studies (Graham 1955; Caddy 1973; Wenner 1983; Jones 1992). Examples exist from other sectors as well. In agriculture, for instance, past studies have focused on understanding the effects of different practices on bird diversity with an eye on how to improve diversity around cropland (Owens & Myres 1973; Best 1983). In forestry, some research has been concerned with how to best measure the biodiversity outcomes of different forest management practices (Lindenmayer, Margules & Botkin 2000; Lindenmayer, Franklin & Fischer 2006), whether altered or managed forests can provide sufficient biodiversity benefits (Hansen *et al.* 1991), and the ecological effects of different harvesting patterns (Franklin & Forman 1987).

The Appalachian mountains are highly diverse in species (Stein, Kutner & Adams 2000), many of which are endemic to the region and/or sensitive to changes in the environment. As such, the negative environmental effects of gas infrastructure development are likely to impact many species and their habitats. Perhaps the most immediate and clear expected impact from gas development is the displacement of rare and vulnerable species. Gas development is occurring in fairly pristine forests (Drohan *et al.* 2012) and other important habitats, meaning that there is a high chance that infrastructure will destroy the sensitive habitats or individuals directly in its path.

At the landscape scale, gas infrastructure disrupts forests, wetlands, and other habitats. Habitat fragmentation is well studied in ecology and conservation (Margules, Milkovits & Smith 1994; Ranta *et al.* 1998; Didham *et al.* 1998; Davies, Margules & Lawrence 2000; Jaeger 2000; Fahrig 2003). The major fragmenting effects of gas infrastructure act through reduction of core habitat, creation of edges, and reduction of connectivity. Well pads and their associated infrastructure are regularly spaced on the landscape, but with little distance between them. This spatial pattern maximally fragments the landscape by creating many long edges in core forests and other core habitats (Drohan *et al.* 2012).

Edges change the light, humidity, wind, sound, temperature, and other factors at the edge of forests and into their interiors (Saunders, Hobbs & Margules 1991; Matlack 1993; Forman & Alexander 1998; Haskell 2000). These edge effects may increase stress on organisms at the habitat edge (Burke & Nol 1998; Gibbs 1998; Lehtinen, Galatowitsch & Tester 1999), may facilitate invaders (Cadenasso & Pickett 2001; Watkins *et al.* 2003; Pauchard & Alaback 2006), and will alter the future species composition of communities at the edge and in the interior (García-Tejero *et al.* 2013).

In addition to creating edges, gas infrastructure disturbances reduce habitat connectivity by creating barriers to dispersal and movement. Some species avoid crossing roads (Rico, Kindlmann & Sedláček 2007; Shepard *et al.* 2008), and we might expect similar responses for well pads and gathering pipelines. When roads or pipelines cross streams, they change stream flows and may prevent movement up- or down-stream (Pépino, Rodríguez & Magnan 2012). These connectivity issues are especially important for Appalachian species when considered in the context of climate change, since this region is projected to be very important for species migrating to track changing conditions (Lawler *et al.* 2013).

Infrastructure construction increases soil erosion, which has negative ecological consequences. Erosion may increase sediment loads in streams and consequently affect the turbidity, light absorption, chemistry and temperature of the water column (Reid & Dunne 1984; Lane & Sheridan 2002). When sediment settles, it can alter the stream substrate (Boxall & Maltby 1995). Many species are sensitive to small changes in water and substrate quality (Curry & MacNeill 2004; Cover *et al.* 2008). Gas development tends to be concentrated in time and space, e.g. many well pads, roads, and pipelines are constructed simultaneously (*pers. observation*), which will enhance short-term erosion impacts and could lead to long-term changes in stream biodiversity. Nearer to infrastructure, erosion changes soil qualities (Verity & Anderson 1990; Ni & Zhang 2007) and may affect which species move in to the disturbed area after construction is complete.

The combination of drilling technology, energy demand, and desire for energy independence means that domestic gas development will continue in the Appalachian region and more broadly for some time. Since we cannot prevent all gas development, our best chance to avoid some of the discussed ecological impacts is to alter future development through improved siting practices.

Chapter Summary

My dissertation is composed of four chapters which aim to advance the science and practice of conservation planning by providing new solutions to real world conservation problems as well as concrete recommendations for how to use my results. The chapters of my dissertation are motivated by four questions:

Chapter 1: How do new surveys of rare species change our conservation priorities?

In Chapter 1, I use a long-duration time series of rare-species surveys to see how a growing dataset changes rankings of watersheds in Tennessee. The chapter provides some insight into how data used by conservation practitioners may affect their decisions in data poor, data rich, and near-future contexts.

Chapter 2: What are the environmental impact tradeoffs between four easy-to-implement natural gas surface infrastructure siting guidelines?

In Chapter 2, I compare negative effects of realistic gas infrastructure layouts that I create using four siting practices and examine tradeoffs between impacts within and across siting practices. The results inform how one strategy used by conservation practitioners – rules of thumb – to affect gas developer behavior may introduce tradeoffs in future negative impacts. The results also point to a need to explore how more advanced conservation planning tools may improve the performance of surface infrastructure by simultaneously optimizing infrastructure locations to reduce potential impacts.

Chapter 3: What is the cost of reducing environmental impacts from surface infrastructure at the lease-hold scale?

In Chapter 3, I assess the magnitudes of avoidable impacts from surface infrastructure in Pennsylvania by using advanced spatial planning software I created for this task to plan well pad, access road, and gathering pipeline locations with environmental objectives and monetary constraints. I find that the cost of reducing impacts varies considerably across sites, and while impacts at a median site can be reduced upwards of 40% before costs become prohibitive, a uniform policy applied to this context may not produce desirable outcomes. The results of this chapter indicate that larger scale actions to reduce environmental impacts from shale gas surface infrastructure will perform best when accounting for heterogeneity across sites.

Chapter 4: How cost-effective are different regulations for reducing environmental impacts from surface infrastructure?

In my final chapter, I assess the cost effectiveness of multiple policy options for reducing aggregate impacts across the same set of sites explored in Chapter 3. I find that a typical uniform, inflexible approach similar to the median site approach mentioned above would lead to much higher system-wide costs for a given level of impact avoidance as compared to a market based, tradable permits approach. These results provide inside for several decision makers, but especially policy makers who are tasked with formulating and implementing conservation-oriented regulations for shale gas development.

Chapter 1: Updating conservation priorities over 111 years of species observations

A version of this chapter was originally published by Austin W. Milt, Sally R. Palmer, and Paul R. Armsworth:

Austin W. Milt, Sally R. Palmer, and Paul R. Armsworth. "Updating conservation priorities over 111 years of species observations." *Journal of Applied Ecology* 51 (2014): 1515-1524.

Austin Milt performed the secondary data collection, data processing, analysis, interpretation, and writing for this article. Sally Palmer and Paul Armsworth contributed intellectually to the design, interpretation, and revision of the article.

1.1 Abstract

Observations of species occurrences are often used to inform spatial prioritizations for the effective use of limited conservation resources. Additional species observations have the potential to change where a conservation group plans to invest. But by how much? How different would conservation priorities be if planners updated current observations with the information they will have next year? We sought to address these questions using a 111 year dataset that reflects commonly used collection and prioritization practices. We quantify changes in the ranking of Tennessee watersheds brought on by annual additions of species observations made between 1900 and 2010. We ranked watersheds by their complementary contribution to overall species richness. We examine the sensitivity of our results to the number of watersheds prioritized. We expected the effect of new observations to diminish as the dataset grew, and we found this to be the case. Importantly, however, additional observations may continue to significantly change conservation priorities in the future if current data collection trends continue. We found that, overall, additional observations can greatly affect priorities and that this result is sensitive to the number of watersheds ranked. Thus the extent of planning activities moderates the effect of including additional data. *Synthesis and applications:* Long-term, opportunistically collected data of species locations are commonly used in conservation planning. We find that when using such data additional species observations significantly affect subsequent priorities. This effect is most pronounced when data are sparse. As such, data collection should be a focus of very early conservation actions in new areas. Even in well-studied areas, however, additional observations may continue to change spatial priorities into the future, and so while data collection can decrease in well studied areas, it should continue at a lower intensity. Our methods could also be used to

determine the balance of data collection and conservation action in a new location.

1.2 Introduction

Spatial prioritization is prominent in the science (Brooks *et al.* 2006; Pressey *et al.* 2007; Watson *et al.* 2011; Coll *et al.* 2012; Rainho & Palmeirim 2013) and practice (Redford *et al.* 2003; Didier *et al.* 2009; Henson *et al.* 2009; Groves *et al.* 2012) of conservation. This is due to the long history of spatial conservation planning methods (Diamond 1975; Margules & Pressey 2000; Watson *et al.* 2011), to the recognition that conservation funding is limited (Ando 1998; Myers *et al.* 2000; Bottrill *et al.* 2008) and to the spatial nature of many conservation decisions (Pressey *et al.* 2007). Prioritization methods vary depending on the conservation goals, expertise and data available to planners. Priorities may be determined by local species richness or biodiversity uniqueness (Csuti *et al.* 1997; Myers *et al.* 2000), by metrics of threat (Pressey *et al.* 2007; Joseph, Maloney & Possingham 2009; Carwardine *et al.* 2012) or many other factors. This variation in method, along with variation in the data used in a particular evaluation, can lead to very different decisions about where, when and how to take action (Wilson *et al.* 2005; Rondinini *et al.* 2006).

Data on species occurrences are commonly used in conservation prioritization (e.g. Zafra-Calvo *et al.* 2010; Simaika *et al.* 2013; Mateo *et al.* 2013). Often, such datasets change in extent, resolution, accuracy, and coverage as more observations are added (e.g. Magurran *et al.* 2010; Ahrends *et al.* 2011; Felinks *et al.* 2011; Martin, Blossey & Ellis 2012). Spatial priorities will be affected by data characteristics, such as spatial resolution (Araujo *et al.* 2005; Arponen *et al.* 2012), type (e.g. presence/absence data vs. abundance data; Gaston & Rodrigues 2003) and bias (De Ornellas, Milner-Gulland & Nicholson 2011; Metcalfe *et al.* 2013). Past studies looking at the effect of changing data on the outcome of conservation planning have tended to stylize the spatio-temporal extent and resolution of data used in conservation (Freitag & van Jaarsveld 1998; Polasky & Solow 2001; Felinks *et al.* 2011) and ignore the somewhat opportunistic nature of data being used by many conservation practitioners.

In this paper, we examine how additional species occurrence records affect spatial conservation priorities. In so doing, we focus on species-centric conservation approaches, as opposed to focusing on conservation goals targeting priority habitats or whole ecoregions (Watson *et al.* 2011). Tennessee, which we use as a case study, is a centre of richness for freshwater fish species and molluscs and a region within the coterminous United States of particularly high species imperilment (Dobson *et al.* 1997; Stein, Kutner & Adams 2000). We examine how annual additions of species observations made from 1900 to 2010 would change the ranking of watersheds being prioritized for conservation action. Specifically, we rank watersheds by complementary richness. As a conservation objective, complementary richness rewards watersheds for covering species not found in other protected watersheds (Vane-Wright, Humphries & Williams 1991). We use a rank correlation statistic to quantify the change in priorities brought on by an additional year's observations. Further, we assess the magnitude, trend, and consistency of priority changes over time. We examine the sensitivity of our results to the number of watersheds prioritized. In the Appendix, we also explore the sensitivity of our results to ranking method, spatial or taxonomic sampling bias, and changes in data reliability due to changing technology and organism or population persistence.

The dataset we use is one currently used by conservation planners in Tennessee. As is often the case with datasets built from historical occurrence records, this one has been collated in a piecemeal and somewhat opportunistic fashion. As a result, the dataset suffers from more spatial, temporal, and taxonomic sampling bias than in systematic surveys. Arguably, it still represents the best information available to conservation planners regarding the distribution of priority species in Tennessee today.

Previous studies with similar methodologies to ours focus on the effectiveness of conservation outcomes under different data quality and quantity scenarios (Freitag & van Jaarsveld 1998; Polasky & Solow 2001; Gladstone & Davis 2003; Gaston & Rodrigues 2003; Grantham *et al.* 2008; De Ornellas, Milner-Gulland & Nicholson 2011). The general approach in empirical studies has been to aggregate data over time, simulate changes to the data (e.g. by subsampling to represent reduced sampling effort), and to evaluate conservation plans on the altered dataset (but see Felinks *et al.* 2011). Doing this allows authors to cover a range of data collection scenarios in order to generalize across many situations that might be encountered. These studies conclude that data quality and quantity are important factors in taking effective conservation actions, but the details are dataset specific. For example, Grantham *et al.* (2009) assess how switching from initial species surveys to habitat protection affects the long-term coverage and retention of proteas. They find that for their case study area, a shorter duration of surveying (~2 years) followed by longer protection is optimal. Their study has important implications for conservation planning since it indicates that long-term data collection need not preclude conservation actions.

We complement previous work in many ways. First, we use a much longer-term dataset spanning 111 years. Second, we build the dataset sequentially over the time period rather than subsampling without regard to time and thus we follow the actual collection of species observations. As such, our analysis includes the co-variation of data characteristics over time. Third, our analyses do not rely on the aggregate dataset as the most accurate knowledge of species distributions over time. Rather, we focus on describing how changing knowledge over time affects conservation priorities. Finally, we use raw species observations as they were recorded rather than modelled data or controls for data biases. We briefly explore how reducing among-watershed sampling bias affects our results (see Appendix S1).

1.3 Materials and Methods

1.3.1 Case Study Area

Tennessee is one of the most biodiverse inland states in the U.S., second only to Alabama in the diversity of freshwater fishes and possessing a comparatively high degree of species endemism (Stein 2002). Over 10% of the state's plant and animal species are considered at-risk, and Tennessee ranks seventh among all states in the number of documented extinctions, a fact largely attributed to the major modification of streams and river systems in the early to mid-20th century (Stein, Kutner & Adams 2000; Stein 2002). Widespread conversion of lands for agricultural purposes has also contributed to fundamental changes in hydrologic regimes in many sub-regions of the state, and excess nutrients and sedimentation from agricultural production contribute to degraded water and habitat quality (Tennessee Department of Environment and Conservation 2014). Increased urbanization within the state's Metropolitan Statistical Areas has resulted in destruction and fragmentation of terrestrial habitats and degradation of streams and wetlands.

Local, national and international conservation organizations such as The Nature Conservancy and the World Wildlife Fund have invested in Tennessee for over 35 years in collaboration with many partners, including federal agencies such as the U.S. Fish and Wildlife Service and state agencies such as the Tennessee Natural Heritage Program and the Tennessee Wildlife Resources Agency. Foundational to this work have been a series of conservation plans designed at ecoregional scales and using species occurrence data to set biodiversity conservation goals (Smith *et al.* 2002; The Nature Conservancy 2006). Beginning in 2005, all state wildlife agencies receiving federal State Wildlife Grant funding were required to submit a Comprehensive Wildlife Conservation Strategy, more commonly known as a State Wildlife Action Plan. The primary emphasis of State Wildlife Action Plans is to improve the habitat and population conditions of "species of greatest conservation need" as defined by the state (Tennessee Wildlife Resources Agency 2005). Designing and executing these plans has resulted in an increased emphasis on the use of field-collected species occurrence and habitat data to identify priority conservation geographies and assess threats to these areas.

1.3.2 Prioritization Data

We used species observation data collected between 1900 and 2010 by Tennessee Natural Heritage Program to test how species observations affect conservation priorities. The Natural Heritage dataset is used in multiple forms of conservation planning in Tennessee, and the State Wildlife Action Plan in particular (Tennessee Wildlife Resources Agency 2005). The 2005 Tennessee State Wildlife Action Plan used these species occurrence data in combination with NatureServe global and state rarity rankings, U.S. Fish and Wildlife Service federal status listings, and other available population status data to assign "species of greatest conservation need" status (SGCN). The species occurrence data have been used as a key component in mapping local richness to understand where high SGCN concentrations occur across the state. These same data have been used to assess the complementarity of larger ecological units for terrestrial and freshwater species.

The Natural Heritage data are opportunistically recorded point observations (EOs) of individual species, most of which have a NatureServe Conservation Status rank higher than S3, and some of which are regularly monitored (TN Natural Heritage Program, *pers. comm.*). Rarely, observations are made as a result of premeditated prediction and collection efforts (TN Natural Heritage Program, *pers. comm.)*. Because of the nature of how these data have been collected, there is no measure of sampling effort embedded in the dataset. The dataset contains both unique and repeat observations. A unique EO represents the spatial location, species identity, and date of an observation of an individual or group. Subsequent observations of the same individual or group are here called repeat EOs. The dataset contains 17 586 unique EOs or 25 838 EOs including repeats. Fig. 1.1 shows the number of EOs recorded in each year including repeats. The dataset contains both terrestrial and aquatic species. The majority of EOs recorded in the dataset are of plants (15 001 records or 58%), although the dataset also represents species from 14 other taxonomic groups recognized under the State Wildlife Action Plan. Approximately 51% of EOs were recorded after 1995.

Conservation organizations often make use of a mix of raw point

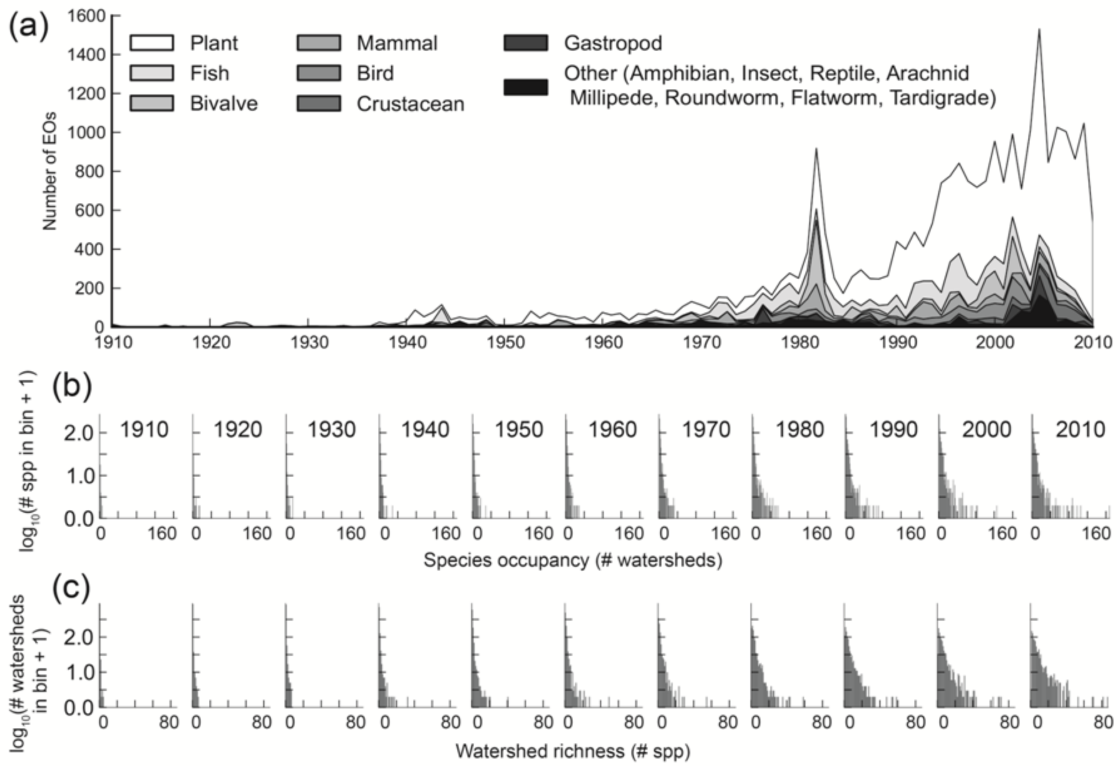


Fig. 1.1. Distribution of element occurrences (EOs) across watersheds and species over the 111 year period; x-axis is year.(a) Number of new EOs recorded throughout Tennessee each year broken down by taxonomic groups. An EO is an observed point location of an individual or population. There are 25,838 EOs in total, with the majority represented by Plant (15,001 or 58%). (b) and (c) show distributions as histograms of (b) species occupancies in Tennessee watersheds and (c) watershed richnesses at 11 time periods.

observations and interpolated layers or modelled distributions (NatureServe, pers. comm.; The Nature Conservancy 2006; Schloss *et al.* 2011; Wilson 2011; Yorkshire Wildlife Trust, pers. comm.). Modelled data, such as those produced by species distribution models, are valuable because they estimate the unobserved range of a species and thus point out potential high-value areas not revealed by raw point occurrences. Also, species distribution models incur fewer omission errors and may reduce the effects of spatial sampling bias on conservation plans (Rondinini *et al.* 2006). On the other hand, raw point observations are simple to use and do not suffer as much from commission errors as species distribution models. Moreover, species distribution models cannot accurately estimate the ranges of very rare or under-sampled species (Olden, Jackson & Peres-Neto 2002; Wisz *et al.* 2008), a particular problem for our dataset and similar contexts, because most species occur fewer than five times in the data (Fig. 1.1b, year 2010).

We used watersheds from the US Geological Survey HUC-12 Watershed Boundary Dataset (http://datagateway.nrcs.usda.gov, accessed 09 May 2013) as our spatial unit of analysis. Watersheds are an appropriate unit for spatial prioritization in conservation at the state level when focusing on terrestrial and aquatic species. At this scale, planners can target particular watersheds for further action. This action may come in the form of whole-watershed management (e.g. best practices by all farmers in the watershed), or it may call for more refined analysis to target protection within the watershed (e.g. protecting stream headwaters through forested-land conservation). For instance, watersheds have been used to delineate conservation priorities for known occurrences of freshwater species in the south-eastern United States (Smith et al. 2002) and elsewhere (Pryce et al. 2006). Watershed boundaries are also less changeable than other spatial units like land parcel boundaries and are therefore fitting for our century-spanning analysis. Of the 1152 watersheds in Tennessee, 925 have at least one observation by the Natural Heritage Program by 2010. These 925 watersheds (Fig. 1.2) acted as our candidate sites for selection and have a median area of 101 km^2 (1st quartile = 77 km^2, 3rd quartile = 142 km^2).

1.3.3 Ranking Watersheds

We explored the situation in which a conservation group aims to cover as many species across the combined set of priority watersheds, rather than only the most species rich watersheds. This requires consideration of how the species assemblages of watersheds complement one another. To implement this, we ranked watersheds based on their

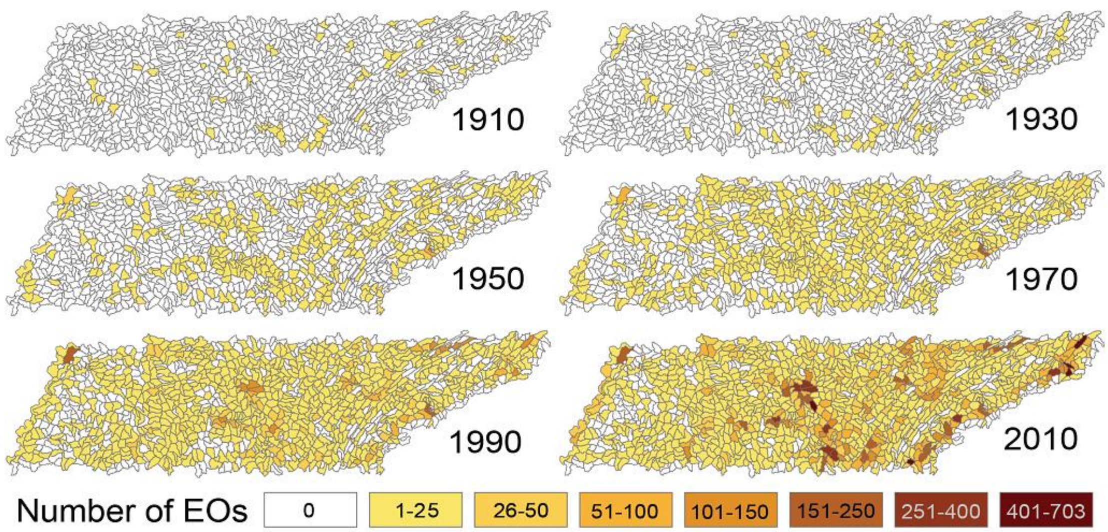


Fig. 1.2. Number of EOs, including repeats, accumulated in each watershed since 1900 at six snapshots in time. Darker colors have a higher density of EOs. For example, the darkest watershed in 1970 has 155 EOs which it accumulated between 1900 and 1970. Of the 1152 watersheds, 925 have at least one EO by 2010, 695 have ≤ 25 and 56 have ≥ 100.

frequency in near-optimal solutions to the maximal coverage problem (MCP; Cabeza & Moilanen 2001). Given a watershed budget (*b*), the globally optimum solution to the MCP is the set of *b* watersheds that together cover more species than any other set of *b* watersheds. Solving the MCP gives a set of watersheds that together perform well as a conservation strategy. It does not automatically give a means to rank individual watersheds. However, Pressey, Johnson & Wilson (1994) introduced the idea of irreplaceability, defined as "the frequency of occurrence of individual [watersheds] in the range of possible representative systems." Irreplaceable watersheds are those that have a high potential to contribute to the conservation goal under many realized priority sets. We drew on this concept when deciding how to rank complementary watersheds.

The algorithm we used to rank watersheds has two parts. In the first part, we used a genetic algorithm optimizer to choose one set of *b* watersheds that maximizes the number of species covered. The difficulty of the MCP means that the genetic algorithm optimizer guarantees, at worst, locally optimum solutions. In conservation the local optimality of solutions can be a strength, because finding many near-optimal solutions rather than the one best solution lends flexibility to the decision making process. For a given budget, we ran the genetic algorithm optimizer 500 times and kept those solutions that achieved ≥ 95% of the richness of the best solution. In the second part, we ranked watersheds by the number of times they appeared across those top-scoring solutions (irreplaceability). In total, we examined 11 sensitivity tests corresponding to 11 watershed budgets (*b*): 1, 5, 10, 20, …, and 90 watersheds.

In Appendix S1 we also explore the sensitivity of our results to other ranking methods. Namely, we assess prioritization based on the local richness of watersheds when ignoring complementarity in order to explore how sensitive our results are to the particular choice of conservation objective that we examine. We also assess variations on this local richness case to the number of watersheds prioritized, controls for data reliability and a control for sampling bias. All ranking was carried out in Python v2.5.4.

1.3.4 Measuring Changes in Priorities

We measured by how much additional species observations cause watershed rankings to change. We used Spearman's rank correlation statistic, ρ, to measure the similarity of two different rankings of candidate watersheds for conservation, where the ranking of watersheds in a year is based on the observed assemblages of species in those watersheds. Our measure of priority change (V) is the difference between rankings, or

$$V_t = 1 - \rho(\text{ranking in year } t, \text{ranking in year } t - 1)$$

We subtract the rank correlation from 1 to ease explanation such that larger values of V correspond to larger changes in priorities. Because ρ can take values between -1 and 1, V can vary between 0 and 2. We would typically expect V to take values between zero – where the rankings are identical – and one – where the two rankings have no relation to one another.

Below, we describe the entire process of ranking watersheds and calculating V with a budget of 10 watersheds:

1. **Initialize**: Using data from 1900–1909, the genetic algorithm optimizer chooses 10 watersheds that maximally cover present species. This is repeated 500 times. Watersheds are ranked by their frequency in the top 95% of the 500 solutions.

2. **Update Records**: Add records from the next year. In the first iteration, we added records from 1910 so that the irreplaceability and ranking of watersheds in 1910 was based on EOs from 1900-1910. The ranking in the second iteration (1911) was based on EOs from 1900-1911, and so on.

3. **Rank Watersheds**: Repeat Step 1 with the updated dataset from Step 2.

4. **Calculate V_t:** Calculate the difference between the rankings between the current year and previous year as defined above (1 - Spearman's ρ). This gives the magnitude of change in priorities for the current year. In the first iteration, we get V_{1910} by comparing the rankings from 1910 and 1909. In Fig. S 1.2, we summarize the result of delaying updates of records and priorities.

5. **Repeat**: Repeat Steps 2–4 through the year 2010, adding the most recent records for each iteration such that

   V_{1910} = 1-ρ(ranking in 1910, ranking in 1909)
   V_{1911} = 1-ρ(ranking in 1911, ranking in 1910)
   ⋮
   V_{2010} = 1-ρ(ranking in 2010, ranking in 2009)

1.3.5 Statistical Analyses

To test if additional data have a significant effect on priorities, we tested whether the 5% confidence limit (5% CL) about the median of *V* contained zero. To calculate the 5% CL, we used bias-corrected accelerated bootstrapping from 10 000 samples of the same size as the original sample (usually 101 data points, one from each year within 1910–2010). Bootstrapping was performed in MATLAB r2012b.

We examined trends in priority changes over time or dataset size using ordinary least squares regression. In this dataset, time and $\log_{10}$(dataset size) are highly collinear ($R^2 = 0.99$), indicating they cannot be included in the same regression (Quinn & Keough 2002). Therefore, we regressed priority changes against time and dataset size separately. Due to the similarity of results when regressing against time or dataset size, we focus on regressions in time in the main text (but see Table S 1.3 for dataset size results). We expected changes in priorities to decrease as we accumulated data. To confirm this, we tested that the 95% confidence intervals of the slopes of the above regressions are negative.

For some of the watershed budgets, we observed that the data could be clearly separated into two distinct segments (e.g. Fig. 1.3a,b). Thus, we created piecewise regressions using the *segmented* package in R v2.12.1 to further examine the results (Muggeo 2008). The piecewise regression optimization is sensitive to initial guesses of breakpoints, so we visually estimated the breakpoints and then used the piecewise model with the convergence closest to our estimations (Muggeo 2008). For each sensitivity test, we tested piecewise models with one or two segments and compared AICc scores to determine which offered the more parsimonious fit to the data. AICc model comparison explicitly considers the trade-off between model fit using maximum likelihood and parsimony through the number of parameters (Crawley 2007).

We also wanted to know if changes in priorities might be sustained over the near future. This was primarily determined by interpolating the predicted value of *V* and its 5% significance in 2010 using the one or two-segment model chosen by AICc competition. We also extrapolated to 2030 to assess a more distant level of change if current conditions hold.

1.4 Results

We found that spatial conservation priorities are generally sensitive to additional data and the number of watersheds in which conservation can take place (Table 1.1: column 2). We also found that the sensitivity of priorities decreased over time (Table 1.1: columns 4,5; Fig. 1.3a,b), but that the trend levelled out after some time and the point at which this occurred depended on the watershed budget (Table 1.1: columns 4,5,6;

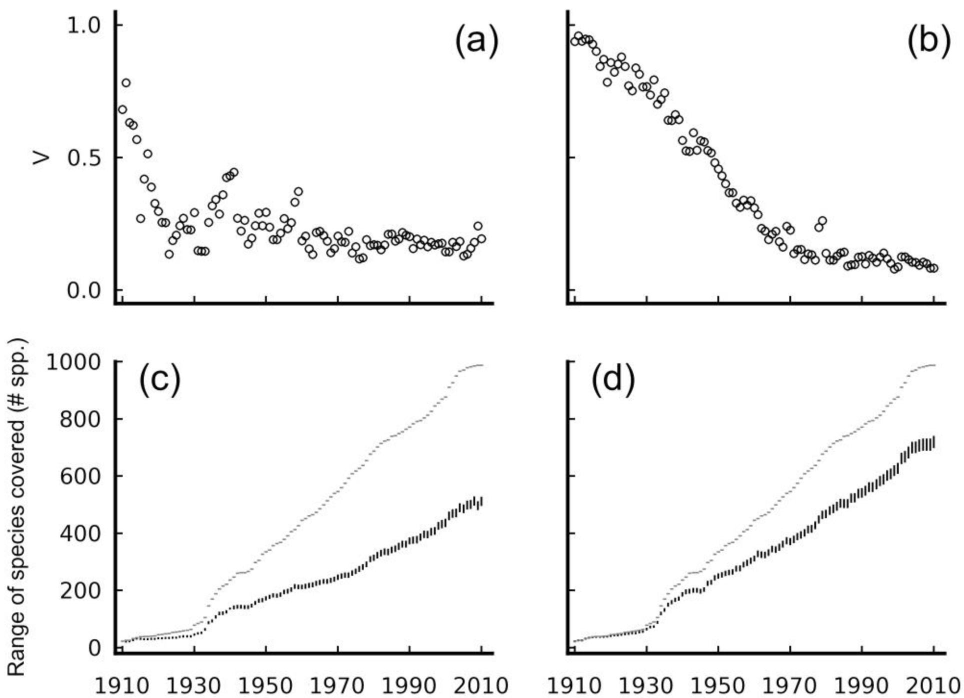


Fig. 1.3. Representative results of ranking watersheds by their frequency in near-optimal solutions to the Maximal Coverage Problem (MCP; Cabeza & Moilanen 2001). Solutions to the MCP choose a fixed subset of all watersheds that maximally cover known species. Conservation priorities are sensitive to additional data, but that sensitivity declines with a smaller watershed budget and over time and eventually levels out. Panels (a) and (c) correspond to rankings created by choosing 20 watersheds to solve the MCP; panels (b) and (d) correspond to a 70-watershed solution. (a), (b) effect of one year of additional element occurrences, from 1910 to 2010. (c), (d) range of species covered by solutions to the MCP. Bottoms and tops of vertical, lower, black bars are minimum and maximum number of species covered by solutions to the MCP. Horizontal, higher grey bars are the number of species with at least one occurrence by that year.

Fig. 1.3a,b). Finally, we found that additional data may continue to affect priorities in the future if current collection conditions hold (Table 1.1: columns 9,10).

Taking two exemplar budgets from those shown in Table 1.1 for illustrative purposes (20 watersheds and 70 watersheds), Fig. 1.3a,b illustrates the magnitude of change in priorities over time and Fig. 3c,d shows the number of species covered over time indicating the performance of the conservation planning process (see also Fig. S 1.3).In Fig. 1.3c,d, the lower set of vertical, black bars shows the numbers of species covered by solutions to the MCP over time. Compare this to the higher, horizontal, grey bars, which show how many species were known to occur in Tennessee in that year. The height of the lower bars relative to the higher bars shows how well the prioritization performed in each year. A comparison between Fig. 1.3c,d shows that a watershed budget of 70 led to greater coverage of known species than using a lower budget of 20 watersheds. We also assess the performance of prioritizations relative to the full dataset in Fig. S 1.3.

In the first part of our analysis, we wanted to know if, in general, one year of additional EOs affects conservation priorities. Provided more than one watershed is being considered for conservation action, we found that one year of additional EOs significantly changed the ranking of watersheds (Table 1.1: second column). Values in the second column of Table 1.1 reveal the magnitude of overall changes in priorities brought on by additional EOs, with larger values indicating more change per year of additional EOs. Note that the overall magnitude of change brought on by additional EOs increases with the watershed budget (moving down second column in Table 1.1).

Next, we assessed how changes in priorities changed over time. Changes in priorities over time were pronounced (Fig. 1.3a,b, Table 1.1, Fig. S 1.4 maps the spatial changes in priorities through time). The exemplar budgets shown in Fig. 1.3a,b reveal how large changes in priorities near the beginning of the time period are followed by a steep decline earlier and steadier decline later, but differ in how steep that initial decline is and when it occurs. These differences are quantified in Table 1.1. The fourth and fifth columns in Table 1.1 show the main results of our analysis of trends in changing priorities. In eight sensitivity tests, the degree to which priorities changed with one year of additional EOs decreased over time (Table 1.1: column 4). In seven of those eight tests, this result was true in both the first and second time periods (Table 1.1: columns 4,5).

We also explored in more detail how trends in changing priorities differed between the first and second time periods in each sensitivity test. We did this by comparing the values in the fourth and fifth columns of Table 1.1 and illustrate this in Fig. 1.3a,b. There was support in most

Table 1.1. Effect of additional data on priorities overall and when regressed against time. Conservation priorities are sensitive to additional data, but that sensitivity declines with a smaller watershed budget and over time and may level out in the future. Sensitivity of the ranking method to the number of sites used in the prioritization was tested. Subsequently, (column 2) the overall magnitude of change brought on by one year of additional element occurrences (EOs) was tested, as well as (columns 3-8) the trend of changing priorities over time. Finally, the regression models were used to predict the change in priorities from additional EOs in (column 9) 2010 and (column 10) 2030 if current collection conditions hold. Sensitivity tests without [$slope_2$] and [break] fields use a one-segment regression

test	med(V)	intercept	$slope_1$	$slope_2$	break	R^2	ΔAICc	V_{2010}	V_{2030}
1-watershed	0	-	-	-	-	-	-	-	-
5-watersheds	**0.14**	-1.0	0.6	-	-	0.03	2	**0.17**	**0.18**
10-watersheds	**0.18**	0.2	0.6	-	-	-0.01	1	**0.19**	**0.19**
20-watersheds	**0.20**	84.6	**-43.9**	**-1.3**	1920	0.73	83	**0.16**	**0.13**
30-watersheds	**0.19**	40.4	**-20.7**	**-1.5**	1937	0.88	118	**0.13**	0.10
40-watersheds	**0.22**	37.9	**-19.4**	**-1.3**	1944	0.94	165	**0.14**	**0.12**
50-watersheds	**0.21**	35.0	**-17.8**	**-0.9**	1954	0.97	199	**0.14**	**0.12**
60-watersheds	**0.20**	29.8	**-15.1**	-0.1	1965	0.98	203	**0.14**	**0.14**
70-watersheds	**0.31**	27.6	**-13.9**	**-1.7**	1971	0.98	177	**0.09**	0.06
80-watersheds	**0.37**	24.9	**-12.5**	**-2.0**	1977	0.99	173	**0.08**	0.04
90-watersheds	**0.40**	23.8	**-12.0**	**-2.3**	1978	0.99	146	**0.09**	0.05

Table columns are [med(V)] = median of V across all years; [intercept] = model intercept; [$slope_{1(2)}$] = slope of the first (second) segment in $V \cdot yr^{-1} \cdot 10^{-3}$; [break] = breakpoint (year) of the two segments; [R^2] = adjusted R^2; [ΔAICc] = ΔAICc of the piecewise model not chosen.; [V_{2010}] = predicted value of V in 2010; [V_{2030}] = predicted value of V in 2030

Values in **bold** are significant at the 95% confidence level.

cases for distinguishing the two time periods evident in each sensitivity test, as can be seen in Fig. 1.3a,b and tested by AICc competition in column eight in Table 1.1. When using a watershed budget larger than 10, there was a larger negative trend in changing priorities in the earlier time period than the later (Fig. 1.3a,b, Table 1.1: columns 4,5). Last, as we increased the watershed budget, the changing effect of additional EOs on priorities decreased more gradually in the first time period (moving down fourth column in Table 1.1).

Finally, we tested if the continuing change in priorities apparent in Fig. 1.3a,b might continue into the future if data collection conditions persist. The results of this analysis can be seen in the last two columns of Table 1.1. Bold values in the last two columns of Table 1.1 reveal that under several sensitivity tests we expect additional EOs to continue to affect priorities if collection conditions persist. For instance, when we used a watershed budget of 60, we predicted an effect of one year of EOs in both 2010 and 2030. Contrarily, in four cases additional EOs affected priorities in 2010 (Table 1.1: column 9) but are not predicted to do so in 2030 (Table 1.1: last column). The even split in the results of this part of our analysis, along with the fact that we performed extrapolation, makes it unclear how common the continuing effect of additional EOs will be in the future for our study system.

In Appendix S1 we describe results for other sensitivity tests including our choice of conservation objective and controls for data reliability and sampling bias.

1.5 Discussion

How do additional observations of species change spatial conservation priorities? The importance of this question should be evident by the growing spatial prioritization literature (Kukkala & Moilanen 2013) as well as the ongoing use of opportunistically collected element occurrences (EOs) for prioritization. We addressed this question by examining how the ranking of watersheds in the U.S. state of Tennessee changed as species observations were recorded over the past century. We ranked watersheds by their complementary contribution to conserving species richness and assessed how our results were affected by the number of watersheds considered for conservation action. Our methods and results can give insight into state-level prioritizations for watershed actions that focus on across-watershed complementary richness, to areas early in data collection and to those with a long history of species observations.

Perhaps our most important finding is that when additional data have a significant effect on priorities in 2010, additional data are also likely (54% of cases) to have a significant effect in the future if data collection conditions hold. Because our complementarity analysis uses species identities to determine priorities, small additions of infrequently occurring species have a larger effect on priorities than when prioritizing by local species richness only (e.g. Appendix S1), and we expect this to persist in the future as long as data collection conditions continue.

Another key take-home message from our analysis is that additional observations tended to have a decreasing effect on priorities as we amassed data. While this phenomenon is well documented in studies focusing on the accuracy of prioritizations (Grantham *et al.* 2008), we offer a novel data context and explanation for why this occurs. The causal mechanism is due to the complementarity goal we used: additional observations increased the evenness of assemblages across watersheds, lowering the probability that one more observation in a watershed made that watershed necessary for a high-richness solution.

Our results have multiple competing consequences for conservation. In cases similar to ours, new observations may continue to significantly determine current spatial priorities and thus should be collected for that purpose. At the same time, current prioritizations will not necessarily match those for next year, so we do not expect additional observations to determine, on their own, long-term priorities. Our results also suggest that initial observations in data-poor regions will have the greatest effect on determining priorities. Subsequent observations then serve to refine those priorities. Finally, we found that our results were sensitive to conservation goals and to controls for data reliability and sampling bias (Appendix S1). Specifically, when ranking by local richness, the degree of decrease in priority changes over time was smaller as was the overall magnitude. Similarly, when ranking by local richness, future priorities are only expected to be affected by additional EOs when using a low tolerance for data reliability. The weaker effect of additional EOs on priorities when ranking by local richness is because the local richness objective ignores what is unique about different watersheds and thus misses influential variation in the data. Taking our results in aggregate, we suggest that the most effort in species observation be put forth early when a conservation group enters a new area. However, species observations should not cease since new observations will help refine priorities and update them as conditions change. In locations other than Tennessee, we expect similar patterns. The time in the collection record at which the qualitative shift from large changes in priorities with a rapid decrease in changes to smaller, but persistent changes in priorities – characterized in Fig. 1.3a,b – will depend on the conservation context. Managers in other locations could therefore use an analysis such as ours to determine the balance of data collection and conservation actions over time.

Measures of performance of conservation plans enable decision makers to assess whether more information is needed before acting (Polasky & Solow 2001), what management actions to take (Walters & Hilborn 1978), and to rank methods for creating new plans (Grantham *et al.* 2010). While we did focus on differences between priorities over time, the number of species covered by choices of watersheds for conservation drove the prioritization process (Fig. 3c,d, Fig. S 1.3). As such, our analysis is most similar to passive adaptive management (Walters & Hilborn 1978; Williams 2011), in that the decisions we make are refined as we gain information, but differs in that we do not assume that choices in one year affect those in the next year.

We necessarily made several decisions that may have affected our results. First, priorities were updated annually, which may partially explain small priority changes overall. Longer update periods increase the median change in priorities (Fig. S 1.2). Second, we chose to use raw point occurrences rather than modelled data to prioritize watersheds. As a result, conservation actions focusing on the highest priority watersheds in any one year would be under-representative of potentially important areas (Rondinini *et al.* 2006). Aims to create comprehensive conservation plans should, when possible, use a mix of raw point occurrences and modelled data (Rondinini *et al.* 2006). That being said, we anticipate many of the effects we find will carry over to cases where practitioners are combining the two data types. Third, our ranking method used two pieces of information to come up with relative rankings of watersheds: the spatial locations and species identities of EOs. This was done intentionally so we could directly relate changes in species observations to changes in priorities. Additional information would increase the direct applicability of our results to Tennessee. For instance, the costs and patterns of land use change and management over time would have made apparent in priority setting the trade-off of these factors with species coverage. We could have also chosen to base our analyses on the un-ranked irreplaceabilities or local richnesses of watersheds rather than transforming to ranks first. Our assumption was that all decision variables required for a ranking of watersheds were encompassed in their irreplaceabilities and thus keeping additional information was unnecessary. In reality, conservation actions will rely on relationships among players, detailed site histories, short-term opportunity, and other such information which is rarely recorded over such long time-spans. Finally, our statistical choices affect the inferences that can be drawn from our analyses. For example, the change in priorities in one year was calculated on two watershed rankings whose datasets overlapped substantially, and the overlap grew over time. Therefore, each ranking was not independent of the earlier rankings. As such, particular significance levels should be interpreted cautiously.

Prioritization is a common and necessary part of conservation planning. Here, we have provided insight into how regular updating of priorities is affected by additional data. Unlike previous studies, our study used a conservation-relevant dataset that spans 111 years, which enabled us to explore long-term trends others could not. Additionally, we used simple prioritizations that did not account for data weaknesses. As such, our results may reflect practice more closely than other studies. Our results suggest that conservation planners can expect additional observations to alter priorities when conservation goals are complementarity-based.

1.6 Appendix

1.6.1 Appendix S1. Additional sensitivity tests.

To examine how the choice of conservation objective affected our results, we present variations to the complementarity-based analyses in the main text. These additional tests are based on the local species richness of sites; we thus name the ranking method and related sensitivity tests as *Local Richness* sensitivity tests. As with the complementarity case, there are multiple ways to construct sensitivity tests. Within the *Local Richness* ranking method are variations on the number of watersheds prioritized, a control for sampling bias, and controls for data reliability.

Note, Fig. S1, S2, S3 and S4 are all separate from the *Local Richness* sensitivity tests and supplement the main text on complementarity based ranking.

1.6.2 Ranking Watersheds by Local Species Richness

1.6.2.1 <u>Local Richness</u>

We first ranked all 925 watersheds by the number of unique species found within each watershed regardless of the abundance of that species. This is our base case for the *Local Richness* ranking method and we call it *925-watersheds* (Table S 1.1).

1.6.2.2 <u>Number of Watersheds</u>

We also looked at sensitivity tests with fewer watersheds and did so for two reasons. First, the number of watersheds affects the size of the dataset in any one year and consequently the relative contribution of one year's data to the current ranking of watersheds. Second, the distribution of EOs across any one subset of watersheds may be substantially different from the entire distribution and could consequently alter the effect of new observations on priorities. To come up with a subset of watersheds, we simulated a conservation plan in which only the most species-rich places are ever considered for conservation (Williams *et al.* 1996). We chose three subsets of watersheds from the total set of 925. We chose watersheds based on the worst ranking they achieved in the base case just described, such that we were left with close to 5, 50 and 100 watersheds. We used the 4, 48 and 93 watersheds that ever achieved, at worst, a rank of 1.5, 12.5 and 22.0, respectively in *Local Richness: 925-watersheds*.

Table S 1.1. Detailed list of sensitivity tests performed for the (1) *Local Richness* ranking method, in which watersheds were ranked by their unique species richness. Sensitivity of the ranking method to (2) the number of sites used in the prioritization process, (3) sampling bias, and (4) data reliability, was tested.

Code	**Description**
(1) *925-watersheds*	all watersheds (n = 925)
(2) *4-, 48-, 93-watersheds*	4, 48 and 93 best-ranked watersheds from *925-watersheds*
(3) *control-bias*	85 watersheds with at least 50 records by 2010, using rarefied datasets
(3) *no-control-bias*	85 watersheds from *control-bias* without controlling for sampling bias
(4) *1-, 2-, 10-, 25-, 50-year*	Data older than 1, 2, 10, 25 and 50 years are discarded

1.6.2.3 Sampling Bias

In the above four scenarios, the data are not standardized by sampling effort. Heterogeneous sampling effort could bias estimates of the relative richness of watersheds (Walther *et al.* 1995) and would consequently reduce the extent to which additional data change priorities. To explore the sensitivity of our results to spatial and taxonomic sampling bias, we performed rarefaction by systematically subsampling records in a subset of watersheds and repeating the prior analysis (Gotelli & Colwell 2001). Because many watersheds in our dataset have too few records to perform rarefaction effectively, we chose to look at only those watersheds (n = 85) which have at least 50 unique records by 2010. We started with the data for those 85 watersheds over the entire 111 year period. From these data, we chose 50 unique EOs from each watershed to create a rarefied dataset with 4250 unique EOs. These were used to rank watersheds by richness as we did before. We repeated this process 1000 times, creating 1000 rarefied datasets, each with a new choice of 50 EOs from each watershed. Statistics were then averaged over all 1000 repeats. Finally, to make direct comparisons between the rarefied datasets and the original dataset, we performed the same ranking with the full set of data (each watershed had at least 50 unique EOs) from the 85 watersheds chosen for the rarefaction process.

1.6.2.4 Data Reliability

Observations made very early in the data record may be less reliable indicators of current presences of species than observations made more recently. This happens because technology improves, individuals and populations die and move, and entire species ranges shift. As such, a conservation group may choose to weight early EOs differently to more recent ones. To examine the sensitivity of our results to this practice, we removed old data while repeating the same ranking exercises. We considered five sensitivity tests which vary by their tolerance for data age. In the strictest test, only the most recent year's data are kept. For instance, in 1950 we assumed that the only relevant data were those collected in 1950. In 1951, we used only the data from 1951, and so on. We call this the *1-year* test (Table S 1.1). In the most lenient test, we kept all data, regardless of age. This test is the same as *Local Richness: 925-watersheds*. The other tests kept the most recent 2, 10, 25 and 50 years of data. Other than removing old data from the dataset, the ranking process was the same as before.

1.6.2.5 Statistics for the Local Richness tests

All values of V from Local Species Richness sensitivity tests (Table S 1.1) were arcsine transformed to increase the normality of regression residuals. We found that V almost never exceeded 1. In the few cases this occurred – 16 data points from Local Species Richness: 1-year (85 data points) and 3 data points from 2-year (98 data points) – we withheld those data points to satisfy the upper bound of the arcsine transform. Finally, we did not perform piecewise regressions for Local Richness sensitivity tests because they did not show two distinct periods.

1.6.3 Results for the Local Richness Tests

1.6.3.1 Overall Results

For any one sensitivity test, one year of additional element occurrences (EOs) significantly changed the ranking of watersheds (Table S 1.2: column 2). This was true for all sensitivity tests except the 4-watershed test for which sequential rankings were usually identical. Further, the overall magnitude of change brought on by additional observations grew with the number of watersheds prioritized (moving down column 2 in Table S 1.2). Of the ten sensitivity tests for which we assessed priority changes over time, nine showed a decreasing effect of additional EOs over time.

1.6.3.2 Number of Watersheds

Reducing the number of watersheds ranked by local richness reduced the overall effect of additional EOs on priorities (Table S 1.2: column 2), but had little other effect on the results (Table S 1.2). When we reduced the number of watersheds being ranked to 4, there was a qualitative change in priority changes over time (Table S 1.2).

Table S 1.2. Effect of additional data on priorities and when regressed against time for (1) *Local Richness* sensitivity tests. Sensitivity of the ranking method to (2) the number of sites used in the prioritization process, (3) sampling bias, and (4) data reliability, was tested. Subsequently, (5) the overall magnitude of change brought on by one year of additional element occurrences (EOs) was tested, as well as (6) the change in that value over time. Finally, the regression model used to create (6) was used to predict the change in priorities from additional EOs in (7) 2010 and (8) 2030 if current collection conditions hold.

test	**5. med(*V*)**	**intercept**	**6. slope**	**R^2**	**7. V_{2010}**	**8. V_{2030}**
1. *Local Richness*						
925-watersheds	**0.13**	2.1	**-1.0**	0.20	0.08	0.06
2. *Number of Watersheds*						
4-watersheds	0	-	-	-	-	-
48-watersheds	**0.07**	1.9	**-0.9**	0.09	0.05	0.03
93-watersheds	**0.08**	2.5	**-1.2**	0.21	0.04	0.02
3. *Sampling Bias*						
no-control-bias	**0.01**	2.9	**-1.5**	0.22	-0.01	-0.03
control-bias	**0.11**	1.0	-0.5	0.02	0.10	0.09
4. *Data Reliability*						
1-year	**1.07**	4.9	**-2.0**	0.07	**0.98**	**0.94**
2-year	**0.65**	2.8	**-1.1**	0.05	**0.62**	**0.59**
10-year	**0.26**	2.2	**-0.9**	0.13	**0.22**	**0.20**
25-year	**0.16**	1.4	**-0.6**	0.07	**0.14**	**0.12**
50-year	**0.13**	1.7	**-0.8**	0.15	0.10	0.08

Table columns are [med(*V*)] = median of *V* across all years; [intercept] = model intercept; [slope] = slope of the regression in $V \cdot yr^{-1} \cdot 10^{-3}$; [$R^2$] = adjusted R^2; [V_{2010}] = predicted value of *V* in 2010; [V_{2030}] = predicted value of *V* in 2030

Values in **bold** are significant at the 5% confidence level.

1.6.3.3 Sampling Bias

Controlling for sampling bias had two notable effects on the results. First, there was a large increase in the overall effect of one year of EOs on priorities (Table S 1.2: column 2). Second, controlling for sampling bias erased the negative trend in changing priorities over time (Table S 1.2: column 4).

1.6.3.4 Data Reliability

Removing data older than some age produced more qualitative and quantitative changes than any other Local Richness sensitivity test. First, a low tolerance for data age produced the largest increase in the overall effect of additional EOs (Table S 1.2: column 2). This effect was lost by the time we allowed for data up to 50 years old since at that point most data were being kept. Second, controlling for data reliability led to stronger declines in the effect of additional EOs, though this was more of a quantitative result (Table S 1.2: column 2) and the result is not apparent inTable S 1.2. Finally, controlling for data reliability may increase the overall effect of additional data on priorities enough to maintain that effect into the future (Table S 1.2: last two columns).

Fig. S 1.1. Figures showing V plotted against time for all sensitivity tests. See main text and Table S 1.1 for test definitions

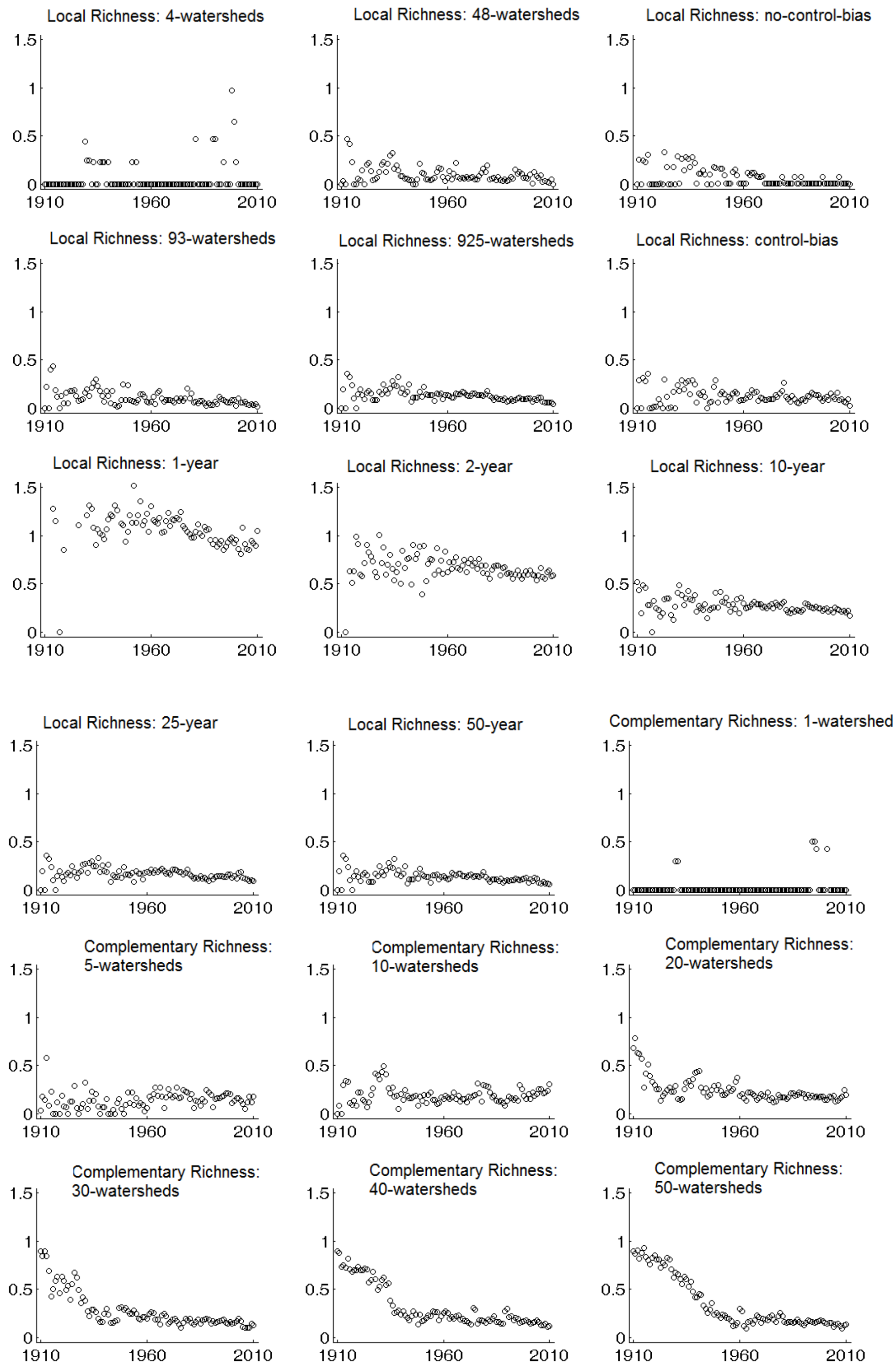


Fig. S 1.1 Continued

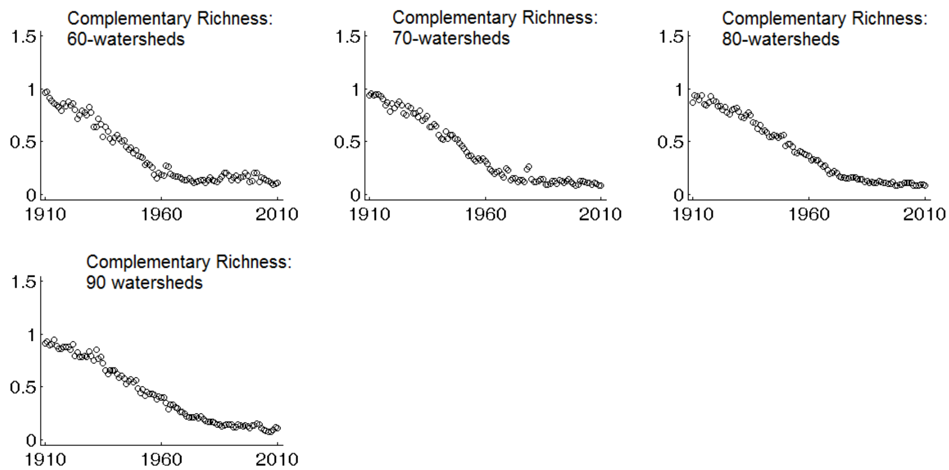


Fig. S 1.1 Continued

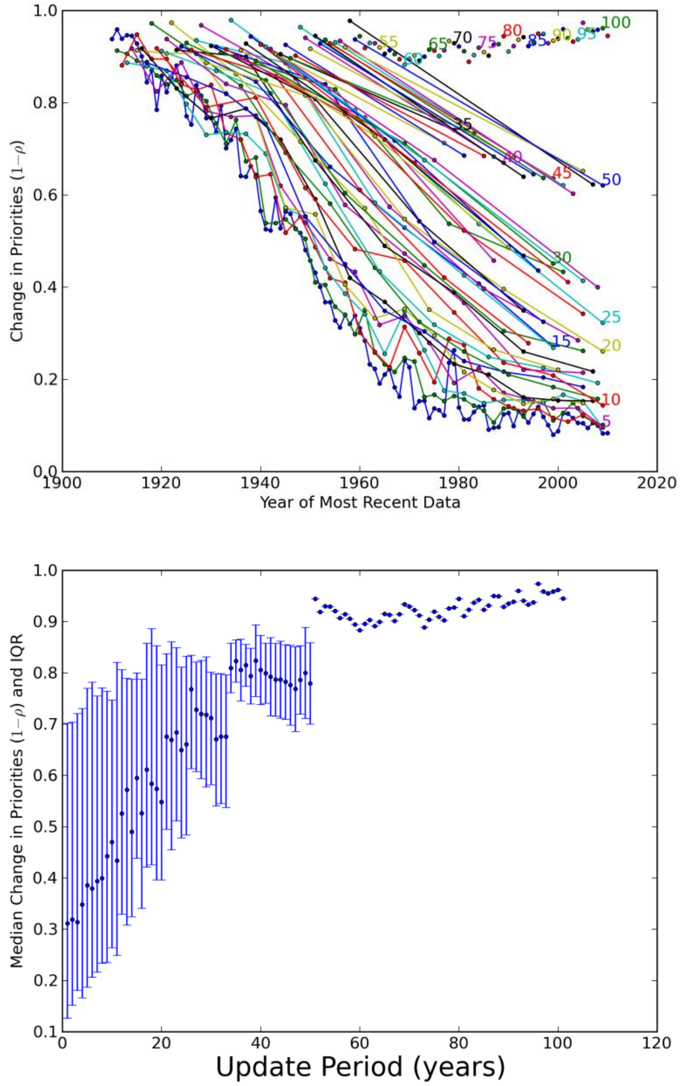


Fig. S 1.2. Changes in priorities over time when budgeting 70 watersheds for complementarity as the length of delay between priority updates changes. *top panel* Each line represents a different update period between 1 and 101 years. The bottom blue line is the same as *70-watersheds* in the main text where we update priorities annually. Higher lines use increasingly longer delays between updating priorities. For instance, the blue line labeled “15” represents the situation where we reassess priorities every 15 years. Intervals of 5-year update periods are labeled with colored numbers. *bottom panel* Median (point), 25% and 75% quartile (error-bars) of the lines in the *top panel*. Larger update periods have too few points to have an IQR.

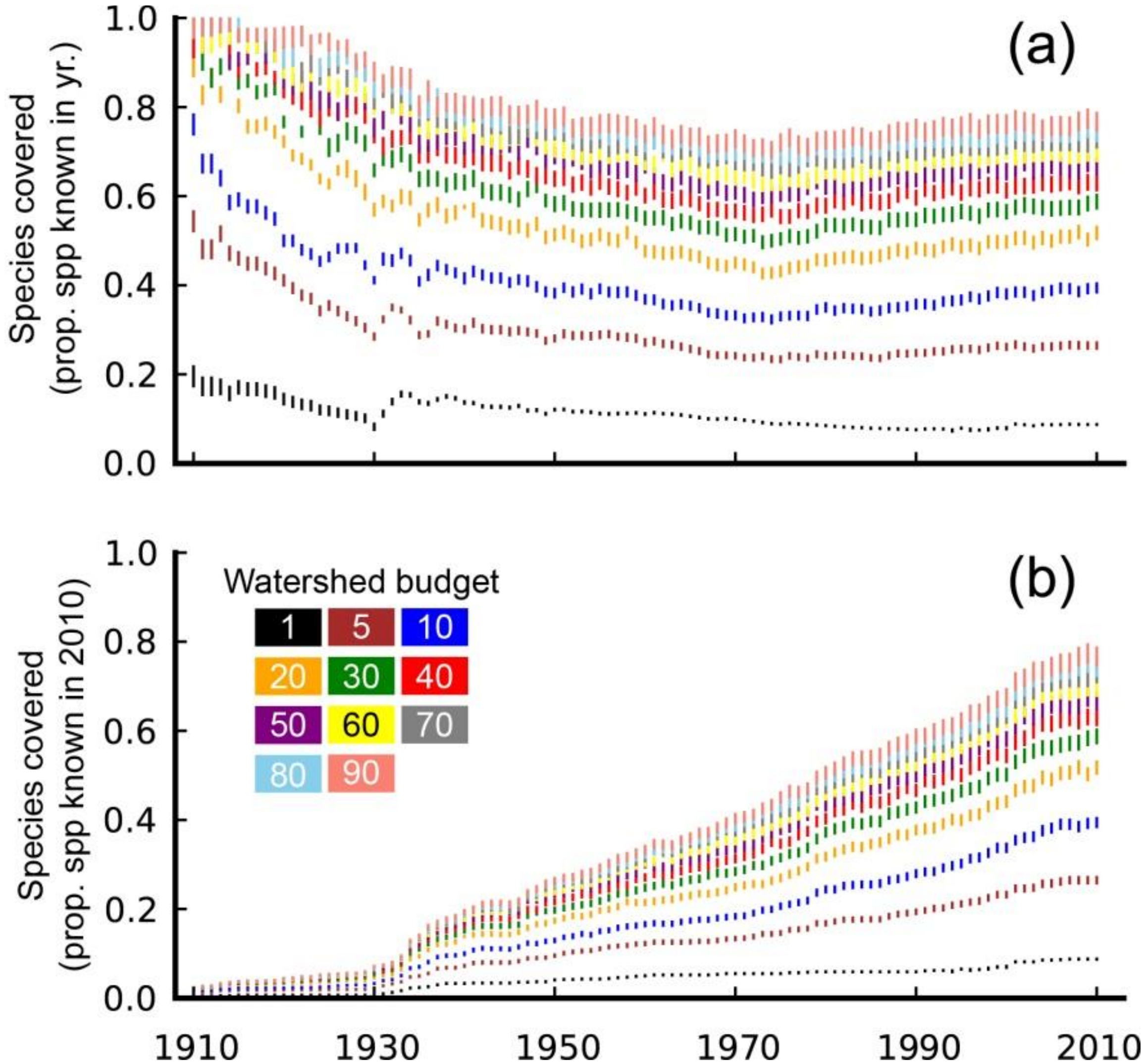


Fig. S 1.3. Performance of conservation prioritizations as element occurrences (EOs) accumulate over time. The measure of performance is the proportion of species with at least one occurrence in Tennessee in *(a)* the current year or *(b)* 2010 that were covered by solutions to the Maximal Coverage Problem (MCP). The bottoms and tops of vertical bars correspond to the minimum and maximum proportions of species covered, respectively. For instance, the blue bar at 1910 in *(a)*, which corresponds to a watershed budget of 10, shows that solutions to the MCP covered ~73-79% of species known to occur in Tennessee in 1910. Note that the ranges of some watershed budgets are partially hidden due to plotting multiple budgets on one figure. The range is always narrow since we kept only the top 5% performing solutions to the MCP.

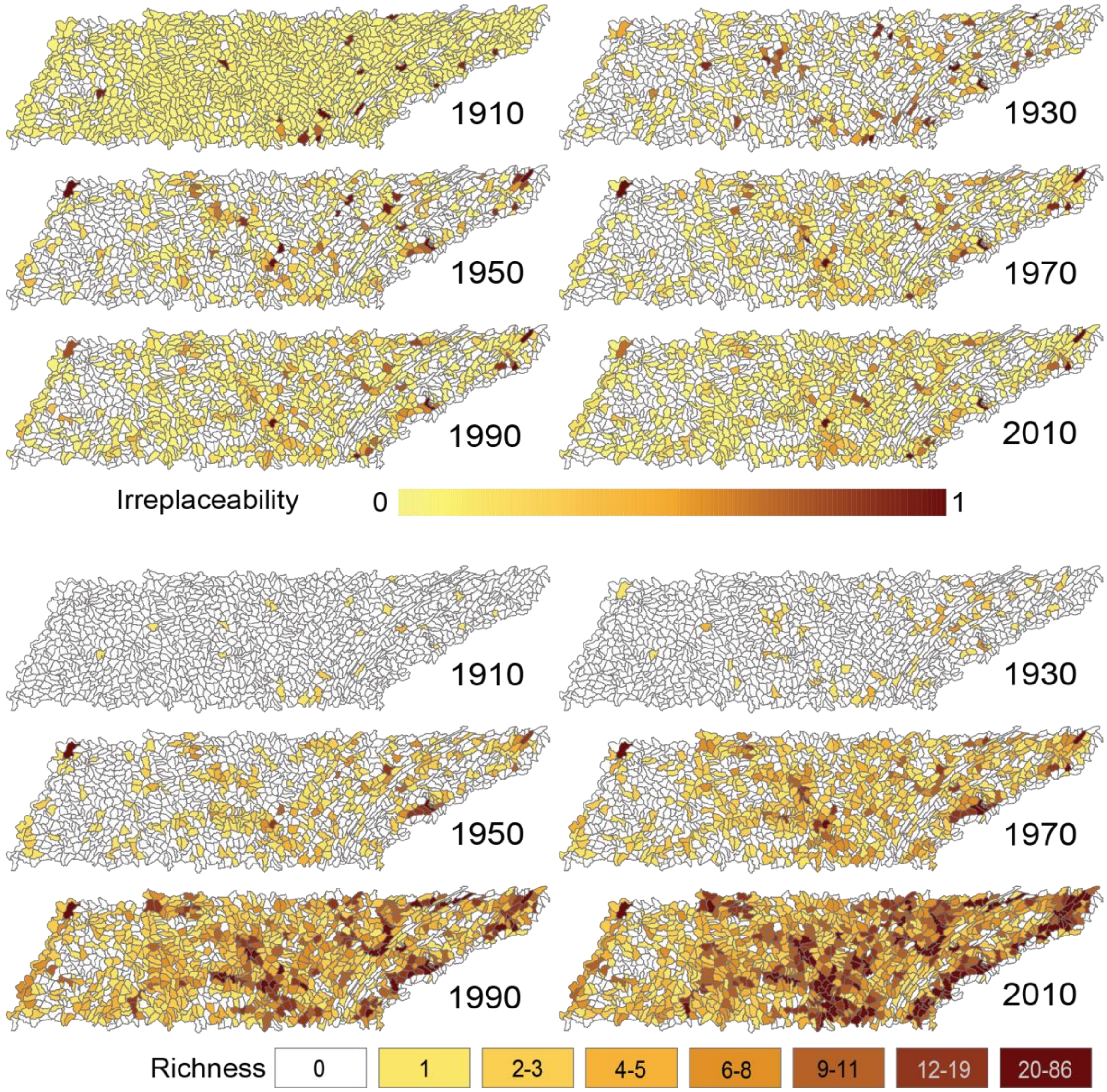


Fig. S 1.4. Spatial distribution of watershed *(top)* irreplaceabilities and *(bottom)* richnesses at six snapshots in time. *(top)* Irreplaceabilities come from the sensitivity test where we choose 70 watersheds to solve the Maximal Coverage Problem and represent the proportion of a sample of near-optimal solutions in which a watershed occurred. The solution to the Maximal Coverage Problem in each year is based on the assemblage of species accumulated in each watershed since 1900 (see main text for details). *(bottom)* Richnesses are the number of unique species accumulated in that watershed since 1900.

Table S 1.3. Overall effect of additional data on priorities and when regressed against $\log_{10}$(size of dataset). Sensitivity of the ranking method to the number of sites used in the prioritization was tested. Subsequently, (column 2) the overall magnitude of change brought on by one year of additional element occurrences (EOs) was tested, as well as (columns 3-8) the trend of changing priorities as the dataset grew. Sensitivity tests without [$slope_2$] and [break] fields use a one-segment regression. Values of *V* were natural logarithm transformed for *Local Richness* tests.

test	med(*V*)	intercept	$slope_1$	$slope_2$	break	R^2	ΔAICc
Comp. Richness							
1-watershed	0	-	-	-	-	-	-
5-watersheds	**0.14**	0.04	-0.21	0.03	2.12	0.04	0
10-watersheds	**0.18**	0.54	1.34	-0.02	1.79	0.10	10
20-watersheds	**0.20**	-2.18	**-0.80**	**-0.04**	2.12	0.73	80
30-watersheds	**0.19**	2.02	**-0.75**	**-0.07**	2.49	0.91	127
40-watersheds	**0.22**	2.11	**-0.53**	**-0.05**	2.93	0.95	154
50-watersheds	**0.21**	1.77	**-0.52**	**-0.05**	3.18	0.96	143
60-watersheds	**0.20**	1.87	**-0.49**	+0.00	3.49	0.97	136
70-watersheds	**0.31**	1.84	**-0.47**	**-0.08**	3.60	0.96	80
80-watersheds	**0.37**	1.84	**-0.42**	-0.05	3.83	0.97	55
90-watersheds	**0.40**	1.74	**-0.41**	-0.08	3.84	0.97	49
Local Richness							
4-watersheds	0	-	-	-	-	-	-
48-watersheds	**0.07**	0.21	**-0.03**	-	-	0.10	-
93-watersheds	**0.08**	0.25	**-0.05**	-	-	0.20	-
925-watersheds	**0.13**	0.24	**-0.03**	-	-	0.17	-
no-control-bias	**0.01**	0.24	**-0.05**	-	-	0.21	-
control-bias	**0.11**	0.18	-0.02	-	-	0.02	-
1-year	**1.07**	1.31	**-0.07**	-	-	0.06	-
2-year	**0.65**	0.79	**-0.04**	-	-	0.04	-
10-year	**0.26**	0.39	**-0.04**	-	-	0.13	-
25-year	**0.16**	0.23	**-0.02**	-	-	0.06	-
50-year	**0.13**	0.23	**-0.03**	-	-	0.12	-

Chapter 2: Synergies and tradeoffs among environmental impacts under conservation planning of shale gas surface infrastructure

At the time of writing, this article was in revision at *Environmental Management*, with the expectation to be published in the near future:

Austin W. Milt, Tamara D. Gagnolet, and Paul R. Armsworth. "Synergies and tradeoffs among environmental impacts under conservation planning of shale gas surface infrastructure." *Environmental Management. In Review.*

Austin Milt performed the secondary data collection, data processing, analysis, interpretation, and writing for this article. Tamara Gagnolet assisted in data collection, and both co-authors contributed intellectually to the design, interpretation, and revision of the article.

2.1 Abstract

Hydraulic fracturing and related ground water issues are growing features in public discourse. Few have given as much attention to surface impacts from shale gas development that result from building necessary surface infrastructure. One way to reduce future impacts from gas surface development without radically changing industry practice is by formulating simple, conservation-oriented planning guidelines. We explore how four such guidelines affect the locations of well pads, access roads, and gathering pipelines on state lands in Pennsylvania. Our four guidelines aim to (1) reduce impacts on water, reduce impacts from (2) gathering pipelines and (3) access roads, and (4) reduce impacts on forests. We assessed whether the use of such guidelines accompanies tradeoffs among impacts, and if any guidelines perform better than others at avoiding impacts. We find that impacts are mostly synergistic, such that avoiding one impact will result in avoiding others. However, we found that avoiding forest fragmentation may result in increased impacts on other environmental features. We also found that single simple planning guidelines can be effective in targeted situations, but no one guideline was universally best at avoiding all impacts. As such, we suggest that when multiple environmental features are important in an area, more comprehensive planning strategies and tools should be used.

2.2 Introduction

High-volume hydraulic fracturing and horizontal drilling combined with higher gas prices in the late 2000s led to the boom of shale gas development in the eastern United States. Hydrofracking is the high-pressure pumping of water and sand to break shale and release gas. Horizontal drilling allows multiple wells to be drilled laterally from a location rather than a single well. In Central Appalachia, especially West

Fig. 2.1. Well pads (rectangular clearings), access roads (linear clearings in bottom-right), and gathering pipelines (other linear clearings) pose many impacts on habitats in the Marcellus formation, including forest and wetland loss and fragmentation, displacement of species of conservation concern, erosion, and freshwater sedimentation and fragmentation. Courtesy of M. Godfrey, The Nature Conservancy.

Virginia and Pennsylvania, U.S.A., where we illustrate a case study, over 14,000 horizontal well permits have been granted since 2008. Many of these (~60%) have yet to be drilled (West Virginia TAGIS Unit 2014; Whitacre 2014), and many additional permits and wells are expected (Evans & Kiesecker 2014). Development of worldwide shale gas reserves (U.S. Energy Information Administration 2013) and accompanying surface infrastructure (Fig. 2.1) should raise concerns about associated environmental impacts. Since much of the potential gas development has not been realized, there are still significant opportunities to reduce future impacts through careful planning.

Terrestrial impacts resulting from the spatial locations of gas surface infrastructure are an understudied and important issue. Impacts occur at all stages of the development process, from pre-production through post-production (Burton *et al.* 2014). Here we focus exclusively on pre-production activities and specifically the construction of well pads, access roads, and gathering pipelines, the impacts of which may play out at different rates and spatial extents. The magnitudes and types of impacts change from stage to stage, but it is clear that site construction incurs the most direct and thus quantifiable land use changes, which here are our focus. To date, most attention has been given to groundwater, surface water, and air quality (Howarth, Ingraffea & Engelder 2011; Entrekin *et al.* 2011; Clark *et al.* 2012; Smith *et al.* 2012; Olmstead *et al.* 2013). Attention to surface impacts has been less common, but is growing (Johnson *et al.* 2010; Davis & Robinson 2012; Slonecker *et al.* 2012; Drohan *et al.* 2012; Evans & Kiesecker 2014). Because the Marcellus shale formation is roughly uniform within a development site (Fig. S 2.1 in Appendix) and because horizontal drilling allows for many wells per pad, shale gas surface infrastructure in the Appalachian region can be evenly spaced and of low density. As such, terrestrial impacts from surface infrastructure are similar to low-density, rural housing development. Like houses, well pads are spaced hundreds of meters apart and connected by gravel roads (Fig. 2.1). Pipeline corridors are similar in width and surface maintenance requirements to underground electric transmission corridors. During construction, drilling, and hydraulic fracturing, gas development may have much larger cumulative impacts than other types of development because of high-traffic trucking of materials, the size of temporary well pad staging areas, and the intensity of drilling noise and light.

Much of shale gas development is occurring in areas of high biological diversity (Gillen & Kiviat 2012; Kiviat 2013). In the Central Appalachian region, especially Pennsylvania, there are large areas of relatively intact forest on protected and unprotected lands. As of late 2013, 32% of state-owned public land area in Pennsylvania host well permits (Whitacre 2014). More development is occurring on private, unprotected lands, much of which have high conservation value (Robles *et al.* 2008). At lease-hold and larger scales, the spatial configuration of gas surface infrastructure may greatly fragment forests (Slonecker *et al.* 2012; Drohan *et al.* 2012; Racicot *et al.* 2014). Stream crossings for roads and pipelines may reduce stream connectivity and increase sediment delivery and risk of spills to streams. The construction of surface infrastructure disturbs soils, which changes local topography and may lead to increased erosion. When erosion and sedimentation controls are inadequate or fail, runoff can increase sedimentation in streams. In the event of hazardous waste spills, soils and streams may be contaminated and experience large die-offs of biota (Lustgarten 2009; Detrow 2012). For more details on potential environmental impacts of shale gas surface infrastructure, see Gillen and Kiviat (2012), Slonecker et al. (2012), Kiviat (2013) and others.

In this paper we explore how multiple environmental impacts are associated with one another in shale gas infrastructure development. We specifically focus on using simple practices to plan well pads, access roads, and gathering pipelines, and we also assess the relative effectiveness of our planning practices at avoiding impacts. Simple planning practices, which focus on achieving one or a few goals, may help developers incorporate additional environmental objectives into surface development without having to radically change practices. However, there is the risk that following simple practices will exacerbate tradeoffs among multiple environmental objectives since simple practices must be, by definition, narrowly defined. We use the Marcellus shale play in Pennsylvania as a case study. We planned well pad footprint locations, access road routes and gathering pipeline routes for twenty state forests and game lands, where shale gas development could occur. We attempted to follow the same planning steps in the order the gas industry uses and adhered to construction constraints imposed by state laws and development practices (Appendix). For each development site, we created four infrastructure layouts corresponding to four simple siting practices. For each infrastructure layout, we computed eight impact metrics that reflect the conservation objectives of stakeholders in the region. We looked for synergies and tradeoffs among impacts, i.e. whether some impacts were positively or negatively correlated. We also assessed whether some of our planning practices performed better than the others for some impacts, i.e. an impact metric was significantly lower when using one practice versus the others.

Similar articles have focused on the hydraulic fracturing process (Kargbo, Wilhelm & Campbell 2010) at relatively small scales (Slonecker *et al.* 2012). Racicot et al. (2014) undertook a similar planning exercise in which they planned well pads, access roads, and pipelines for a small region in Quebec, Canada. They looked at the potential impacts of surface infrastructure under various regulatory constraints, focusing on impacts with, versus without, additional ecological restrictions. Racicot et al. (2014) also analyzed the extent to which fairly simple ways of affecting development patterns in turn affected potential environmental impacts of that development. Our paper is novel in several ways. First, we look at impacts on a larger suite of environmental and human features, choosing to explore the broad spectrum of potential impacts since features will vary in importance from place to place. Second, we assess the synergies and tradeoffs among our suite of impacts and explore how our findings may affect more general planning of gas infrastructure. Third, we look at impacts over a large spatial extent, which may allow us to draw more general conclusions for the greater Appalachian region. Finally, we assess the relative effectiveness of simple planning practices across the suite of metrics, enabling us to see how simple practices may target one or several impacts.

2.3 Methods and Materials

2.3.1 Shale Energy Industry Reference

In planning surface infrastructure within our study sites, we attempted to follow the planning practices of the shale energy industry. Our main contact was Triana Energy, LLC, which is a privately held oil and gas exploration and production company based in Charleston, West Virginia, USA and operating in both West Virginia and Pennsylvania. We used Triana's experience with shale energy development to inform the dimensions of production units, construction limits such as road grades, and surface infrastructure planning more generally.

2.3.2 Study Site Selection

We chose twenty case study sites from the 319 State Forests, State Parks, and State Game Lands underlain by Marcellus Shale in Pennsylvania (Fig. 2.2). Selection of study sites was done independently of current development status, severed mineral ownership, and the probability of future shale development. We chose Pennsylvania state lands as study sites for a few reasons. First, these public lands are managed by the state and are large tracts of consolidated land, which makes large-scale planning for multiple pads and thus potential gains from our study larger than in other areas. At the same time, mineral rights are severed from surface ownership on some public lands in Pennsylvania, which means that shale gas development can and is occurring on these lands. In areas with intact surface and subsurface ownership, some state forest lands and state game lands have also been leased for shale gas development. Finally, we chose to use highly forested public lands in a state with a large forest system because we wanted to highlight the potential impacts on relatively intact terrestrial habitats.

State lands were first buffered by 100 m to remove small gaps between otherwise contiguous lands. We then combined any buffered lands that overlapped. We reduced our set of lands to those that could support between one and five full-size production units. A production unit is not physical infrastructure, but represents the area drained of gas by the wells of a single well pad (Fig. S 2.1). To reduce the set of lands, we first reduced the set of combined lands to those between 3 km^2 and 30 km^2 (n=216). Next, we maximally covered each buffered land with 914 × 3352-m (3000 × 11,000-ft) production units rotated 27° (Triana Energy, LLC *pers. comm.*); the size of the production units is based on an assumption of 6 wells per pad. We considered a production unit economically feasible if at least 90% of its area was within the buffered land. Production unit placement was done visually and by hand. We categorized the buffered lands into one of five size classes by the number of full-size production units we estimated they would support. Finally, we chose at random four buffered lands from each size class as our study sites. We note that the number of production units we visually estimated would fit in each buffered land is usually smaller than the total number we later placed when implementing our four siting practices. All GIS analysis was done in ArcMap 10.1 (ESRI).

Cultural features were mapped using 2008 National Agricultural Imagery Program aerial images. All non-industrial cultural features, such as homes, agricultural buildings, retail businesses, and recreational fields overlapping the 20 chosen sites were mapped with free-hand drawn polygons. These were used to calculate distance to cultural features within the study sites.

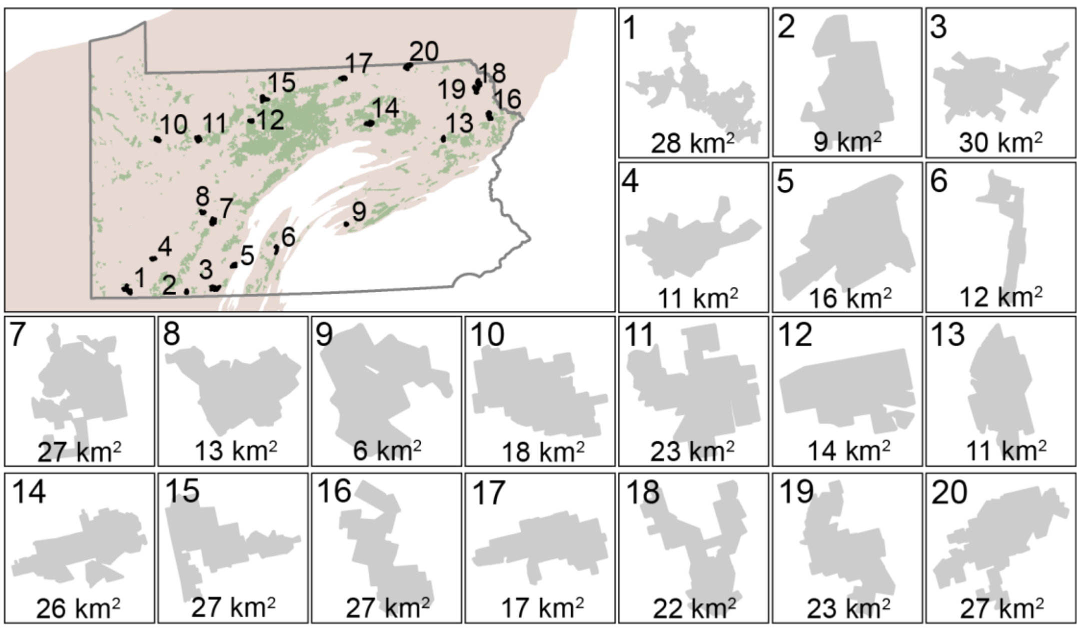


Fig. 2.2. (large map) Pennsylvania State Forests, State Parks, and State Game Lands (green or darker gray) within the Marcellus gas formation (beige or lighter gray) serve as candidate lease areas from which we chose twenty public lands (black outlines) for our analysis. We chose to use state lands that could support one to five full-size production units after buffering individual lands and combining overlapping lands. (small maps) chosen sites with numbers corresponding to large map and site areas in km^2.

2.3.3 Siting Practices

We chose to use four siting practices that are easy to understand, easy to implement, and that we expected to differentially achieve conservation goals. Our four siting practices are 1) site pads close to roads to reduce the impacts of access roads, 2) site pads close to pipelines to reduce the impacts of gathering pipelines, 3) site infrastructure away from water to reduce risks to aquatic systems, and 4) site infrastructure in disturbed areas to reduce impacts on intact habitats. We sited infrastructure four times at each study site, once per siting practice. In addition to following these practices, we adhered to existing regulatory and construction constraints (Appendix). We also attempted to follow existing industry planning practices, which are to site well pads/production units first, then to select routes for access roads and gathering pipelines (Triana Energy, LLC, *pers. comm.)*.

Before implementing a siting practice, we placed production units to maximally cover the study site. We started by placing full-size production units in the manner described in *Study Site Selection*. We then placed half-size production units in the same fashion. There is subjectivity in exactly where production units are placed, but we expect a human will find the maximum number of production units and near-optimal locations.

Next, well pads were placed within pad envelopes in the production units. A pad envelope is a smaller area within a static production unit that represents potential locations for a well pad while still draining the production unit (Fig. S 2.1). Moving the pad within its envelope changes the routes of wells. In a full-size production unit, the pad envelopes are 1524 × 122 m. In a half-size production unit, pad envelopes are 122 × 122 m. After restricting ourselves to the pad envelopes, we sited pad points (at the center of a raster pixel) to primarily adhere to the given siting practice. Secondarily, well pads were placed to center on the envelopes (or one envelope in the half-sized production unit), which reflects the developer practice of making all lateral wells close to a desired length (Triana Energy, LLC *pers. comm.)*. Finally, if there were still options (especially in the case of half-production units) we put the well pads in the location closest to the nearest pad (siting done by visual inspection). We also adhere to Pennsylvania regulatory setbacks for well pads and other infrastructure.

In most production units (63 of 93 or 68%), the "in disturbed areas" practice gave no direction about where to place well pads, because there were no cultivated or developed areas to bias placement. As such, most of the placement for these was done based on trying to center the pad in the envelope and/or put it near other pads. However, subsequent road and pipeline routes did depend on this practice.

Well pads were connected to the existing road network using a least cost path method (ESRI 2013). The cost surface varied depending on the siting practice. When siting pads close to (1) existing roads or (2) existing pipelines, the road cost surface was determined by the distance, elevation, and slope over which the road was built. These two factors reflect monetary costs of road construction materials and soil movement. When siting infrastructure (3) away from water, the cost surface was the inverse of the Euclidean distance from the water, such that cost increased with proximity. Where the distance to water was 0, we set the inverse distance to be 100 times larger than the maximum value in the rest of the raster. Finally, when siting infrastructure (4) in disturbed areas, we used the 2006 National Land Coverage Dataset to weight disturbed classes (Developed, Planted/Cultivated) as 100 times less costly than undisturbed classes and used that as the cost surface. Because the existing pipeline network is sparse, some study sites do not intersect with the network. We therefore buffered the study sites by 15 km – chosen to ensure every study site intersects with at least one pipeline – before creating the placement regions for the pipelines and roads. We then built pipelines and roads outside of study sites where needed to complete the line, but clipped them off at the study site boundary and measured impacts only inside the study site. Pipelines and roads were sited using the same methods. Finally, we note that because the least-cost path method checks for a spatially additive impact, there is an implicit assumption that placing more infrastructures is always more impacting. As such, the placement algorithm is biased towards shorter roads and pipelines. In reality, some impacts like *forest frag 1* below are spatially non-additive such that lower impacts *can* be achieved with more infrastructure in some cases.

2.3.4 Environment Impact Metrics

For each infrastructure layout, we calculated several environmental and cultural impact metrics. Metrics were calculated on a rasterized version of the study area, where cells are 30x30 m. We calculated the following metrics after each siting practice was implemented in a study site:

1. cultural – risk to human cultural features; sum over the raster of proximity to cultural features, where proximity is defined as the inverse of the Euclidean distance of infrastructure to the nearest mapped cultural feature.

2. erosion – erosion potential; sum over the slope raster of all cells occupied by infrastructure.

3. forest loss – total number of previously forested cells in which infrastructure is developed

4. forest frag 1 – effective mesh size; area of each patch if all the forest were combined and then divided into S equally sized patches with the same degree of landscape division as the original set of patches; similar to average patch size, but with more consistent responses to fragmentation (Jaeger 2000);

$$m = \frac{A_t}{S} = \frac{1}{A_t} * \sum_{i=1}^{n} A_i^2$$

   where A_t is the number of cells in the analysis region, and A_i is the area of each forest patch *after* fragmentation.

5. forest frag 2 – perimeter to area ratio; number of forest edge cells – those which border a non-forest cell – divided by total number of forest cells after development.

6. rare spp – risk of displacing rare or other target species; sum over raster of expected number of occurrences of known locations (EOs) of rare species. To create this surface, a map of habitat types from the Northeastern Terrestrial Habitat Mapping Project () was overlaid with EOs from the Pennsylvania Natural Heritage Program (Table S 2.1). Each habitat type was assigned the number of EOs it contained. This was done without regard to species identity. This number was then divided by the total area of the habitat type to reduce areal effects. Finally, values were multiplied by one million to aid understanding.

7. water 1 – reduction of stream connectivity; number of stream crossings; number of stream cells occupied by infrastructure.

8. water 2 – risk to aquatic systems; sum over raster of proximity to water bodies, where proximity is defined as the inverse of the distance a liquid would flow from a cell over the surface to reach the first-encountered water body. This was calculated using ArcGIS's Flow Direction tool.

2.3.5 Analyses

We tested two major hypotheses:

1. Siting practices differentially affect metrics of impact. The goals here were to assess whether the choice of practice is important and if some practices are more generally effective than others.

2. Impacts are correlated across sites and practices. The goal here was to assess which impacts trade-off when planning infrastructure by simple planning guidelines.

Before performing the tests below, we square-root transformed *forest loss* and *cultural* and $\log_{10}$ transformed all other data to increase normality. To test the first hypothesis, we built eight repeated measures ANOVAs, one for each metric as the response, where the study site was the "individual" on which repeated measures were being taken. The "treatments" on each individual were the practices themselves, and the ANOVA was used to test if practices created significant variation in the impact response of the site. We then used a post-hoc Tukey HSD test to determine, for those significant models, which practices differentially affected each impact metric. The ANOVAs were carried out using ezANOVA in R (v3).

To test the second hypothesis, we calculated Pearson's correlation coefficient between each pair of impact metrics, within each practice. For instance, one correlation was between *forest loss* and *erosion* when siting infrastructure in disturbed areas. We used a Bonferroni corrected alpha of 0.0018 (alpha = 0.05/28 = 0.0018) to test significance. We focused only on post-development correlations of impacts, but do explore the marginal change in impacts associated with following these simple planning practices in Appendix.

2.4 Results

Infrastructure layouts were qualitatively different based on which simple planning practice was followed (Fig. 2.3). For instance, Fig. 2.3 shows four layouts resulting from our four planning practices for one development site. Note especially the lengths of gathering pipelines and that their routes differed markedly based on the locations of well pads, which were themselves in markedly different locations based on the planning practice. These same patterns hold for the other development sites (Fig. S 2.3).

Differences in the locations of infrastructure led to varying impacts. While placing infrastructure away from water generally performed better across impacts than placing pads near existing roads or pipelines, or placing infrastructure in disturbed areas, no one practice was universally better across impacts than the others (Fig. 2.4). This should be expected since our simple practices purposefully target one or two objectives and cannot accommodate more without becoming intractable. Further, most impact metrics we measured were positively correlated, such that doing better at avoiding one impact meant doing

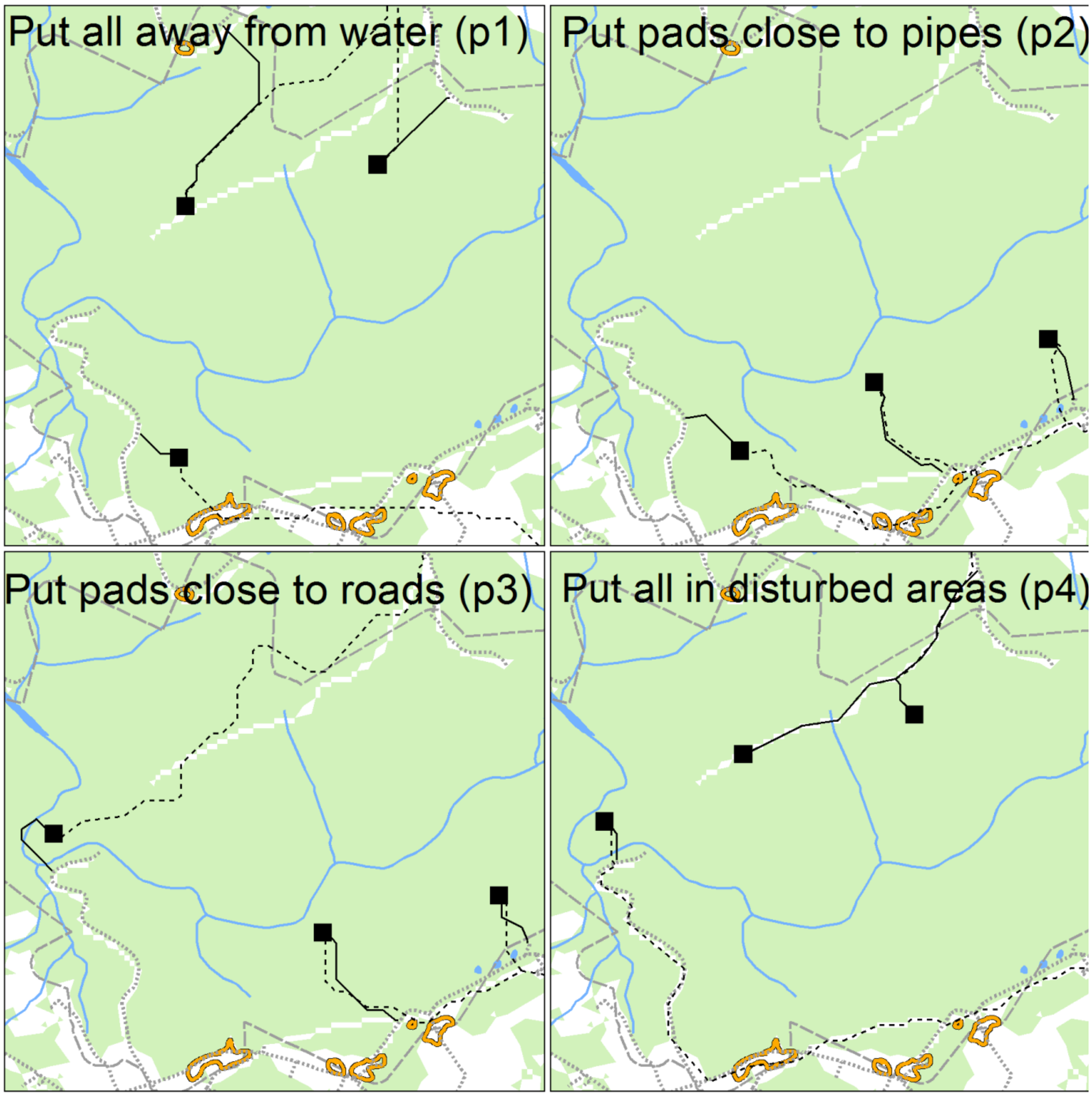


Fig. 2.3. Four layouts of well pads (■), roads (−), and pipelines (--) at one of our twenty sites (medium gray --). Other layouts presented in the Appendix. Existing roads are medium gray dots (…); proposed infrastructure is black. Each layout is the result of following a different simple siting practice (codes here). Each layout, if developed, would result in different impacts on forests (green or light gray), aquatic habitats (blue or medium gray), erosion, and cultural features (orange or dark gray with black outline) (Table S 2.2 for values).

better with others. However, one measure of forest fragmentation had consistent tradeoffs with other impacts (Fig. 2.4). This is an important caveat for Pennsylvania, where potential forest fragmentation is large (Johnson *et al.* 2010; Evans & Kiesecker 2014).

We found that for this study region, set of impact metrics, and simple siting practices, impacts are more synergistic than antagonistic. This is revealed in Fig. 2.4 by the dominance of positive correlations (+) between impacts. The strongest and most numerous positive correlations occurred between the *water 2* impact metric, a measure of proximity of infrastructure to water bodies, and other metrics. Conversely, there were no significant positive correlations between the *rare spp* metric and other metrics. However, there is a negative correlation between *forest frag 1*, a measure of forest fragmentation, and both *forest loss* and *erosion* potential. There is also a less strong but persistent tradeoff with water impacts. Tradeoffs involving forest fragmentation are especially important in Pennsylvania, where there are large areas of intact forest. Thus, if we prioritize minimizing forest fragmentation, this may come at the cost of increasing forest loss, erosion, and stream impacts.

The bar charts and letters in the bottom of Fig. 2.4 reveal that while practices differ in how they affect each impact metric (different groups within each impact metric), no practice performed better (had lower means) than others across all impacts. This is expected, since our simple practices were chosen to achieve specific goals. For instance, putting infrastructure away from water (Fig. 2.4 "p1") produced the lowest water impacts, but some higher forest impacts. That example also reveals that simple planning practices can be relatively effective at avoiding targeted impacts. Another example is that putting infrastructure in disturbed areas resulted in relatively lower levels of *forest frag 2* and *forest loss*. Putting infrastructure close to (*p2*) existing pipelines or (*p3*) existing roads did not produce clear patterns in impacts relative to the other practices. This may indicate that planning practices that do not clearly target an environmental feature like water or forest are less likely to be effective in a predictable manner. That said, both practices performed the best with regards to *erosion* potential. Finally, the blank bars in Fig. 2.4 show that there was no significant difference in performance of planning practices with regards to *forest frag 1* or *rare spp*. Both *forest frag 1* and *rare spp* are metrics that will respond to fairly large landscape changes. Effective mesh size (*forest frag 1*) responds most when large patches are divided, which is relatively unlikely in this context. The *rare spp* metric is a coarse resolution metric.

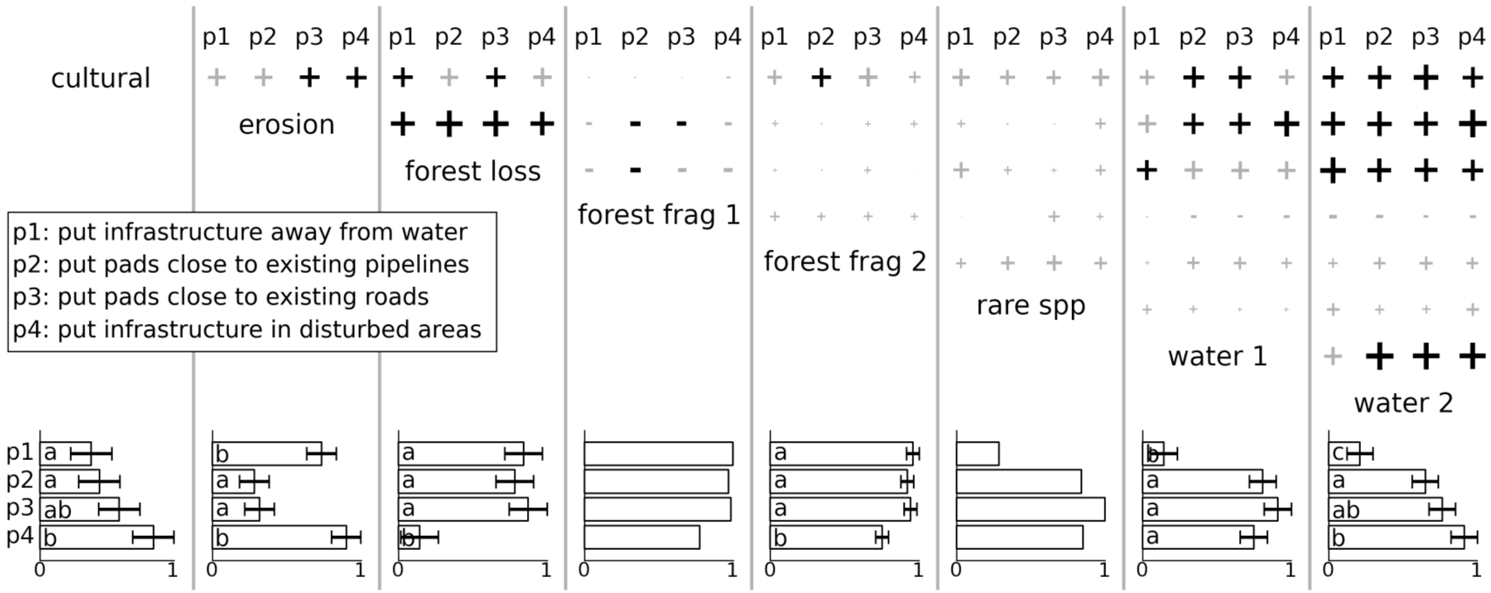


Fig. 2.4. Synergies and tradeoffs among impacts and practices. (grid) Correlations between impact metrics across all 20 sites. Negative correlations (-) are tradeoffs; synergies are +. Symbol size reflects correlation magnitude. Significant correlations (black) were assessed with Bonferroni corrections. (bar charts) Performance of practices. Bars are mean values (+/- Tukey HSD confidence intervals) within the impact listed above on the diagonal, e.g. the left set of bars corresponds to the *cultural* impact of different siting practices. Non-overlapping errors bars and group letters indicate significant differences. Values are scaled 0 to 1, where 1 is worst. *forest frag 1* and *rare spp* repeated measures ANOVAs were not significant, so no Tukey HSD tests were performed.

2.5 Discussion

To inform the use of simple guidelines for planning shale gas surface infrastructure, we wanted to know if the use of such guidelines accompanies tradeoffs among environmental impacts. We also sought to assess the relative effectiveness of guidelines for avoiding impacts. Using Pennsylvania public lands as a case study, we sited well pads, access roads, and gathering pipelines according to four guidelines that are easy to understand and implement. We calculated and correlated eight proxies for environmental impacts that would result from developing such infrastructure.

We found mostly weak positive correlations between impacts, indicating that for our study area, multiple impacts could be avoided simultaneously with environmentally-oriented planning. More generally, some impacts will be negatively correlated – i.e. trade-off – in that trying to avoid one impact by changing the location of infrastructure necessarily leads to increasing another impact. In those cases, the best strategy to avoid impacts is to choose locations that balance antagonistic impacts. In this study, one measure of forest fragmentation - effective mesh size - was negatively correlated with several other impacts. This is likely due to both spatially intrinsic tradeoffs between impacts and the use of simplified planning guidelines (Fig. S 2.2). This result is important in Pennsylvania where reducing impacts on forest fragmentation are highly important and may come at the cost of increasing other impacts, regardless of planning practices. These patterns generally hold when looking at marginal – as opposed to absolute – changes brought on by development (Fig. S 2.2).

We also found that simple guidelines for planning surface infrastructure can be relatively effective when focusing on one or a few impacts. For instance, our analysis revealed that, for this context, putting infrastructure away from water was more effective at avoiding water and cultural impacts than other simple planning guidelines. Unfortunately, no one simple guideline was universally better at avoiding impacts than all other guidelines. As such, when more targeted, site-specific planning is possible, simple planning guidelines that encompass the planning context may be good enough.

Since our results suggest that regulations targeting single impacts will not be universally effective, new regulations either need to be more comprehensive or site specific, e.g. through a review process similar to Pennsylvania's Natural Diversity Inventory Environmental Review Tool (http://www.gis.dcnr.state.pa.us/hgis-er/Login.aspx). For instance, our results indicate actions that target reductions in forest impacts by placing infrastructure in disturbed areas (e.g. Fig. 2.3) will actually increase erosion potential, stream crossings, and cultural impacts relative to alternative strategies (Fig. 2.4). Actions that incentivize developers to take an impact-comprehensive planning approach could be more effective, but will be more challenging to design and implement than simpler, targeted ones. Alternatively, implementing a site-level review process where regulator and developer work together to identify potential impacts and avoid them could be effective, though at an increased time and resource commitment for both sides.

Our study joins the large and growing body of work on measuring and assessing the relationships between multiple indicators of environment and society for decision making. Increasingly common are studies of tradeoffs among ecosystem services (Bennett, Peterson & Gordon 2009; Raudsepp-Hearne, Peterson & Bennett 2010; Moilanen *et al.* 2011; Maskell *et al.* 2013; Howe *et al.* 2014). At the same time, multi-objective planning is an essential part of conservation more generally (Cattaneo *et al.* 2006). Our study incorporates elements of land use decision making, another arena in which ongoing research highlights that understanding the effects of land use on multiple indicators is important (Phalan *et al.* 2011; Sayer *et al.* 2013).

We chose impact metrics to cover features and topics of concern to stakeholders and to illustrate synergies and tradeoffs between multiple objectives. These are clearly not exhaustive. We focused on impacts at one particular stage – site development – while impacts occur at all stages. In reality, impacts from surface infrastructure development extend beyond the scope of our study, and include both positive and negative effects. For instance, while we focus on known rare species, development also impacts common species and species not prioritized by conservation. We focus on known rare species because these are a greater focus for decision makers. Indeed, developers in Pennsylvania are required to avoid known locations of rare species through the Pennsylvania's Natural Diversity Inventory Environmental Review process mentioned above. We also use a simple measure of impacts on humans that ignores some positive and negative socioeconomic effects of development (Sovacool 2014).Many such effects play out over scales largely independent of the scale of analysis here, while others could form alternative or additional components of a more comprehensive analysis at this study's scale. Regardless, we were able to identify some synergies and tradeoffs with the impacts we did include.

Future analyses similar to ours could benefit from a few methodological changes. First, while public lands served as an informative case study for testing impacts from development, shale gas is being extracted all over the Central Appalachian region, including private lands. Private lands tend to be more fragmented in ownership and land use, which would likely affect the resulting infrastructure layouts and impacts from that development. It is unclear how transferrable our results are to the private lands context. Second, we did not rigorously test the change in impact correlations before versus after surface infrastructure was placed, choosing to qualitatively assess the change in correlations (Fig. S 2.2). While this does not affect our practice-specific conclusions described above, it does affect our ability to say with statistical confidence whether simple guidelines induced tradeoffs and synergies among impacts.

Multiple routes can be taken to reduce impacts from development including regulation, land protection, and changing industry practices. Each of these routes is currently underutilized in the Central Appalachian region, but we expect this study will help inform future actions in each. First, federal and state regulations exist to restrict the placement of surface infrastructure but do not address some important ecological impacts such as habitat fragmentation. Furthermore, most of Pennsylvania's current setback requirements that aim to protect sensitive habitats potentially can be waived with a request and justification by the gas developer. Second, the efficacy of traditional land acquisition and easement is uncertain in shale gas development sites where surface and subsurface rights to land are sometimes separately owned. Our results suggest an effective strategy for conservation groups and landowners - in conjunction with willing gas industry partners - is to inform site-specific planning where priority environmental or biological features are present. Third and finally, the gas industry currently does not have strong incentives to go above and beyond regulation to further reduce impacts. Increasing environmentally-oriented planning will require some effort to lower the threshold of entry for the gas industry. Simple planning practices can be more easily assimilated into existing planning than more complex tools and practices. However, more advanced tools and methods may be needed to help industry planners incorporate environmental objectives into their planning, especially when tradeoffs exist among impacts.

2.6 Appendix

2.6.1 Change in impacts due to development

In the main text we focus on correlations between impact metrics *after* development has occurred. In doing so we put emphasis on the absolute tradeoffs/synergies between impacts and consequently take a more holistic approach to reducing impacts rather than focusing on the practice-specific effects of development on impact metrics. Another relevant approach is to test the change in correlations between impacts induced by following our four simple planning practices. We briefly explored how correlations between impact metrics changed from before to after development occurred.

Our methods for measuring correlations before development were very similar to that presented in the *Analyses* section in the main text: we measured the values of our impact metrics at each of the twenty development sites and assessed correlations between them. There are two important aspects that differ here. First, six of our eight impact metrics always have a value of 0 before development. Only *forest frag 1* and *forest frag 2* have a non-zero value before development. In order to measure correlations between impacts for the six other metrics, we calculated, for each development site, the average value of the impact surface raster, since those six metrics are just spatially additive functions of the surface rasters. For instance, for the *erosion* metric, we took the average value of the slope raster. We only included in the average those values in the raster that were feasible locations for roads since the feasible road area is an intermediate between the pad and pipeline areas and we wanted to include those areas that might actually be developed. The second difference from our main text analysis is we did not differentiate between planning practices before development, since there would be no difference anyway.

After calculating pre-development values and correlations, we looked at the difference in correlations between post and pre-development. Since we did not differentiate between planning practices before development, we combined the post-development metric values across planning practices so there is only a single comparison for each pair of metrics.

Fig. S 2.2 summarizes the results of this analysis. Before development (top of Fig. S 2.2), many impacts trade-off with one another, probably necessarily due to being affected in spatially disjoint areas. As the Difference table reveals (bottom of Fig. S 2.2), planned development would often increase synergies (red, positive correlations) between impacts (in 16 of 28 comparisons). At the same time, planned development would lead to relative tradeoffs (blue, negative correlations) between impacts (in 12 of 28 comparisons). For the most part, the general patterns that we present in the main text (somewhat summarized in the After sub-table in Fig. S 2.2) hold when looking at marginal changes brought on by development (summarized in the Difference sub-table).

In some cases (e.g. forest frag 2 vs. forest loss), planned development could reverse the correlation between impacts (e.g. from -95 to +13). Finally, the green, italicized numbers indicate that very few correlations were significant after Bonferroni correction. We did not do statistical testing for the significance of the Difference table values.

2.6.2 Minimum Development Constraints

We adhered to existing regulatory and construction constraints on the placement of gas surface infrastructure. Namely,

2.6.2.1 pads

- 0 m from lease boundary
- 0 m from pad zones
- 100 m from wetlands > 1 acre
- 30 m from stream
- 152 m from cultural feature

2.6.2.2 roads

- 0 m (but not in) wetlands > 1 acre
- slope < 10000%. Note that generally roads would not be graded to larger than 15% slope and in steep places, construction of roads on steep slopes would be accomplished by switchbacks, which the least-cost path algorithm in ArcMap is not able to accomplish. As such, a few planned roads in our analysis may have been unrealistically steep.

2.6.2.3 pipes

- 0 m (but not in) wetlands > 1 acre

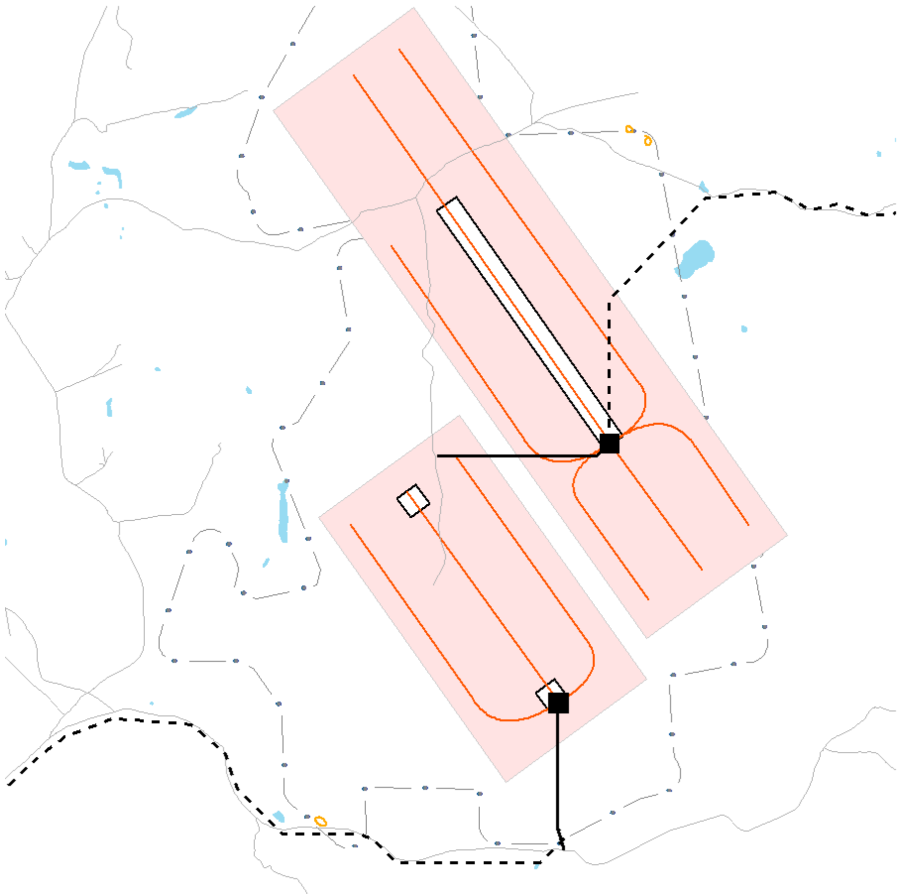


Fig. S 2.1. Illustration of production units (pink rectangles), pad envelopes (white rectangles with black outline) and how they relate to the well pad (black squares) and associated well laterals (red lines) within a development site (dot-dash gray line). The larger, full-sized production unit has six wells, while the half-sized drainage unit has only three. As a result, the full-sized unit places the well pad in the center, while the half-sized unit places the well pad one of the ends. Note that regulations require wells to stay within the development site boundary, so one well from the full-sized unit is cut-off. We did not explicitly site well laterals, so the production unit boundaries do not correspond to what would be the final well trajectories. We include the well laterals here for illustration.

BEFORE

	erosion	forest loss	forest frag 1	forest frag 2	rare spp	water 1	water 2
cultural	-48	-22	65	30	56	9	-3
erosion		23	-24	-24	***-71***	32	49
forest loss			-30	***-95***	19	-43	-26
forest frag 1				30	29	13	-3
forest frag 2					-8	49	36
rare spp						***-39***	-41
water 1							***91***

AFTER

	erosion	forest loss	forest frag 1	forest frag 2	rare spp	water 1	water 2
cultural	62	53	-18	52	53	52	***74***
erosion		64	-50	15	14	29	***66***
forest loss			-61	13	27	37	58
forest frag 1				29	17	-24	-39
forest frag 2					44	24	33
rare spp						***23***	36
water 1							***84***

DIFFERENCE

	erosion	forest loss	forest frag 1	forest frag 2	rare spp	water 1	water 2
cultural	110	75	-82	22	-3	43	78
erosion		41	-26	39	85	-4	17
forest loss			-31	107	8	80	84
forest frag 1				-1	-12	-36	-36
forest frag 2					52	-25	-3
rare spp						62	77
water 1							-7

Fig. S 2.2. Results of comparison of impact correlations before and after development. Red cells are positive correlations. Blue cells are negative correlations. Bold, green, italicized text denotes significance at the 95% confidence level. Correlation coefficients are multiplied by 100 for ease of presentation

Table S 2.1. Datasets used to place infrastructure and calculate impact metrics.

Dataset	Used for	Source
contiguous forest patches	siting infrastructure, forest metrics	National Land Cover Dataset 2006 (classes 41, 42, 43, 90): http://www.mrlc.gov/nlcd06_data.php
Wetlands in PA	restricting infrastructure, siting infrastructure, water 2 metric	National Wetlands Inventory: http://www.fws.gov/wetlands/Data/Data-Download.html
habitat classifications for NE USA	rare spp metric	Northeastern Terrestrial Habitat Mapping Project: http://conserveonline.org/workspaces/ecs/documents/ne-terrestrial-habitat-mapping-project
Spatial locations of rare species observations in PA	rare spp metric	Pennsylvania Natural Heritage Program data request. Data only available through PNHP formal request.
Streams, rivers, lakes	restricting infrastructure, siting infrastructure, water metrics	Pennsylvania State University (via PASDA): Networked streams of Pennsylvania http://www.pasda.psu.edu/default.asp
Digital Elevation Model, elevation at a point in raster format	restricting infrastructure, siting infrastructure, erosion metric	USGS (1 arc second ~30 m): http://ned.usgs.gov/
"cultural" features, i.e. schools, recreational fields, dwellings, reservations, etc.	restricting infrastructure, cultural metric	Pennsylvania National Agricultural Imagery Program 2008, 1m resolution aerial imagery http://www.pasda.psu.edu/default.asp

Table S 2.1 Continued

Dataset	Used for	Source
existing roads	siting infrastructure	US Census Bureau TIGER/Line 2008 (all counties All Lines RDFLAG = "Y"): http://www.pasda.psu.edu/default.asp
existing pipelines	siting infrastructure	MapSearch pipelines. Propriety data.
conservation lands in PA	choosing sites, siting infrastructure	PASDA-Pennsylvania Conservation Stewardship (1998): http://www.pasda.psu.edu/default.asp

Fig. S 2.3. In this series of figures we show the infrastructure configurations created for 3 representative development sites. Each figure contains four layouts of well pads (■), roads (–), and pipelines (--) at one of our twenty sites (gray). Existing roads and pipelines are brown; proposed infrastructure is black. Each layout is the result of following a different simple siting practice. Each layout, if developed, would result in different impacts on forests (green), aquatic habitats (blue), erosion, and cultural features (orange).

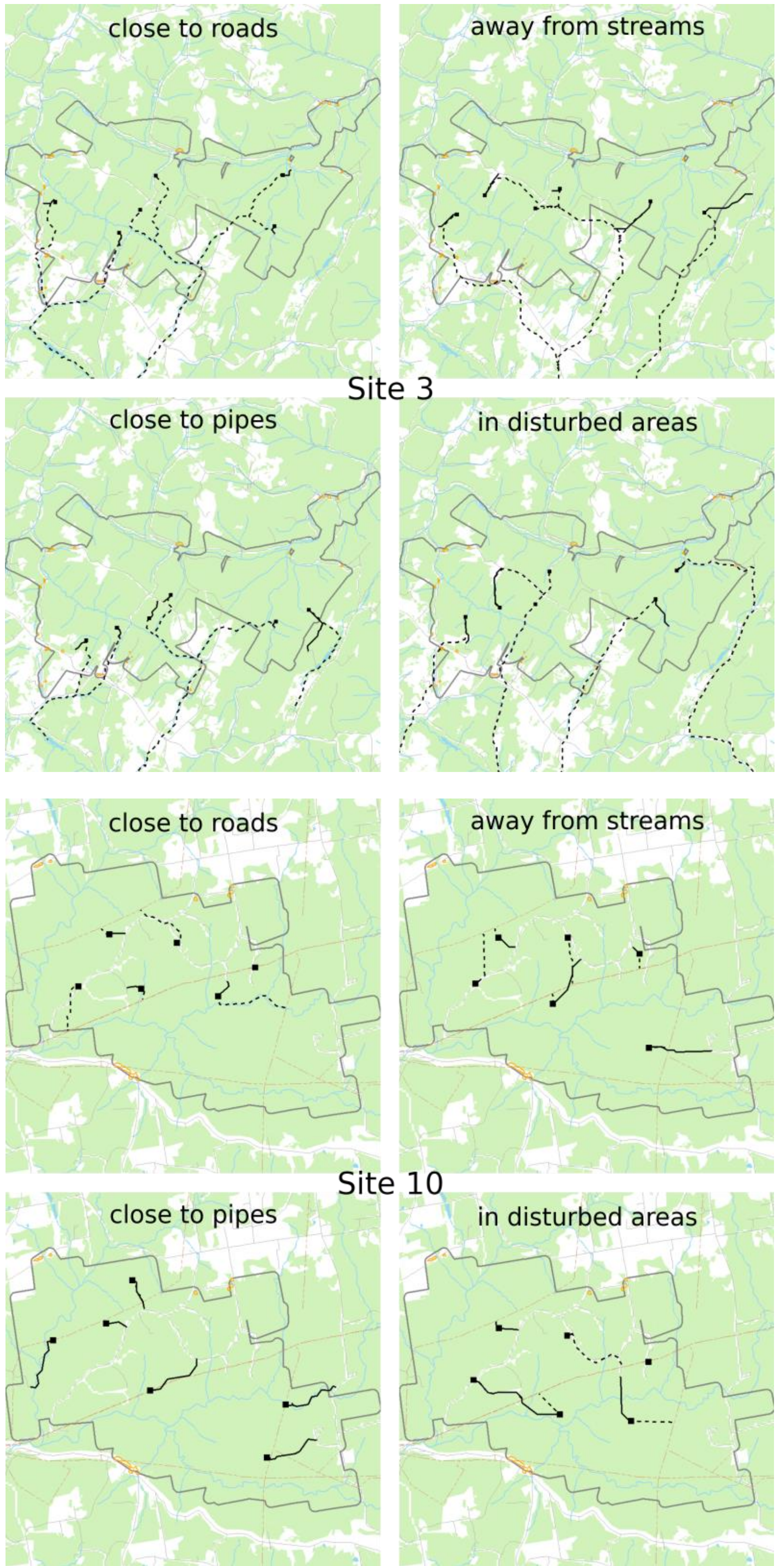


Fig. S 2.3 Continued

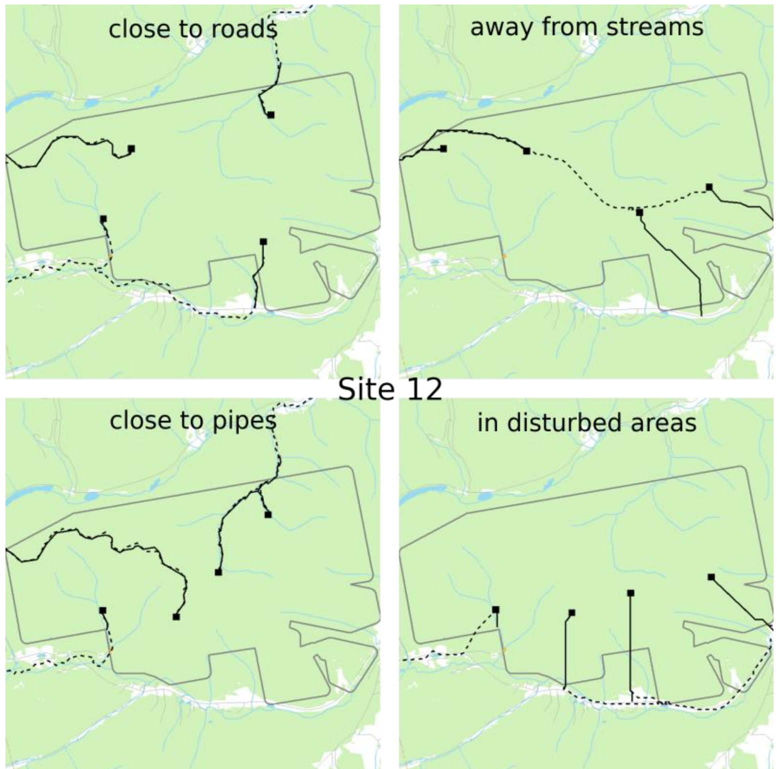


Fig. S 2.3 Continued

2.6.3 Analyses Results

Table S 2.2. Impact metric values for each of the layouts shown above and including the "no-development" scenario (practice "pre"). Note the pre-development values are not measured the same way as post-development values (see Additional Methods above). These are untransformed values.

Table S 2.3. Continued

		impact metrics							
site	**practice**	**cultural**	**erosion**	**forest loss**	**forest frag 1**	**forest frag 2**	**rare spp**	**water 1**	**water 2**
1	p1	0.74	9511.66	317	21739.13	0.256	6241.51	0	0.97
1	p2	0.62	3630.59	189	21739.13	0.251	2545.89	67	2.10
1	p3	1.30	4000.03	249	21739.13	0.252	2145.84	170	3.59
1	p4	1.41	8078.43	165	21739.13	0.250	2277.06	82	3.55
1	pre	0.001	18.796	0.478	22222.22	0.248	2.894	0.056	0.007
2	p1	0.22	2401.26	159	15151.52	0.111	2061.91	0	0.30
2	p2	0.11	2045.48	168	14492.75	0.111	1869.36	0	0.24
2	p3	0.40	1602.99	202	14925.37	0.116	2540.80	34	1.57
2	p4	0.49	2466.99	138	15151.52	0.103	2173.17	8	0.84
2	pre	0.001	10.317	0.810	15384.62	0.088	5.434	0.026	0.004
3	p1	1.81	12048.07	519	52631.58	0.163	3350.06	3	1.83
3	p2	1.21	6698.81	542	52631.58	0.164	4812.26	251	4.35
3	p3	1.74	6857.29	499	52631.58	0.162	2980.85	249	3.58
3	p4	1.91	13684.22	340	52631.58	0.153	2313.02	85	5.55
3	pre	0.001	18.063	0.676	52631.58	0.146	2.363	0.047	0.006
4	p1	1.09	2875.96	211	16129.03	0.189	4421.34	0	0.42
4	p2	1.19	2456.47	233	15873.02	0.190	4311.15	20	0.97
4	p3	0.80	3845.82	315	16949.15	0.199	5655.39	78	1.53
4	p4	0.90	3884.77	113	15873.02	0.178	2227.30	9	1.32
4	pre	0.001	14.505	0.595	17543.86	0.172	3.250	0.042	0.006

Table S 2.3. Continued

site	practice	impact metrics							
		cultural	erosion	forest loss	forest frag 1	forest frag 2	rare spp	water 1	water 2
5	p1	0.70	5258.18	286	4424.78	0.277	1974.14	4	0.81
5	p2	1.18	1764.74	184	4347.83	0.264	2001.67	48	1.72
5	p3	1.36	2051.58	288	4366.81	0.273	647381.94	84	3.45
5	p4	0.98	4700.61	118	4310.34	0.255	216134.83	24	2.31
5	pre	0.002	8.333	0.444	4545.45	0.249	3.203	0.052	0.005
6	p1	0.14	3020.49	81	13333.33	0.124	655.23	0	0.12
6	p2	0.07	598.92	63	13157.89	0.123	1339.40	0	0.14
6	p3	0.11	1055.46	72	13333.33	0.124	1030.55	0	0.15
6	p4	0.10	1894.63	12	13333.33	0.120	184.30	7	0.65
6	pre	0.001	14.574	0.717	13333.33	0.120	4.975	0.038	0.005
7	p1	0.41	2799.89	340	35714.29	0.169	6299.90	2	0.62
7	p2	0.42	2563.12	350	35714.29	0.171	22066.97	11	0.94
7	p3	0.48	3034.22	361	35714.29	0.170	82877.62	24	1.45
7	p4	1.64	7882.51	175	35714.29	0.160	77877.20	15	3.06
7	pre	0.001	9.227	0.664	37037.04	0.155	5.241	0.040	0.005
8	p1	1.02	3084.90	220	7518.80	0.284	18796.90	0	0.55
8	p2	0.96	1034.21	156	7692.31	0.280	353768.56	35	1.10
8	p3	1.26	1445.75	197	7692.31	0.285	370628.50	52	1.41
8	p4	0.95	5759.22	111	5952.38	0.270	170054.69	21	3.11
8	pre	0.002	9.944	0.516	7874.02	0.266	3.827	0.048	0.006
9	p1	0.13	3450.25	127	10101.01	0.110	1829.85	0	0.15
9	p2	0.04	973.69	60	10204.08	0.098	757.35	5	0.32
9	p3	0.12	1132.58	74	10204.08	0.102	1333.76	7	0.61
9	p4	0.04	1223.11	57	10204.08	0.098	660.54	3	0.22
9	pre	0.002	14.545	0.797	10309.28	0.090	5.222	0.034	0.005

Table S 2.3. Continued

site	practice	impact metrics							
		cultural	erosion	forest loss	forest frag 1	forest frag 2	rare spp	water 1	water 2
10	p1	0.16	1961.45	202	23255.81	0.135	3519.31	0	0.31
10	p2	0.15	1620.88	232	23255.81	0.138	3556.95	38	1.38
10	p3	0.17	1713.43	188	23809.52	0.134	2609.62	40	0.94
10	p4	0.20	2591.10	190	23255.81	0.134	2181.54	8	0.99
10	pre	0.001	8.693	0.718	23809.52	0.124	6.660	0.041	0.006
11	p1	0.19	1664.89	169	41666.67	0.075	2359.23	0	0.47
11	p2	0.15	1641.51	143	41666.67	0.073	2017.07	22	0.78
11	p3	0.25	1429.56	143	41666.67	0.073	1140.00	43	0.87
11	p4	0.15	3040.84	127	41666.67	0.072	801.58	17	1.43
11	pre	0.001	8.208	0.851	41666.67	0.067	6.698	0.035	0.005
12	p1	0.44	12666.93	485	25641.03	0.114	2431.98	4	1.16
12	p2	0.50	5597.44	458	25641.03	0.102	1365.40	150	2.49
12	p3	0.53	4871.99	374	25641.03	0.095	1091.21	138	2.54
12	p4	0.28	10429.01	224	25641.03	0.085	289.85	34	2.90
12	pre	0.001	25.140	0.872	26315.79	0.069	0.762	0.046	0.006
13	p1	0.09	1059.29	132	12500.00	0.122	1832.60	0	0.14
13	p2	0.23	1348.39	215	12345.68	0.132	5484.25	21	0.86
13	p3	0.11	881.23	166	12345.68	0.127	2872.77	2	0.29
13	p4	1.53	3905.93	153	12345.68	0.119	3148.98	8	1.14
13	pre	0.002	8.363	0.753	12658.23	0.108	6.427	0.030	0.004
14	p1	0.44	11737.51	617	58823.53	0.102	8313.79	3	1.11
14	p2	0.58	8949.48	685	58823.53	0.105	9147.52	83	3.41
14	p3	0.74	7990.75	684	58823.53	0.106	9351.45	112	3.82
14	p4	1.74	17552.61	557	40000.00	0.101	18896.08	68	5.48
14	pre	0.001	16.131	0.805	58823.53	0.080	1.990	0.035	0.005

Table S 2.3. Continued

site	practice	impact metrics							
		cultural	erosion	forest loss	forest frag 1	forest frag 2	rare spp	water 1	water 2
15	p1	0.45	8437.78	446	66666.67	0.149	790.53	0	1.57
15	p2	0.78	6223.11	704	66666.67	0.156	745.11	96	3.59
15	p3	0.53	4550.11	551	66666.67	0.152	572.20	89	2.23
15	p4	0.99	9245.37	352	66666.67	0.146	745.94	32	3.52
15	pre	0.000	16.526	0.698	71428.57	0.137	0.825	0.048	0.006
16	p1	2.29	12147.42	963	29411.76	0.150	18662.62	5	1.76
16	p2	1.57	5443.25	615	28571.43	0.136	24734.86	54	3.60
16	p3	1.89	5108.42	713	29411.76	0.140	25629.98	91	4.65
16	p4	2.89	8104.22	365	29411.76	0.125	126910.11	22	3.44
16	pre	0.001	8.541	0.719	30303.03	0.116	6.552	0.021	0.004
17	p1	0.17	4283.59	365	20000.00	0.111	4162.54	0	0.48
17	p2	0.16	1700.87	335	20000.00	0.108	5030.55	0	0.46
17	p3	0.18	2199.23	353	20000.00	0.110	5172.63	0	0.56
17	p4	0.51	8496.39	260	10526.32	0.102	32221.05	23	2.68
17	pre	0.001	18.599	0.747	20408.16	0.088	1.540	0.024	0.004
18	p1	0.67	7658.05	508	30303.03	0.242	5762.46	6	1.13
18	p2	1.36	5157.50	573	28571.43	0.245	78492.70	190	4.58
18	p3	1.44	5279.11	742	27027.03	0.256	2969.75	326	6.60
18	p4	1.02	9820.28	249	29411.76	0.227	18051.86	27	4.18
18	pre	0.001	9.923	0.481	30303.03	0.217	2.988	0.028	0.004
19	p1	0.65	6139.97	477	43478.26	0.172	6276.80	0	0.79
19	p2	1.08	3558.60	510	43478.26	0.174	34686.67	84	2.33
19	p3	1.11	4106.18	582	41666.67	0.177	66242.91	170	3.41
19	p4	1.23	7077.80	353	43478.26	0.164	50512.85	27	3.66
19	pre	0.001	9.296	0.655	43478.26	0.151	4.238	0.028	0.004

Table S 2.3. Continued

site	practice	impact metrics							
		cultural	erosion	forest loss	forest frag 1	forest frag 2	rare spp	water 1	water 2
20	p1	0.44	3992.51	301	22727.27	0.253	3108.47	0	0.64
20	p2	0.77	3079.36	344	22222.22	0.255	4552.39	125	3.04
20	p3	0.70	2865.79	345	22727.27	0.255	3511.98	146	3.11
20	p4	0.85	7910.96	160	22727.27	0.244	1394.25	81	5.04
20	pre	0.001	10.247	0.366	23255.81	0.240	1.446	0.043	0.005

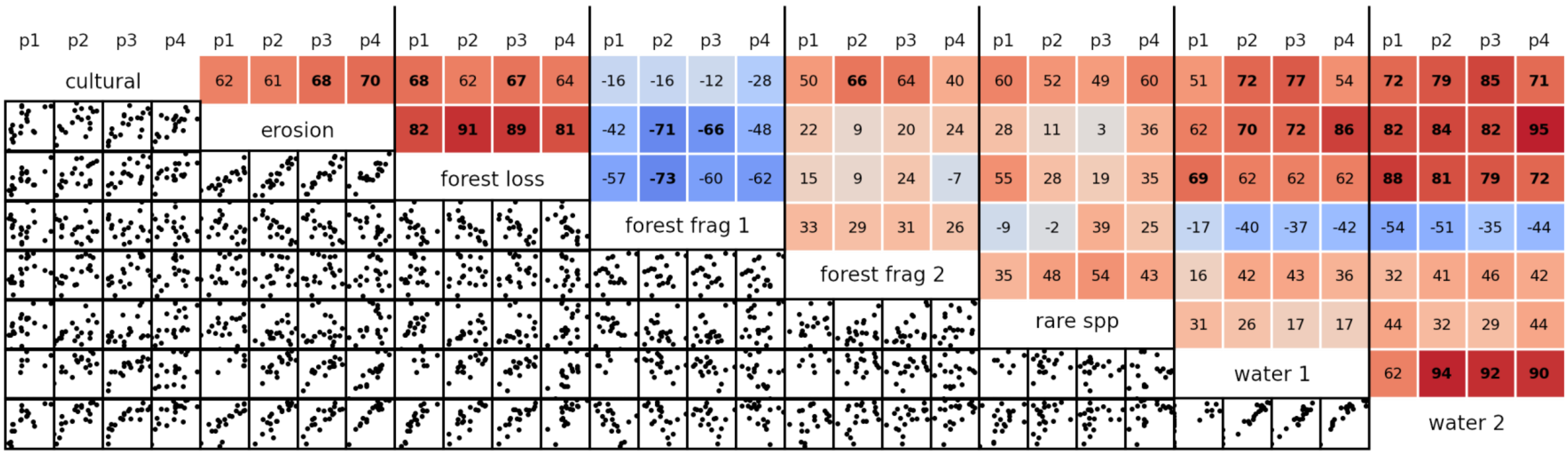


Fig. S 2.4. (above diagonal) Pearson's correlations between impacts within practices. Each correlation is taken across all 20 sites. Colors correspond to the strength and direction of the correlation. Numbers are correlation coefficients multiplied by 100. Significance of correlations was computed using a Bonferroni correction where k = 28. Significant correlations are in bold. (below diagonal) Scatter plots of impacts within practices and impacts. These were used to create the correlation coefficients above the diagonal. Practice codes are p1: put infrastructure away from water, p2: put pads close to existing pipelines, p3: put pads close to existing roads, p4: put infrastructure in disturbed areas.

Table S 2.3. Results of repeated measures ANOVAs and Tukey HSD tests for practices within impacts. Note, for "forest frag 1" and "rare spp," the non-significant ANOVA does not mean that development had no effect on those metrics. It only means that different practices did not differentially affect the metric.

	Metric	**SS Error**	**p-value**	
	cultural	1.48	9.31E-04	
	mean	**2.5% CL†**	**97.5% CL†**	**Group**
p1	0.71	0.65	0.78	a
p2	0.74	0.67	0.81	a
p3	0.81	0.74	0.87	ab
p4	0.92	0.85	0.98	b

	Metric	**SS Error**	**p-value**	
	erosion	0.95	1.37E-12	
	mean	**2.5% CL†**	**97.5% CL†**	**Group**
p1	3.66	3.61	3.71	b
p2	3.41	3.36	3.47	a
p3	3.43	3.38	3.49	a
p4	3.75	3.70	3.80	b

	Metric	**SS Error**	**p-value**	
	forest loss	173.03	9.76E-11	
	mean	**2.5% CL†**	**97.5% CL†**	**Group**
p1	17.84	17.11	18.57	a
p2	17.49	16.76	18.22	a
p3	18.01	17.28	18.74	a
p4	13.83	13.10	14.55	b

	Metric	**SS Error**	**p-value**	
	forest frag 1	0.08	5.78E-02	
	mean	**2.5% CL†**	**97.5% CL†**	**Group**
p1	4.35	model was not significant, so no Tukey HSD test performed		
p2	4.35			
p3	4.35			
p4	4.32			

Table S 2.3 Continued

	Metric	SS Error	p-value	
	forest frag 2	0.01	3.12E-08	
p1	**mean**	**2.5% CL†**	**97.5% CL†**	**Group**
p2	-0.81	-0.82	-0.81	a
p3	-0.82	-0.82	-0.81	a
p4	-0.81	-0.82	-0.81	a
	-0.84	-0.85	-0.83	b

	Metric	SS Error	p-value	
	rare spp	11.11	3.02E-01	
	mean	**2.5% CL†**	**97.5% CL†**	**Group**
p1	3.55	model p-value is >0.05, so no Tukey HSD test performed		
p2	3.75			
p3	3.80			
p4	3.75			

	Metric	SS Error	p-value	
	water 1	8.95	3.51E-16	
	mean	**2.5% CL†**	**97.5% CL†**	**Group**
p1	0.24	0.07	0.40	b
p2	1.45	1.29	1.62	a
p3	1.64	1.47	1.81	a
p4	1.34	1.18	1.51	a

	Metric	SS Error	p-value	
	water 2	1.66	3.40E-14	
	mean	**2.5% CL†**	**97.5% CL†**	**Group**
p1	-0.23	-0.30	-0.16	b
p2	0.12	0.05	0.19	a
p3	0.21	0.14	0.29	a
p4	0.33	0.26	0.40	a

† These confidence limits are derived from the Tukey HSD test. They represent half the confidence interval based on the standardized error of the repeated measures ANOVA and the critical value from the studentized range distribution, which depends on the critical p value (0.05), the number of degrees of freedom in the model error (57), and the number of practices being compared (4). The Tukey HSD test could produce different confidence ranges for each pair of practices being compared, but all confidence ranges

are the same in this case because we have equal sample sizes (n=20) across all practices. The relationship between the model mean squared error (SS Error / 57 [deg. freedom]) and the confidence intervals when comparing two practices is.

$$\bar{x}_i - \bar{x}_j \pm q(\alpha, r, df_w)\sqrt{\frac{MS_w}{2} \cdot \left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

where $\bar{x}_i, \bar{x}_j$ are the means of impact values for practices i and j, q is the critical value from the studentized range distribution, MS_w is the mean squared error of the within-groups mean squared error, and n_i, n_j are the number of data points from practices i, j.

Chapter 3: The cost of avoiding environmental impacts from shale gas surface infrastructure at the lease-hold level

A version of this article will be submitted for peer reviewed publication.

Austin W. Milt, Tamara D. Gagnolet, and Paul R. Armsworth. "The cost of avoiding environmental impacts from shale gas surface infrastructure at the lease-hold level."

Austin Milt performed the secondary data collection, data processing, analysis, interpretation, and writing for this article. Tamara Gagnolet assisted in data collection, and both co-authors contributed intellectually to the design, interpretation, and revision of the article.

3.1 Abstract

Shale energy development is receiving increasing attention due to its potential to supply short term domestic fossil energy and due to concerns about its environmental impacts. All energy development harms the environment, but surface impacts from shale energy development might be minimized through careful spatial planning of infrastructure at the lease-hold level. Doing so would come at some financial cost. Here we estimate the relative financial cost of reducing impacts on forests, wetlands, rare species, and flowing freshwater from shale energy development within lease-hold scale sites. At a median site, up to 40% of impacts could be avoided before further avoidance became cost prohibitive. However, this aggregation conceals considerable variation among sites. Low-cost reductions in impacts are possible in many areas and not others, such that increasingly ambitious commitments to avoiding potential impacts could drive many sites out of production. Feasible regulations may be able to target one or two impacts that dominate aggregate impacts, though this depends on the choice of metrics and expected impacts in a region. Our results indicate that regulations seeking to reduce impacts from future development may be possible for moderate reductions in impacts and doing so may be relatively inexpensive, though not extremely so. Cost effective regulations will need to account for heterogeneity in the ability of and relative cost at sites to avoid impacts. Our analysis is unique in its combination of scale, comprehensiveness in surface infrastructure planning, and explicit consideration of the tradeoffs between reducing impacts and increasing construction costs. As such, we are uniquely able to inform the implications of reducing environmental impacts from shale gas surface infrastructure and regulations that seek to do so.

3.2 Introduction

Shale gas production in the U.S. has grown markedly since the start of the boom around 2008 and as of 2012 made up a larger portion of overall gas production than any other source (*Annual Energy Outlook 2014* 2014). Countries other than the U.S., most notably the U.K. (Hays *et al.* 2015) and China, are currently deciding on how to proceed with their own unconventional gas development. Rising with shale gas production and exploration are concerns about its environmental, human health, social, and economic consequences (Hays *et al.* 2015). Unconventional gas surface infrastructure can negatively affect terrestrial and freshwater biodiversity through habitat loss and fragmentation (Gillen & Kiviat 2012; Kiviat 2013; Jones *et al.* 2014) and pollution and sedimentation (Kassotis *et al.* 2013; Olmstead *et al.* 2013).

All energy infrastructure development produces environmental impacts, but the amount of impact and ability to avoid impacts are important factors in determining how much development, if any, is permissible. Types and amounts of impact will be context specific. Here, we focus on the surface terrestrial and freshwater impacts of well pads, access roads, and gathering pipelines accompanying directional drilling and hydraulic fracturing at the lease-hold level. Lease-holds are aggregations of tens of adjacent mineral and/or other subsurface rights of land parcel size into single planning units. Our sites are approximations of lease-holds. More specifically, we quantify and plan infrastructure to avoid forest loss and fragmentation, reductions in stream quality, wetland encroachment, and risk to rare species. Because we suspect most readers have not visited shale gas construction sites (Fig. 3.1), we invite the reader to imagine them as similar rural home development. Well pads are dispersed in space, separated by hundreds to thousands of meters. Access roads, which connect well pads to the existing road network, are generally short (~0.1 km here), narrow (12 m here), gravel roads. And gathering pipelines, which allow gas to be transported off-site, are similar in corridor size and straightness to buried electrical transmission lines.

When relatively large areas are planned as a single unit, there may be scope to move planned well pads, access roads, and gathering pipelines to partially avoid local impacts. The spatial planning problem this presents is not trivial to solve. A cost- or impact-minimizing configuration of wells, well pads, access roads, gathering pipelines, and other infrastructure requires the spatial coordination of all these infrastructures. Here, we present an analysis using spatial optimization software that plans surface infrastructure locations with the interdependence of infrastructure in mind, and does so while reducing potential impacts on the environment.

Fig. 3.1. Well pads (rectangular clearings), access roads (linear clearings in bottom-right), and gathering pipelines (other linear clearings) pose many impacts on habitats in the Marcellus formation, including forest and wetland loss and fragmentation, displacement of species of conservation concern, erosion, and freshwater sedimentation and fragmentation. Courtesy of M. Godfrey, The Nature Conservancy.

Reducing impacts from surface infrastructure is likely to increase construction costs. While infrastructure planning practices will vary by place and planner, planning is unlikely to fully assimilate environmental objectives. Further, costs and impacts may be negatively correlated in space, such that reducing one increases the others, as is commonly the case with multiple spatial objectives (e.g. Raudsepp-Hearne, Peterson & Bennett 2010; Ruijs *et al.* 2013; Qiu & Turner 2013). As with other financially focused endeavors, infrastructure planning will focus on minimizing costs while adhering to constraints imposed by regulation and construction limitations, some of which are oriented towards protecting the environment. More environmentally-oriented planning will move the gas industry away from the financial bottom-line. Therefore, quantifying the costs of reducing environmental impacts may be vital to changing industry practice. In Pennsylvania where our study concentrates, directional drilling and hydraulic fracturing are the most expensive parts of development (Hefley & Seydor 2015), but these costs do not vary much in space under normal development conditions. We concentrate on major costs that vary with the spatial locations of surface infrastructure: moving earth, clearing land, stream crossing infrastructure, and construction materials plus any associated labor (Triana energy, *pers. comm.*).

With the quantification of site-level costs of reducing environmental impacts as a goal, we present a case study of environmentally-oriented shale energy surface infrastructure planning in Pennsylvania. In particular, we (1) developed a novel, advanced spatial optimization algorithm to plan well pad locations and access road and pipeline routes at 84 sites in 5 counties in Pennsylvania (Fig. 3.2), and (2) quantified the tradeoff between reducing environmental impacts and increasing construction costs of alternative development plans. We aggregated the site-level tradeoffs between impacts and costs to arrive at a general tradeoff graph for our Pennsylvania sites that shows the relative cost of reducing impacts by a specific amount.

When the results of our analysis revealed overall promising but heterogeneous potential to reduce impacts at reasonable costs, we further explored causes for site-level differences in cost-impact tradeoffs. There are many possible sources of heterogeneity of costs and impacts across sites, all of which play out by affecting two things: the cost of moving infrastructure relative to the resulting change in impacts and the number of alternative configurations of infrastructure. For instance, larger sites may have more potential locations for infrastructure and thus increase the amount that impacts can be avoided. Some impacts may dominate others in a site, but may be limited in the potential to avoid them, and this would result in little scope to reduce impacts. Some impacts may trade off with others, such that moving infrastructure has

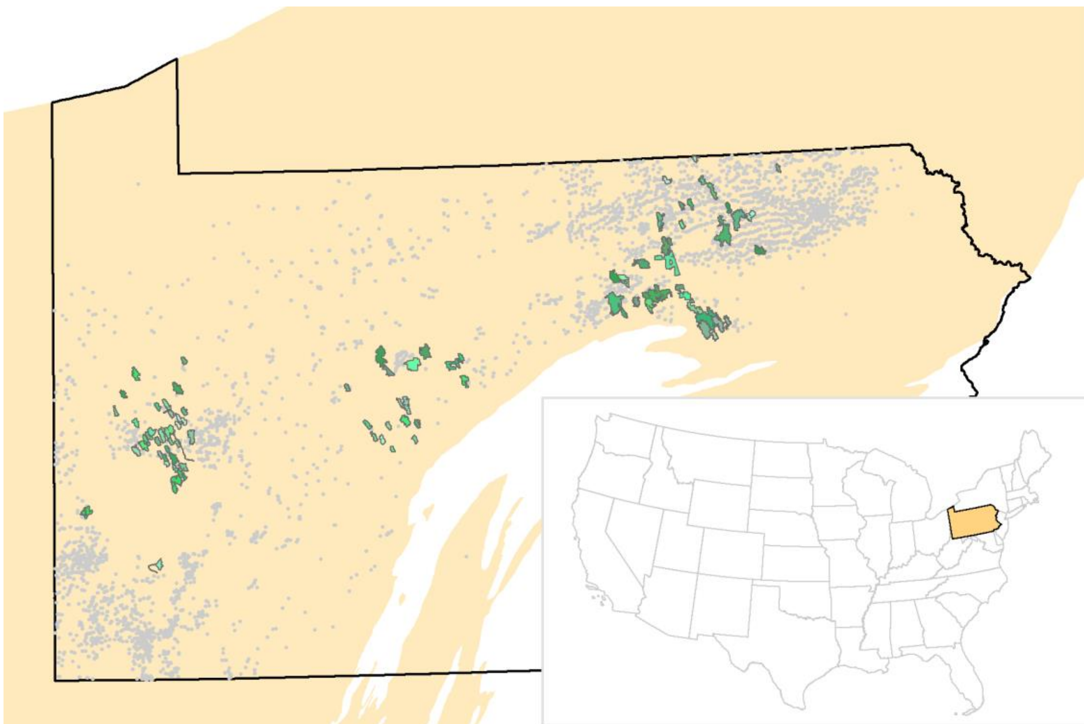


Fig. 3.2. Development sites (green polygons), with grouped well permits (gray) and the Marcellus shale play (beige). Sites were derived by overlaying production units on existing well permits and taking contiguous land parcels under those production units by a single operator. In all, 84 sites were developed.

little effect on aggregate impacts. Finally, there may be attributes of the financial bottom-line configuration of infrastructure that prevent large reductions in potential impacts, such as that configuration is already in a low-impact area. We explore all these possibilities and others.

3.3 Methods

We used Pennsylvania as a representative state for shale energy development in the eastern U.S. Over 9,620 horizontal wells were drilled in Pennsylvania between 2008 and 2014 according to the Pennsylvania Department of Environmental Protection's permit reporting database. As we discuss in Chapter 2, Pennsylvania is an ideal state to study development because of the large amount of development that has occurred, especially since it overlaps with areas of high conservation priority.

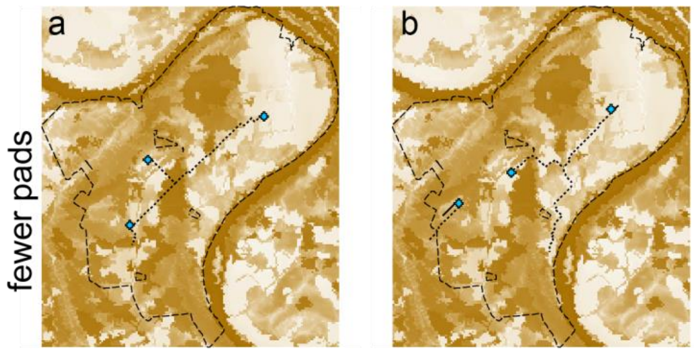
We planned and analyzed shale energy development at the lease-hold scale. This is the scale at which well pads, access roads, and gathering pipelines are currently planned. As such, it is the planning scale which most directly affects the terrestrial and freshwater environment. The lease-hold scale is larger than parcel scale and tends to be several hundred to 10,000 hectares.

Our sites were derived from locations of horizontal wells drilled from 2008-2013 (Table S 3.2). Grouped well points were merged into a single point to estimate locations of well pads. Each well pad point was overlaid with a 6-well production unit 3352.8 m tall by 914.4 m wide (3000 by 11000 ft) and rotated 27° counter-clockwise (Triana Energy, LLC, *pers. comm.*). The Production Unit is the area under each well pad that can be drained of its gas when using a six-well configuration. The angle of rotation matches the grain of the shale. Though this angle changes from place to place, it takes geologic surveys to know what the angle should be, and so we use an angle representative of examples we have seen from the gas industry. Production units were overlaid on land parcels. The set of contiguous land parcels shared by the production units of a single operator became one site. In all, this process produced 176 sites, 84 of which the software was successfully able to place infrastructure in (Fig. 3.2). Those areas which the software failed to place infrastructure in typically had no feasible layouts possible because of lack of road or pipeline access to potential well sites or existing road and pipeline networks.

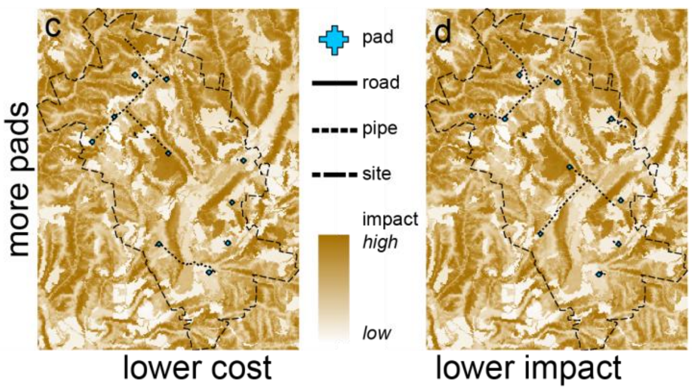
Several additional datasets are needed to place infrastructure in a site, including environmental data that inform infrastructure constraints, impacts, and construction costs (Table S 3.2). One such dataset is the existing pipeline network, which serves as connection points for gathering pipelines. In our experience, there are no sufficiently complete pipeline datasets that can be acquired for a state-wide, multi-company, site-level analysis such as ours. Instead, we used an admittedly sparse pipeline dataset and rather than force gathering pipelines to connect to existing pipelines far outside of the development boundary, we required they connect to a portion of the development boundary in the direction of the nearest existing pipeline. This served our analysis well, since our focus is on costs and impacts within single sites.

The *Impact Score* of a layout is a weighted sum of individual, normalized impact metrics (p. 94) and which we used to place infrastructure. We used five metrics in our planning and analyses. First we calculated the amount of forest acreage lost (*forest loss*) by development of forest pixels. Second, we calculated the total edge-to-area ratio (*forest frag.*) of forest after construction as one measure of forest fragmentation. Third, we calculated wetland encroachment (*wetlands*) as the percent of a 61-91 m (200-300 ft) buffer around wetlands occupied by infrastructure. Fourth, we calculated potential sedimentation in water bodies (*sediment*; p. 98). Fifth and finally, we calculated the expected impact on rare species (*rare spp.*) as the expected number of known rare species occurrences impacted by infrastructure based on habitat associations across the state. For this analysis, impact metrics were weighted equally, except for forest acreage lost and forest perimeter to area ratio. Each of these received half the weight of the others such that each category of impacted features was given equal weight. We describe how impact metrics were normalized on p. 108. Bungee – spatial planning software described in the next paragraph – allows users to adjust the weights of impact metrics, thereby tailoring the behavior of the tool to the priorities of stakeholders in the planning area. Such weights could come from valuation exercises or other studies (e.g. Banzhaf *et al.* 2014).

We developed novel spatial planning software – Bungee – with the chief aim of informing the gas industry of ways to reduce environmental impacts from surface infrastructure at reasonable costs. Bungee stands for “Balancing Unconventional Natural Gas Extraction and the Environment,” but we use it colloquially without capitalization. Bungee performs spatial optimization of surface infrastructure at the lease-hold level (p. 107). Bungee attempts to minimize environmental impacts from infrastructure while limiting the total construction budget. For a single site, Bungee proposes multiple alternative layouts – configurations of well pad locations and access road and gathering pipeline routes – that differ in their impacts and costs (e.g. Fig. 3.3). Associated with an infrastructure layout are summary statistics about the individual impact metrics over which the software optimized as well as the estimated



Metric	Layout a	Layout b
Impact Score	2.08	0.78
Construction Cost ($M)	3.9	5.4
Cost per Pad ($M)	1.3	1.8
rare spp. (spp. × 10^{-3})	0.57	0.00
wetlands (%)	0.01	0.11
sediment ($t \cdot y^{-1}$)	13.4	0.1
forest loss (ha)	2.0	1.1
forest frag. ($k \cdot km^{-2}$)	18.7	18.1



Metric	Layout c	Layout d
Impact Score	2.16	0.90
Construction Cost ($M)	12.1	14.9
Cost per Pad ($M)	1.2	1.5
rare spp. (spp. × 10^{-3})	1.09	0.14
wetlands (%)	0.05	0.07
sediment ($t \cdot y^{-1}$)	124.9	0.0
forest loss (ha)	6.6	3.9
forest frag. ($k \cdot km^{-2}$)	13.6	14.1

Fig. 3.3. Example layouts and their estimated impacts produced by our infrastructure planning software. The choice of layouts was made to illustrate how layouts differ spatially when going from (left column) lower costs and higher impacts to (right column) higher costs and lower impacts. (table) Spatial differences in layouts lead to differences in impact and costs. Individual impact metrics (p. 94) are *rare spp.* – expected number of known rare species locations encountered, *wetlands* – percent of buffer around wetlands occupied by infrastructure, *sediment* – sediment load in moving water bodies by disturbing soil, *forest loss* – area of forest cleared, *forest frag.* – forest edge-to-area ratio after development. *Impact Score* is an aggregate of the individual impact metrics. *Impact Scores* at different sites are not comparable. We use a monotonic, non-linear color scaling on the impact surface to enhance heterogeneity of impact values.

construction costs. Impact metric values and costs are based on GIS layers and consultation with The Nature Conservancy and the gas industry. Our main gas industry contact for this study was Triana Energy, LLC, which is a privately held oil and gas exploration and production company based in Charleston, West Virginia, USA and operating in both West Virginia and Pennsylvania. We used Triana's experience with shale energy development to inform the dimensions of production units, construction limits such as road grades, and surface infrastructure planning more generally.

Infrastructure layouts were planned using our hierarchical, heuristic, least-cost based spatial optimization software, Bungee. Before placing well pads, access roads, and gathering pipelines, Bungee determines both the number and approximate locations of well pads in a site by maximally packing production units within the site polygon. As such, the final number and locations of well pads was not directly determined by the existing well permits. To pack production units into sites, Bungee iteratively adds production units to the site and uses a simulated annealing algorithm to make room for new production units. Production units were allowed to partially overlap other production units as well as go outside the site as long as the total unused area of each production unit did not exceed 20%. We further explain the packing process on p. 109 and Bungee documentation (available upon request).

Infrastructure layouts were placed after production units, which determine the approximate areas where well pads will go. Each infrastructure layout consists of a set of pad locations, access road connection routes, and gathering pipeline connection routes (e.g. Fig. 3.3). The infrastructure layout creation portion of Bungee relies on (1) feasible locations for infrastructure, (2) an impact objective, and (3) cost constraints. The first of these requirements is determined by regulatory setbacks and technological constraints. The impact objective is the *Impact Score*, which we have already described. The third requirement is determined by the cost of an estimated least-monetary-cost layout. All layouts are produced by first proposing locations of well pads. Once locations are chosen, well pads become terminals for roads and pipelines. Roads and pipelines are routed using a least-cost path algorithm that is a modified version of Dijkstra's algorithm (Dijkstra 1959). The surface over which access roads and pipelines are routed is a spatially additive approximation of the *Impact Score*, which is non-additive in nature. Access roads and gathering pipelines must connect to their respective existing infrastructure networks. Each layout must adhere to a monetary construction budget which is iteratively relaxed to trace the tradeoff curve for a site. The optimal layout at one construction budget is determined by a genetic algorithm optimization of the locations of well pads and order of construction of linear infrastructure (p. 107).

This method will likely produce local (rather than global) optima, so we planned layouts five times for each site and took those layouts that were not simultaneously more impacting and more costly than any other layout, i.e. they are closer to the Pareto-frontier.

3.4 Results

3.4.1 Infrastructure Layouts

Using existing well locations to approximate lease-hold scale site boundaries, we planned shale energy surface infrastructure in 84 sites in 5 counties in Pennsylvania (Fig. 3.2). We were limited to five counties with shale energy development because of limits on parcel data. At each site, we used Bungee to automatically assess the tradeoffs between reducing environmental impacts and increasing construction costs. The software produced 2-20 layouts at each site. This number varies because (1) the software is heuristic, (2) the flexibility of planning at each site differs, and (3) the shape of the tradeoff curve varies by site.

There are spatial patterns in layouts across sites (Fig. 3.3). At low costs, access roads and pipelines tend to be straight (Fig. 3.3a,c) and short. As the amount spent increases, the routes of linear infrastructure begin to meander as the software tracks the lower-impact areas (Fig. 3.3b,d). These differences between layouts are often subtle, but even subtle differences can greatly affect impacts and costs. For instance, in the table in Fig. 3.3 we show how the differences in layouts (a) and (b) increase costs by ~1.5 million USD while reducing most impacts, but increasing impacts on wetlands. Sites differ by the number of pads since sites have different sizes and shapes. All layouts within a site are constrained to have the same number of pads.

3.4.2 Tradeoff Curves

We used an aggregate measure of environmental impact (p. 100) along with estimated construction costs of each layout to produce a tradeoff curve for a site (Fig. 3.4). We linearly interpolated between points corresponding to the infrastructure layouts to construct each site's tradeoff curve. Although tradeoff curves are generally thought of as smoothly concave or convex, our tradeoff curves are by nature of their construction not smooth and vary in their convexity. Discontinuity is created by the discreteness of layouts, which produce the points used to construct the tradeoff curves. The set of Pareto-improved layouts within a site tend to be separable into subsets, where each subset contains similarly planned infrastructure. Small variations on layouts produce

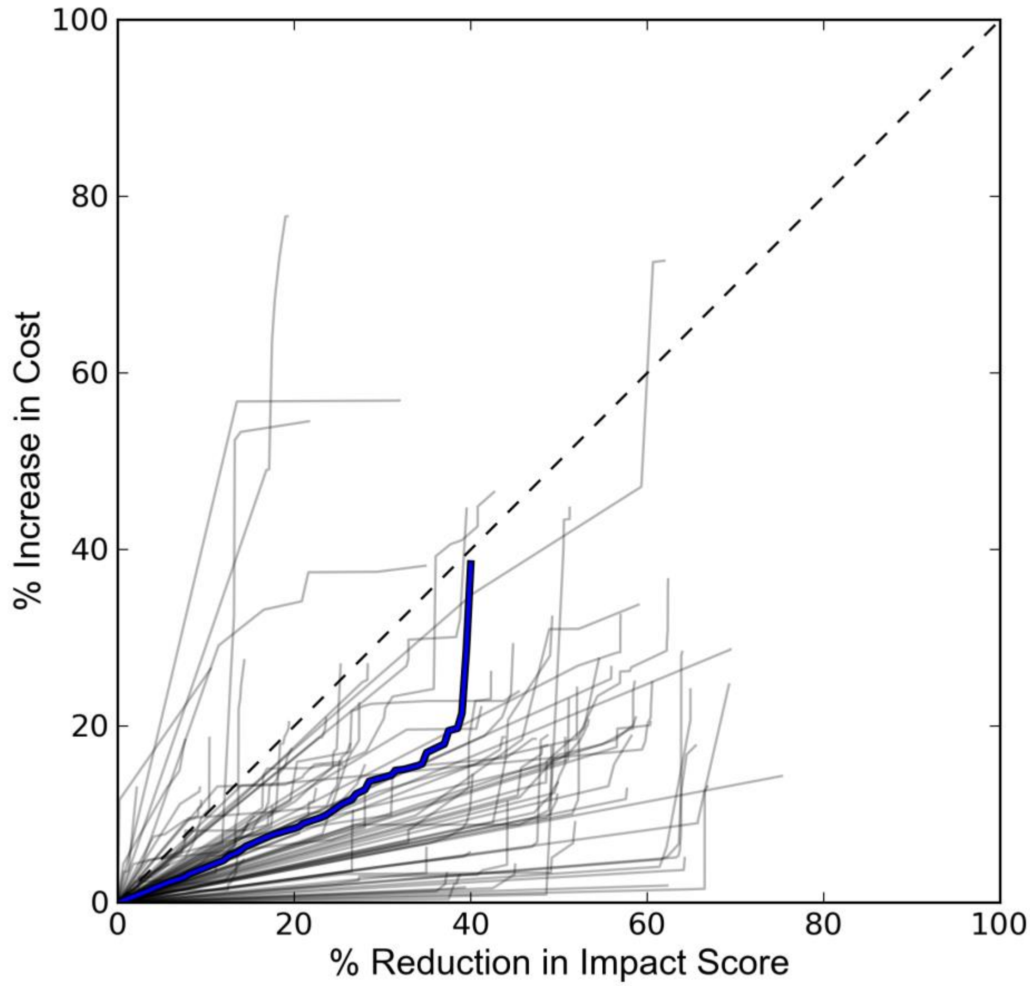


Fig. 3.4. Cost of reducing impacts at the development-area scale as the commitment to reducing impact increases. (gray lines) Tradeoff curves for individual sites, illustrating that the general shape of tradeoff curves, the maximum reduction in *Impact Score* and the resulting increase in construction costs all vary across sites. The lines are interpolated between individual layouts and are shown for clarity, while in reality the tradeoff curves will not be smooth. (blue line) Median of the tradeoff curves where truncated curves are given a very high cost at higher levels of impact reduction. This illustrates the overall cost of reducing the *Impact Score* by some amount when using one type of uniform policy across sites.

small Pareto-improvements, such that points in the tradeoff curve for similar layouts are close together. Spatial differences between subsets of layouts tend to be large and result in large changes in Pareto-optimality. Thus, tradeoff curves tend to be characterized by clusters of points separated by large gaps, i.e. are not smooth. Variations in tradeoff curve convexity arise from a complex combination of the spatial relationship between costs and impacts in a site, as well as environmental, developmental, and regulatory constraints on infrastructure placement, which in combination are difficult to tease apart. In order to compare across sites, which may have very different costs and impacts, we calculated the percent increase in construction costs and percent reduction in *Impact Score* relative to the least-cost layout; the ratio of the first value to the second gives the impact elasticity of cost for a particular layout. As such, the dotted line in Fig. 3.4 is the line of unit elasticity, along which a 1% decrease in the *Impact Score* requires a 1% increase in construction costs. Layouts above the unit elasticity line are proportionally more expensive than the resulting reduction in impact, while layouts below are proportionally cheaper for the impact reduction. As Fig. 3.4 shows, sites varied markedly by the shape of their tradeoff curves. Some curves lie primarily above the unit elasticity, while most lie below. Some curves are concave overall, while most are convex. The variation in the shapes of the tradeoff curves is to be expected, since our sites may differ by how constrained infrastructure locations are, to what degree impacts and costs vary over the site, and the relative magnitude of change in impacts and costs brought on by moving infrastructure (i.e. % change). Below we do more analyses to understand what caused the general patterns we observed.

3.4.3 Cost of Reducing Impacts

On average, we found that although large reductions in *Impact Score* can often be achieved for little cost, this was true only for some sites. In Fig. 3.4 (gray lines), we plot the distribution of tradeoff curves showing the cost of reducing the *Impact Score* by a specific amount. We also plot a summary tradeoff curve that aggregates across sites. The blue line in Fig. 3.4 shows the median value of the cost of reducing the *Impact Score* when we artificially assign a very high cost to reducing impacts for curves that have no layout at the current impact reduction. This line represents overall costs of reducing impacts when using a median-based uniform regulation across sites. The median increase in construction cost increases almost linearly up to ~38%, such that a given percent reduction in the *Impact Score* requires half that percent increase in costs. The median escalates quickly around 40% reduction in the *Impact Score*. Further reductions in the *Impact Score* would require an inordinate amount of money on the median site. In other words, the impact elasticity of costs is about one-half up to ~38% and becomes almost perfectly elastic after 40%. Of our 84 sites, 43 could achieve this 40% reduction in impact at a cost of 0.8-3.8 million USD per pad.

There was considerable variation in site-level ability to and relative cost of reducing potential impacts. Note that raw *Impact Scores* cannot be directly compared across sites, so we summarize by percentage changes relative to the least-cost layout. The site with the largest potential to avoid impacts could reduce the *Impact Score* up to 75% for 4.1 million USD (1 million USD per pad). At the other extreme, another site could reduce the *Impact Score* only 2% at a cost of 7.7 million USD (1.1 million USD per pad). Most sites (75 of 84 or 89%) had tradeoff curves mostly below the line of unit elasticity, indicating it is relatively cheaper to reduce potential impacts by some amount in those sites. However, as can be seen in Fig. 3.4, curves below the line of unit elasticity are fairly evenly distributed in their overall shapes.

3.4.4 Understanding Causes of Heterogeneity of Tradeoff Curves

What site-level differences distinguish those sites with short versus long tradeoff curves and cheap versus expensive reductions in impacts, and which, in aggregate, lead to the median line in Fig. 3.4? We chose to explore our results in three additional ways. First, we explored how site characteristics might influence our results. Second, we looked at the breakdown of impact metrics contributing to reductions in the *Impact Score* to see which impacts are responding the most to moving infrastructure. Third, we qualitatively assessed the spatial characteristics of impact surfaces and least-cost layouts at sites to determine the amount of reduction likely at a site and thus the shape of the site's tradeoff curve.

Simple ordinary least-square regression revealed tenuous support that site attributes reflecting the flexibility of planning can predict the amount that impacts can be avoided or a combination of the cost and impact without planning infrastructure (p. 92). The strongest of these results relied on site-level cost and impact surfaces. The cost surfaces are pixel-by-pixel estimates of construction cost at a site, with one cost surface for each type of infrastructure. The impact surfaces are pixel-by-pixel approximations of the *Impact Score* made by assuming each pixel is developed independently of others (p. 108). We found that higher variation in the impact surface as well as higher variation in the ratio of impact to cost (return on investment; ROI) surfaces explained larger impact reductions and lower relative costs of avoiding impacts (p. 92 and Table S 3.1). With R^2=15% for both models, variation in the impact surface explained most of the variation in our two responses ($p<0.001$). However, there was obvious uneven sampling over the range of values of the variation in the impact surface (Fig. S 3.2). We accounted for this by splitting the data into two parts with even sampling and modeling the parts separately. We then compared the original one- and new two-part models using AICc competition, which evaluates the parsimony of models by rewarding models for higher explanation of variance and punishes them for using more parameters (Crawley 2007). Model competition supported the two-part model, which was non-significant. This result weakens support that variation in the impact surface can be used to explain our responses. Variation in the ROI surface explained 9% of the variation in our responses ($p\approx0.004$). Bonferroni correction for multiple comparisons did not reverse the significance of the ROI models. Thus, areas with higher variation in ROI will, on average, be able to reduce impacts further and at lower relative costs. Other site attributes, including site area, number of well pads, the slope and correlation between impact and cost surfaces, and existing road and pipeline density were unrelated to our responses (Table S 3.1).

Individual impact metrics differentially contributed to reductions in the *Impact Score* relative to the least-cost layout (Fig. S 3.3). Since the *Impact Score* is a sum of the (normalized) impact metrics, we can attribute changes to the *Impact Score* to changes in each of the impact metrics. Reductions in the *rare spp.* metric contributed to 40% +/- 5% (mean +/- std. err.) reduction in the *Impact Score* across sites (Fig. S 3.3). Following *rare spp.* were *sediment yield* (25% +/- 5%), *forest loss* (19% +/- 3%), *forest p2a* (-0.1 % +/- 0.2%), and *wetlands* (-1.2% +/- 0.8%). Three things explain the variable contributions of impact metrics to reductions in the *Impact Score*. First and foremost, the disparity between *rare spp.* and other metrics is likely due to the large variance of *rare spp.* over small spatial scales. The pixel values of *rare spp.* may vary multiple orders of magnitude between adjacent pixels, such that a small shift in infrastructure results in a large reduction in impacts. Couple this with the lack of a direct cost analog to the *rare spp.* metric – unlike *forest loss*, whose cost analog is forest clearing costs – and relatively inexpensive reductions to *rare spp.* can be made with small changes to infrastructure. Similarly, some of the contribution of *forest loss* can be attributed to high variance over small scales, though it has a cost analog, such that the least-cost layout already avoids those areas to some extent. Second, some impacts are small in the least-cost layout, such that there is little scope to reduce those impacts further. The *wetlands encroachment* metric is restricted to areas surrounding wetlands, which are relatively sparse. As a result, *wetlands encroachment* impacts from the least-cost layout tend to be small and as a result impact-reducing layouts have little scope to reduce this metric. Finally, but rarely, reductions in one impact metric resulted in increases in another relative to the least-cost layout, which explains why, on average, *wetlands encroachment* had a negative contribution to reductions in the *Impact Score*.

In most cases, the difference between sites that could achieve large reductions in the *Impact Score* for little cost (25 of 84 or 29%), i.e. those with tradeoff curves in the lower right of Fig. 3.4, and those sites which could do very little (17 of 84 or 20 %), could be explained by visually comparing layouts overlaid on the spatially additive impact surface. The spatially-additive impact surface approximates the *Impact Score* on a per-pixel basis, while the true *Impact Score* depends on the entire layout. From this exercise, two conditions distinguish the aforementioned groups: first, whether the least-cost layout has at least some infrastructure in high-impact areas, and second, whether there are lower-impact areas for infrastructure to be placed. Those 25 sites in the first group tended to have roads or pipelines in high-impact areas in the least-cost layout (21 of 25 sites). Much less often (4 of 25 sites) well pads, but not roads or pipelines, were in high-impact areas in the least-cost layout. Those 17 sites in the second group, which achieved either no large or very costly reductions in the *Impact Score*, tended to be constrained by impacts in one way or another. Many sites (9 of 17) were constrained by a lack of low-impact alternatives for pipelines and pads. Often, sparse existing pipeline infrastructure, to which gathering pipelines had to connect, forced gathering pipelines through high-impact areas. Alternatively or in addition, feasible well pad locations, from which gathering pipelines start and well pads are located, were in high-impact areas. In other sites (6 of 17), reductions in some impacts led to increases in others, such that the aggregate *Impact Score* did not change much. Finally, in 3 of 17 sites, the least-cost layout was already in a low-impact area such that there was little scope to reduce impacts further.

3.5 Discussion

Shale gas development will continue worldwide, but there are opportunities to reduce potential impacts through environmentally oriented planning. Here we presented an analysis that looks at the monetary cost of such planning at the site level, and found scope to reduce impacts, though doing so is not generally cheap. More specifically, we found that most sites do not have the scope to reduce the *Impact Score* more than ~40%. For those sites, reducing the *Impact Score* up to ~38% requires <20% additional investment, but further reductions become very expensive. For many of those sites that can achieve larger reductions in impact at reasonable costs, doing so is inexpensive. For instance, a reduction in the *Impact Score* of up to 60% required <5% increase in costs at several sites (Fig. 3.4). There was only marginal support that simple statistics on the spatial variation of impacts and costs within a site can be used alone to partially predict the nature of tradeoffs in a site, while other site characteristics do not appear to be informative. Some impact metrics dominated others in their contribution to avoiding aggregate impacts.

Our results can be used most directly by policy makers. There is scope to reduce the aggregate *Impact Score* up to ~38% with ~20% increase in costs when following a median-based uniform policy. However, a 20% increase in costs is not small, requiring ~400,000 USD per well pad to reduce impacts by 38% in our case study. That said, the construction costs represented here are for surface infrastructure alone, so total costs to a developer would increase by less than 20%. As such, regulations that target a 38% reduction in aggregate impacts without compensating developers need not result in a 20% increase in gas prices. Regulations that target one or two impacts independently may do fairly well at reducing aggregate impacts. We found that ~65% of reductions across sites in the *Impact Score* over the least-cost baseline were attributable to just two impact metrics, which measured impacts on high-quality habitats and freshwater sedimentation (Fig. S 3.3). Further, only very rarely did impacts trade-off with one another when compared to the least-cost layout. This is a qualitatively similar result to Chapter 2 even though the assessment and set of impacts differ somewhat. As such, it is possible that regulations could target one or two metrics to reduce aggregate impacts in a predictable way up to a point. In Pennsylvania, for instance, restrictions on the amount of infrastructure placed in high-quality habitats and high-slope areas could be feasibly implemented, since these data can be at least partially remotely sensed and impacts assessed by overlaying planned infrastructure with these data. There is already a minimum setback requirement for infrastructure placed in high-quality watersheds in Pennsylvania. Outside of our study area, it will be important to assess potential impacts before implementing such a regulation, since the necessary qualities of the impacts here may not extend elsewhere. Finally, we found that some site attributes partially explain the ability to reduce impacts at a site as well as the relative cost of doing so. However, the low (R^2~9%) explanatory power of these attributes likely precludes them from being used independently from other analyses to predict potential tradeoffs at a site.

It may be possible to regulate aggregate impacts efficiently, i.e. to simultaneously reduce multiple distinct impacts with a single regulation. If this can be done, it presents two major advantages. First, regulating aggregate impacts enables regulators and/or developers to tailor individual impacts at the site level rather than be forced to focus on impacts they may not be able to avoid. Second, a general regulatory tool for aggregate impacts could work in multiple regions where regulating single impacts might not. As our results show (Fig. 3.4), uniform policies, e.g. that require all sites reduce impacts by a certain amount, may be unnecessarily costly. More flexible regulations may be able to take advantage of the heterogeneity of site-level tradeoffs between cost and impact to reduce impacts across all sites. There are many economic tools and existing programs that can cost efficiently allocate single impacts across sites and can serve as a basis in this context (Ferraro 2008). Market-based mechanisms such as cap and trade can exploit heterogeneity across sties, incentivizing those who cannot reduce impacts by much or for cheap to compensate those who can and will. We explore this further in Chapter 4.

There are two major barriers to implementing any aggregate impact regulation in this context. First, there must be some information exchange between regulators and developers. We think at minimum that regulators must share their environmental priorities in order for developers to assess alternative infrastructure layouts that may reduce aggregate impacts. Existing command-and-control regulations do this to some degree by restricting infrastructure in areas of high environmental value. To avoid aggregate impacts, regulators would need to make explicit how infrastructure layouts are used to evaluate aggregate impacts. In addition to sharing environmental objectives, we think developers will need to share some layouts, which is already done for other permits. Regulators need such layouts to estimate impacts in the absence of additional regulations so they can effectively limit impacts. In the cap and trade system we explore in Chapter 4, regulators would not need to know construction costs, though this additional information may be helpful. Bungee, the planning software we presented here, provides an existing system for this information exchange and is set up in such a way that planning and monitoring would be very inexpensive.

The second barrier to implementing a regulation of aggregate impacts is the difficulty of finding and using a metric of aggregate impact. The *Impact Score* – the measure of aggregate impact we use here – is a site-specific metric. It lends itself well to optimization at the site-level and could be used to implement a uniform policy such as described above, but it could not be used in its present form to allocate allowed impacts across sites. We return to this point in Chapter 4, in which we faced this issue. One example of a aggregate score currently used to allocate conservation funds is the Conservation Reserve Program's Environmental Benefits Index (United States Department of Agriculture Farm Service Agency 2011). The Environmental Benefits Index is measured qualitatively and is used to rank applicants, whereas a metric better suited to our system would be calculated directly from estimated impacts and that could be traded or allocated continuously across sites.

Bungee can be used by regulators, policy makers, conservation NGOs, large land owners, and the shale energy industry to reduce potential impacts from future development. Regulators whose job it is to assess development plans for permits could use our software to evaluate the relative impacts of a proposed infrastructure layout to a lower-impact, more “ideal” layout. Then, threshold allowances could be applied to judge with some objectivity if the proposed layout meets environmental standards. Policy makers could use the software to estimate potential gains from new regulations. Applied in a new context, Bungee could inform the magnitude of subsidies or taxes to affect a target level of impact avoidance by the shale energy industry. Conservation groups who are working in collaboration with the gas industry can use Bungee to inform planning practices at a site. Such collaborations between conservation and industry may be necessary since conservation groups will be better informed about the important impacts in an area and how to evaluate them. Large landowners who have access to ArcGIS can use our software to propose alternative layouts when shale development plans clash with site features the landowner wants protected. That said, our experience has been that the shale energy industry is attuned to such conflicts and, when well informed, wants to avoid them. Finally, the gas industry is perhaps the best positioned to benefit from the use of our software. The scale at which it operates, its flexibility in use, and its incorporation of costs mean that Bungee can be an effective tool for going above and beyond for a progressive company.

Our study is a novel contribution to the shale energy policy literature. In addition, Bungee is a novel planning software in multiple regards. Our methods focus on site-level planning, which is the scale at which decisions about shale gas development most directly affect the environment. As such, we are uniquely able to inform policies and actions at this scale. Bungee attempts to simultaneously site multiple types of infrastructure with very different planning characteristics. The spatial optimization problem it solves and the methods used to do so have, to our knowledge, never been so comprehensively attempted in conservation planning. While the software is specialized to work in the shale gas planning context, the methods could be easily transferred to other development problems that involve connecting potential development sites to infrastructure networks while considering costs and impacts.

Several steps could be taken to improve future analyses such as ours. First, we feel that a more informative analysis would explicitly compare existing shale gas development to proposed layouts at the development-area scale. Such an analysis would require having site boundaries from many shale energy companies as well as more complete pipeline data. Both datasets are difficult to come by and cannot be easily derived from available remotely sensed data. As a compromise, we used publicly available well permit locations and our knowledge of drilling practices to estimate development boundaries. This enabled us to plan in locations where development has occurred, but not to compare to existing development directly. Second, the planning framework and methods in Bungee are advanced, but the software performs a fairly narrow heuristic search for potential infrastructure layouts. We were forced to compromise known optimality for reasonable run speeds. Our results must therefore be interpreted as conservative, since it is likely the layouts produced are not Pareto-optimal. Improvements to the optimization method could improve the optimality of results without prohibitively increasing the run time.

3.6 Conclusion

Shale energy development will likely play an increasingly important role in energy production over the coming decades. Our study and others like it can contribute to a conservation oriented development paradigm, in which the cost of impacting the environment is explicitly accounted for and which factors into decision making on a large scale.

3.7 Appendix

3.7.1 Additional/Detailed Methods and Results

3.7.1.1 Regressions to explain impact reductions

We wanted to know if site attributes related to the flexibility of infrastructure planning would explain how much impacts could be reduced, what this would cost, or a combination of the two. In other words, can we predict the end-points of the tradeoff curves in Fig. 3.4 without planning infrastructure? Addressing this question would allow us to understand what makes some sites able to avoid impacts cheaply as opposed to not being able to avoid impacts at all or for it to be very expensive to do so.

To answer the above question, we identified several predictors related to the flexibility of planning of infrastructure. We define flexible sites as those where there are many feasible, spatially distinct infrastructure layouts with a wide range of *Impact Scores*. An inflexible site is one where feasible layouts are tightly constrained to specific areas or where there is little heterogeneity in impacts and costs. First and second, we used the density of existing road and pipeline networks as predictors since larger/denser existing infrastructure provides more connection points for proposed infrastructure. This was calculated as the number of pixels of existing road (or pipeline) infrastructure per pixel of the site. Third, we used the variation of the pipeline impact surface across pixels (see 3.7.7.5 Estimate additive impact surfaces), since higher variation of impacts might create more alternative routes for pipelines and roads. This was calculated as $CV(I) = \sigma_I/\bar{I}$. We used the pipeline impact surface because it covered the entire site and the spatial structure of the impact surfaces are very similar across types of infrastructure. Fourth, we use the variation of the pipeline cost surface for a similar reason to the pipeline impact surface and calculated the same way. Fifth, we used variation in the ROI surface, which is the ratio of the impact surface to the cost surface. We expected the ROI surface to be the strongest predictor since Bungee is planning according to impacts and costs (see 3.7.7.6 Create final layouts and 3.7.9.3 Linear infrastructure route optimization). Sixth, we used the correlation between the impact and cost surfaces. Highly negatively correlated impact and cost surfaces might lead to costly avoidance of impacts, whereas highly positively correlated surfaces might lead to an inability to reduce impacts since the least-cost layout would already be in low-impact areas (Babcock *et al.* 1997). Seventh and eighth, we used the size

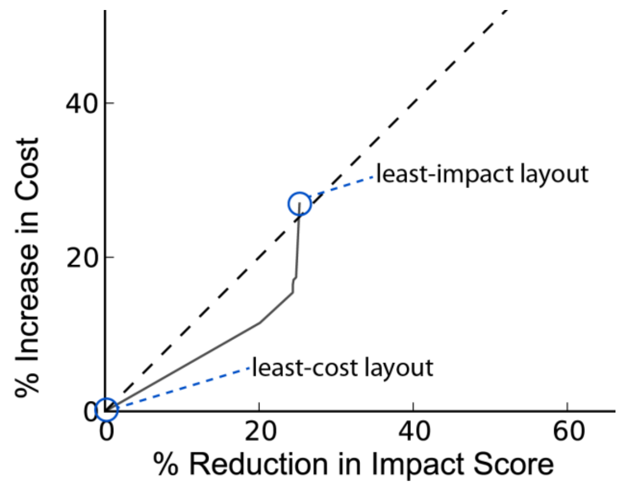


Fig. S 3.1. Example showing how response variables for regressions were derived. The relative cost (Y) and impact reduction (X) of the least-impact layout are calculated proportionally to the least-cost layout represented by (0, 0).

(in hectares and number of well pads) of sites, since larger areas, by definition, have more locations for infrastructure to be placed.

For the regressions, two response variables were identified that describe attributes of the least-impact layout relative to the least-cost layout. These two layouts form the end-points of the tradeoff curves in Fig. 3.4 (see Fig. S 3.1 for example). First, we used the impact reduction of the least-impact layout measured as the X position of the least-impact layout (Fig. S 3.1). Second, we calculated the ratio of the cost increase and impact reduction of the least-impact layout, also known as the impact elasticity of cost of the least-impact layout (Y/X in Fig. S 3.1). As Table S 3.1 below shows, four models were significant ($p<0.05$). Visual inspection of the plots using $log_{10}(CV(impact))$ as a predictor (Fig. S 3.2) show uneven sampling across the values of $log_{10}(CV(impact))$. We explored the effects of treating the dataset as composed of two parts, each with even sampling across $log_{10}(CV(impact))$. To do so, we split the dataset at $log_{10}(CV(impact)) = -1.0$ and ran separate ordinary least-squares regressions on each part. In both cases, the two-part model had an AICc score more than five units below the one-part model (titles in Fig. S 3.2), indicating a two-part model is more parsimonious. Combined with the non-significance of the two-part model, this result weakens support that $log_{10}(CV(impact))$ explains variation in our responses.

3.7.1.2 Summary of metric contributions to reductions in Impact Score

We wanted to see if reductions in the Impact Score relative to the least-cost layout were being driven mainly by one or a few impact metrics. If so, then regulations concentrating on those most responsive metrics could reduce aggregate impacts. In addition, if there was little scope to reduce one of those metrics at a site, then there may be little scope to reduce the Impact Score. We also wanted to see if some metrics tended to trade off with others, which would reduce the effectiveness of some impact-specific regulations and would explain why some sites have short tradeoff curves.

To address these interests, we broke down the Impact Score for each infrastructure layout into its weighted, normalized impact metric constituents. We calculated the proportional change of a metric in each layout from the least-cost layout. Next, we divided this proportional change for each metric by the total change across metrics for the layout to get each metric's contribution to the reduction in Impact Score. To partially control for site differences, we averaged across layouts within a site to get the site's mean contribution to reductions in the Impact Score (gray circles in Fig. S 3.3). We then averaged across sites to get the mean contribution of each impact metric across all sites (bars in Fig. S 3.3).

3.7.2 Impact Metrics

This section contains detailed descriptions of each impact metric used in the analysis, including how each metric is calculated. Metrics were chosen based on their likelihood to occur, their magnitude when they do occur, and their priority for Pennsylvania.

3.7.2.1 forest frag

This is a measure of forest fragmentation, specifically edge-to-area ratio. We chose this metric because of the high potential for forest fragmentation caused primarily by new pipelines and access roads. We calculate *forest frag* as follows:

1. forested areas overlain by proposed infrastructure are removed from the *forest* raster (Table S 3.2) the number of pixel edges joining a forested area and non-forested area are tallied
2. the total number of forested pixels are tallied
3. *forest frag* = [step 2] / [step 3]

3.7.2.2 forest loss

This is a measure of direct forest habitat loss from clearing forested areas. We chose this metric because of the high potential for forest loss caused by surface development. We calculate *forest loss* as the total number of forested cells in the *forest* raster overlain by proposed infrastructure (Table S 3.2). This value is converted into hectares for statistical analyses and reporting.

Table S 3.1. Simple ordinary least-squares regressions of attributes of the least-impacting layout relative to the least-cost layout against site attributes. Predictors are *log_{10}(CV(impact))* = log-transformed coefficient of variation of the pipe impact surface, *CV(ROI)* = coefficient of variation of ROI surface, and *CV(cost)* = coefficient of variation of cost surface. Responses are *impact reduction* = percent impact reduction of least-impact layout as in Fig. 3.4, *log_{10}(elast(impact, cost))* = log-transformed ratio of least-impact layout cost to least-impact layout impact relative to least-cost layout in Fig. 3.4, *cost increase* = percent cost increase of least-impact layout as in Fig. 3.4, *corr(impact, cost)* = Pearson's correlation of the pipe impact surface with the pipe cost surface, *CV(cost)* = coefficient of variation of the pipe cost surface, *pipe density* = density of existing pipeline network (# pipe pixels / area of site in pixels), *road density* = density of existing road network (# road pixels / area of site in pixels), *log_{10}(# pads)* = log-transformed number of well pads in site, and *log_{10}(area)* = log-transformed area of site in hectares. Several variables were log-transformed to increase normality of regression residuals.

Predictor	**Response**	**Rationale**	**Slope**	**Std. Error**	**R^2**	**p**
log_{10}(CV(impact))	impact reduction	layouts produced at sites with larger variation in impact will span a greater range of impact including relatively lower impact options	0.25	0.066	0.15	0.0003*
log_{10}(CV(impact))	log_{10}(elast(impact, cost))	the average – represented by least-impact layout – unit cost of reducing impacts at a site is constrained by impact, and layouts produced at sites with larger variation in impact will span a larger range of impact	0.21	0.055	0.15	0.0003*
CV(ROI)	impact reduction	the maximum amount of impact reduction at a site is constrained by low cost options for reducing impacts	0.77	0.264	0.09	0.0044
CV(ROI)	log_{10}(elast(impact, cost))	the average – represented by least-impact layout – unit cost of reducing impacts at a site is constrained by low cost options for reducing impacts	0.64	0.221	0.09	0.0049

* Results for one-part model. Two-part model favored by AICc competition was not significant.

Predictor	**Response**	**Rationale (all results non-significant)**
corr(impact, cost)	impact reduction	sites where low-cost pixels are also low-impact will not be able to reduce impacts much because the lowest-cost layout will also be low-impact

Table S 3.1 Continued

Predictor	Response	Rationale (all results non-significant)
CV(cost)	impact reduction	layouts produced at sites with larger variation in cost will span a greater range of impact including relatively low-impact options
pipe density	impact reduction	variations on layouts are constrained by connection points for gathering pipelines, so sites with more pipeline connection options will be able to reduce impacts more
road density	impact reduction	variations on layouts are constrained by connection points for access roads, so sites with more road connection options will be able to reduce impacts more
$\log_{10}$(# pads)	impact reduction	larger sites (that support more pads) will have a larger number of feasible layouts that can span a larger range of impacts, including relatively low-impact options
$\log_{10}$(area)	impact reduction	larger sites will have a larger number of feasible layouts that can span a larger range of impacts, including relatively low-impact options
CV(cost)	$\log_{10}$(elast(impact, cost))	the average – represented by least-impact layout – unit cost of reducing impacts at a site is constrained by cost, and layouts produced at sites with larger variation in cost will span a larger range of costs
pipe density	$\log_{10}$(elast(impact, cost))	variations on layouts are constrained by connection points for gathering pipelines, so sites with more pipeline connection options will be able to reduce impacts cheaply
road density	$\log_{10}$(elast(impact, cost))	variations on layouts are constrained by connection points for access roads, so sites with more road connection options will be able to reduce impacts cheaply
$\log_{10}$(# pads)	$\log_{10}$(elast(impact, cost))	larger sites (that support more pads) will have a larger number of feasible layouts that can span a larger range of impacts and do so at lower costs
$\log_{10}$(area)	$\log_{10}$(elast(impact, cost))	larger sites will have a larger number of feasible layouts that can span a larger range of impacts and do so at lower costs

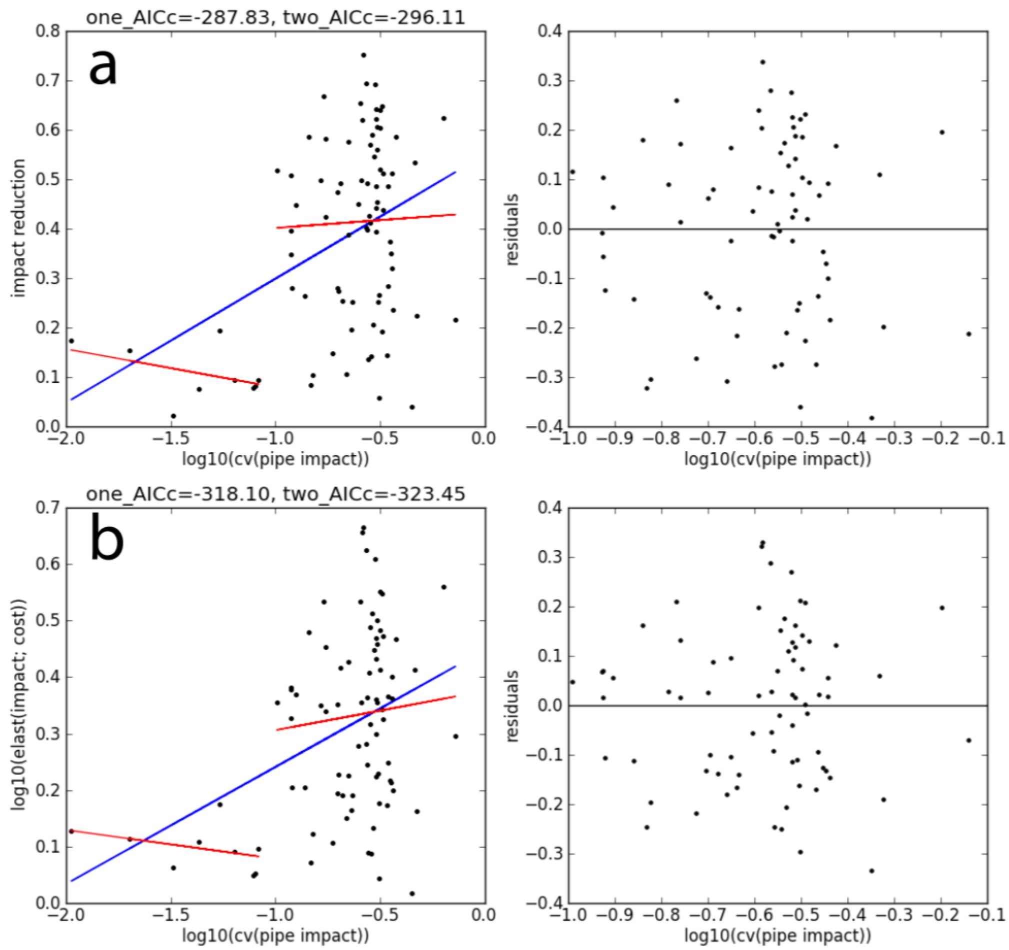


Fig. S 3.2. Illustrating uneven sampling present in regressions of the variation in the additive impact surface with a) maximum impact reduction, and b) the impact elasticity of cost. Blue lines show regression lines for full dataset, while red lines show regression lines when splitting data into two parts to explore effects of treating data in two separate parts. Two-part regressions are not significant. Residuals on right are for the full dataset.

3.7.2.3 sediment yield

This is a measure of potential sediment mobilization and load on streams caused by disturbing soil during construction. We chose this metric because of the high biodiversity value of streams in Pennsylvania and their sensitivity to changes in quality. This sediment yield rasters differ between well pads and linear infrastructure. For linear infrastructure, sediment yield in a pixel is calculated as the number of metric tons per year of sediment resulting from replacing existing land cover with a well pad for each individual pixel. This was modeled using the methodology described in Fernandez et al., 2003.[1] For well pads, the sediment yield in a pixel is the sum of the previous raster's pixels covered by a well pad centered on that pixel.

3.7.2.4 wetlands encroachment

This is a measure of indirect impacts on wetlands through degradation or removal of buffering habitat. We chose this metric because of the importance of wetlands in Pennsylvania. We calculate *wetlands encroachment* as follows:

1. *wetlands* (Table S 3.2) are buffered by different amounts for each type of infrastructure. For well pads, the buffer is 61 m. For roads and pipelines, the buffer is 91 m. This produces three presence-absence rasters that include the buffer and original wetlands.
2. Each present pixel from step 1 is given the same value, which is the percentage of all buffers that each pixel represents, such that the sum across all three rasters is 100%.
3. Each infrastructure type is overlaid with its corresponding raster from step 2 and the sum of those pixels is calculated.
4. *wetlands encroachment* = [sum of three values from step 3]

3.7.2.5 rare spp.

This is a measure of risk to rare species in Pennsylvania, or alternatively to the habitats in which rare species are found. We chose this metric because of the high priority of rare species in Pennsylvania. This metric is calculated as follows:

[1] Fernandez C, Wu JQ, McCool DK, Stöckle CO (2003) Estimating water erosion and sediment yield with GIS, RUSLE, and SEDD. *J Soil Water Conserv* 58 (3):128–136. Available at: http://www.jswconline.org/content/58/3/128.abstract.

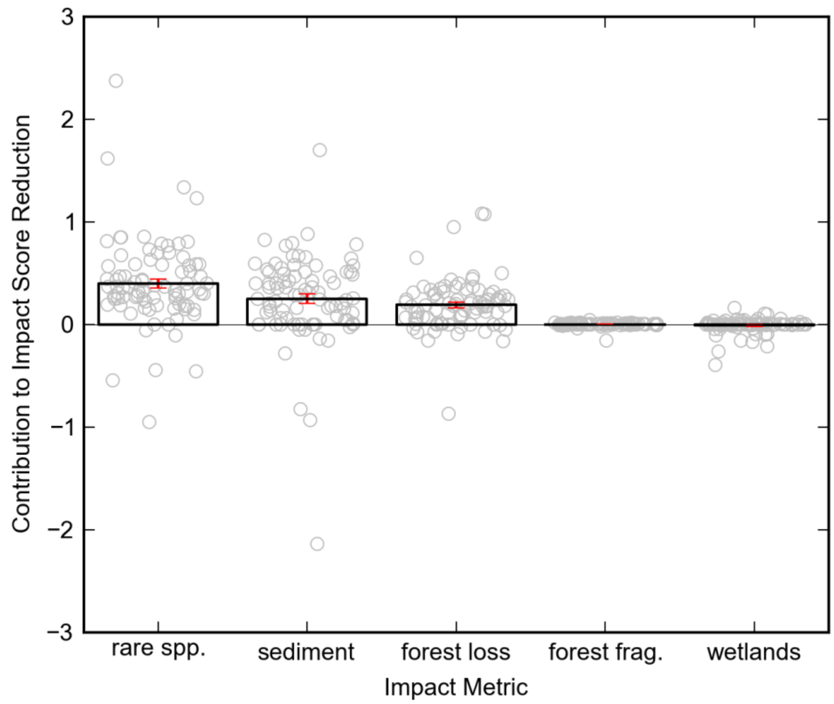


Fig. S 3.3. Contributions of individual impact metrics to reductions in the Impact Score relative to the least-cost layout (n=84 per bar). Bar heights are means of site mean contributions (gray circles) along with the standard error (red error bars). Significant differences are *rare spp.* > *sediment* ≈ *forest loss* > *forest frag.* ≈ 0 > *wetlands*. Jitter was added to gray circles to ease interpretation.

- A map of habitat types from the Northeastern Terrestrial Habitat Mapping Project (Table S 3.2) was overlaid with element occurrences from the Pennsylvania Natural Heritage Program (Table S 3.2). Each habitat type was assigned the number of element occurrences it contained. This was done without regard to species identity. This number was then divided by the total area of the habitat type to reduce areal effects. Finally, values were multiplied by one million to aid understanding.
- *rare spp.* = [sum of pixels of raster from step 1 overlain by any infrastructure]

3.7.2.6 Impact Score

This metric is an aggregation of the others and is used for the optimization of infrastructure layouts, as well as a summary of the total potential impact of a layout. Because individual impact metrics may be spatially non-additive, the *Impact Score* is generally spatially non-additive. In other words, it cannot be calculated exactly until all infrastructure has been planned. The *Impact Score* is basically a weighted sum of the individual impact metrics. However, several requirements of the optimization process cause the final form of the equation for the *Impact Score* to be more complex. Below, we describe each of these requirements and describe how they affect the *Impact Score*. Further down we summarize and show mathematically how each requirement affects the *Impact Score*.

Strictly positive values

Dijkstra's least-cost-path algorithm, on which road and pipeline route planning is based, requires strictly positive values in each pixel of the planning surface. We required the *Impact Score* to be non-negative. When using the spatially additive approximation of the *Impact Score* to plan road and pipeline routes (see Additive Layout Creation Methods), we added a small constant to all pixels so that any zero values would become positive.

Direction of impact undetermined

In this study, all metrics are worse when the values are larger. However, built-in to Bungee is the potential to handle metrics that are worse when they are smaller. We assumed that 'impact' is always either monotonically increasing or decreasing with metric values. To handle the direction of impacts somewhat generally, Bungee compares metric values to a baseline value, which is the best-case or, alternatively, the no-impact scenario. In this analysis, we use a no-impact baseline for all metrics, such that the baseline values for all metrics are zero.

Metric scales differ

Because metric scales differ, they cannot be directly combined. Instead, they are first normalized by dividing by a normalization constant (see 3.7.7.4 Determine normalization constants), which puts them on a similar scale close to one.

Impact priorities differ

The conservation priorities of impacts are different, since some are more important to stakeholders than others. We account for this in Bungee by allowing for metric weights to be specified. In this study, we weight each category of impacted features equally: *forest loss* and *forest frag* each received a weight of 0.5, while all others received a weight of 1.0. Bungee automatically makes all impact weights sum to 1, so this results in weights of 0.125 and 0.25.

Normality of values

Although not a requirement of the optimization process, impact metrics that are more normally distributed – as opposed to highly skewed or multi-modal – enable the optimization algorithm to explore more of the solution space. We used a natural-log transform to help make metrics more normally distributed.

Impact Score calculation

With the above considerations in mind, the *Impact Score* is calculated as follows:

$$Impact\ Score = \sum_{i \in I} w_i \ln\left(\frac{|f(\boldsymbol{X}, \boldsymbol{N}_i) - b_i| + 1}{|m_i - b_i| + 1} + 1\right)$$

where impact i in the set of impact metrics I is evaluated using the function f. The function f, as described for each impact in the sections above, operates on the Boolean rasters and well pad centroids $\boldsymbol{X}$ for the infrastructure layout and the set of additional inputs N_i needed for the metric. The metric is compared to its baseline value, b_i, and that quantity is divided by the normalization constant, m_i. Finally, the metric is weighted by its priority w_i.

The considerations described in the previous subsections of 3.7.2.6 Impact Score enter into the formula in various places:

- w_i comes from Impact priorities differ
- b_i comes from Direction of impact undetermined

- m_i comes from Normality of values
- The natural-logarithm comes from Normality of values
- The three ones come from Strictly positive values. The one in the denominator avoids divide-by-zero issues. The one in the numerator reduces skewing caused by the one in the denominator. The furthest-right one ensures the normalized metric will be non-negative after the natural-log transform.

3.7.3 Cost Metrics

There are many costs associated with shale energy surface infrastructure development. For our analysis, the important ones are those that vary in space or with changing amounts of infrastructure, since this produces variable costs with changing infrastructure locations. Bungee also includes other non-spatial, large costs associated with surface infrastructure. This serves two purposes: to produce more accurate estimates of development cost and two more realistically bound construction costs when reducing impacts.

In describing the cost metrics, we will refer to both a standard cost per some unit and the per-pixel cost. The former cost is easier to interpret and translates directly from the construction action incurring that cost, while the latter is used by Bungee to estimate the cumulative cost of a layout by summing over the pixels occupied by infrastructure. In general, we assume that well pads pixels occupy the entire 30 × 30 m pixel, while road and pipeline pixels occupy the entire length of a pixel, but only part of its width. The width of road and pipeline corridors was set to 15 m such that each type of infrastructure occupies 50% of a pixel. We make note where these assumptions do not hold.

Unlike the impact metrics, cost metrics are assumed to be a) spatially additive, and b) independent across infrastructure types. As such, each type of infrastructure has its own cost surface, which is a summation of the individual cost metrics. All cost values are taken from discussion and data sharing with Triana Energy, LLC.

3.7.3.1 base

Base costs are the costs of construction materials and labor beyond all other costs. Generally we think of *base* costs as the cost of infrastructure in flat areas without trees or other significant features that require special construction. This cost includes things like the cost of gravel, pad materials, etc. This cost surface does not vary in space, but is different for each type of infrastructure. The *base* costs of well pads, access roads, and gathering pipelines was 35.84 USD m^{-2} (32,256.00 USD $pixel^{-1}$), 18.84 USD m^{-2} (8,478 USD $pixel^{-1}$), and 656.18 USD m^{-1} (19,685.4 USD pixel^{-1}) respectively. Note that pipeline *base* cost is defined on a per-length rather than per-area basis.

3.7.3.2 slope

Slope costs are the costs of construction and materials over and above base costs when developing on non-flat areas. These costs typically just apply to roads and pipelines, which may be developed in non-flat areas. They are analogous to cut-and-fill costs for well pads, although the two types of costs may only partially overlap. The *slope* costs are defined by how much additional cost is incurred by developing on an increasingly steep slope. The *slope* costs for both access roads and gathering pipelines was 0.22 USD $\%^{-1}$ m^{-2} (100.90 USD $\%^{-1}$ pixel^{-1}). The *slope* cost is multiplied by a percent-slope raster to get the spatially variable *slope* cost surface.

3.7.3.3 forest clearing and timber

Forest clearing and timber costs are the costs of clearing trees and reimbursing landowners for foregone timber profits. We used the binary *forest* (Table S 3.2) raster to assess whether a cost is incurred and an equal cost per-area (16,679.61 USD ha^{-1}) for all types of infrastructure, though the size of infrastructure types produces different per-pixel costs. The *forest clearing and timber* costs for well pads, access roads, and gathering pipelines was 1,501.16 USD pixel^{-1}, 750.58 USD pixel^{-1}, and 750.58 USD pixel^{-1} respectively.

3.7.3.4 stream crossing

Stream crossing costs include the cost of materials, labor, and permits for construction of stream-crossing infrastructure such as culverts or small bridges for access roads and pipelines. We assumed that each pixel of water incurred the same *stream crossing* cost. The *stream crossing* cost for both access roads and gathering pipelines was 50,000 USD crossing^{-1} (50,000 USD pixel^{-1}).

3.7.3.5 cut and fill

Cut and fill costs are the equipment rental and labor costs of moving soil around within a site in order to flatten construction areas. Though the construction of access roads and gathering pipelines does require cut and fill, we assumed these costs were sufficiently described by the *slope* costs. As such, only well pads were assumed to incur a *cut and fill* cost, which was set to 6.54 USD m^{-3} (5,885.74 USD m^{-1} pixel^{-1}). This cost was then multiplied by a cut-and-fill depth raster to get the spatially varying *cut and fill* cost. The cut-and-fill depth raster is based on the *DEM* raster (Table S 3.2) and was calculated as the height of soil added to or removed from a pixel to bring it to the average height of its eight neighbors.

3.7.3.6 pad and well permit

Pad and well permit costs are the costs of permitting wells on a well pad. In our analysis, each unique well pad is always assumed to have the same number of wells regardless of where Bungee proposes to place it. As such, this cost does not vary in space. However, it does vary with the number of wells. We set the *well permit* cost to be a baseline of 65,000 USD pad^{-1} with an additional 6,000 USD $well^{-1}$. In our analysis, all well pads had six wells such that the total *pad and well permit* cost was 101,000 USD pad^{-1}.

3.7.4 Setbacks and Other Restrictions

Pennsylvania regulations prevent the placement of some surface infrastructure in certain areas. Often times it is difficult to assess *a priori* whether a particular restriction will be activated, and in general, most such restrictions can be waived upon application from the developer. Regardless, we implemented several setbacks and other restrictions under the assumption that developers would follow existing regulations without applying for exceptions. In all, we implemented the following setbacks and other restrictions:

- pads 30 m from water bodies (Table S 3.2)
- pads 30 m from wetlands (Table S 3.2)
- pads 100 m from development-area boundaries
- pads 152 m from buildings and other 'cultural' features (Table S 3.2)
- roads and pipelines 15 m from buildings and other 'cultural' features (Table S 3.2). Note this is to ensure realistic construction rather than adhere to a regulation.
- pads and roads sited in areas between 0 and 20% slope (Table S 3.2). Note this is to ensure realistic construction rather than adhere to a regulation.

3.7.5 Derived Pipeline Methods

Gas wells must be connected to the existing pipeline network to transmit gas to the market. Many of our derived sites are many kilometers away from the nearest gas pipeline; a distance unrealistically far for development due to its cost and logistical difficulty. We know that the derived sites have been partially developed because we use only active or once-active wells to derive them. Because gas pipeline data are proprietary, no GIS data warehouse has complete pipeline data. The

3.7.6 Datasets

Table S 3.2. Description of datasets used in the analysis, including those used by default in Bungee and those specific to this analysis.

Dataset	Used for	Source
well permits	estimating development boundaries	Pennsylvania Department of Environmental Protection: Office of Oil and Gas Management: http://www.portal.state.pa.us/
forest patches	siting infrastructure, forest metrics	National Land Cover Database 2006 (classes 41, 42, 43, 90): http://www.mrlc.gov/nlcd06_data.php
wetlands in PA	restricting infrastructure, siting infrastructure, wetlands encroachment metric	National Wetlands Inventory and National Land Cover Database 2006 http://www.fws.gov/wetlands/Data/Mapper.html
habitat classifications for NE USA	rare spp. metric	Northeastern Terrestrial Habitat Mapping Project: http://conserveonline.org/workspaces/ecs/documents/ne-terrestrial-habitat-mapping-project
spatial locations of rare species observations in PA	rare spp. metric	Pennsylvania Natural Heritage Program data request. Cannot release data.
streams, rivers, lakes	restricting infrastructure, siting infrastructure, water metrics, *sediment yield* raster	National Hydrography Dataset: ftp://nhdftp.usgs.gov/DataSets/Staged/States/FileGDB/HighResolution/
Digital Elevation Model (*dem*), elevation at a point in raster	restricting infrastructure, siting infrastructure, cut-and fill cost, percent-*slope* raster, *sediment yield* raster	National Elevation Dataset within the National Hydrography Dataset: http://www.horizon-systems.com/NHDPlus/NHDPlusV2_data.php

Table S 3.2 Continued

Dataset	Used for	Source
'cultural' features, i.e. schools, recreational fields, dwellings, reservations, etc.	restricting infrastructure	U.S. Board on Geographic Names' Geographic Names Information System (GNIS): http://geonames.usgs.gov/domestic/download_data.htm
existing roads	siting infrastructure	US Census Bureau TIGER/Line 2008 (all counties All Lines RDFLAG = "Y"): http://www.pasda.psu.edu/default.asp
existing pipelines	siting infrastructure	MapSearch pipelines. Propriety data.

pipeline dataset we purchased from MapSearch has many larger pipelines that could be connected to, but misses many of the other available tie-ins (Table S 3.2).

To get around the sparsity of our pipeline dataset, we made the assumption that new gathering pipelines would be developed in the direction of the nearest known pipeline in our dataset. We based our pipeline connections on the rectangle bounding the site with its sides parallel to the cardinal directions ("bounding rectangle" hereafter). First, we assumed any gathering pipelines could connect to any existing pipelines intersecting the bounding rectangle. When no known pipelines intersected the bounding rectangle, we assumed new gathering pipelines would be developed in the direction of the known pipeline closest to the centroid of the site. In some cases, the nearest pipeline was over ten kilometers away. We forced gathering pipelines in this latter category to connect to a section of the bounding rectangle intersected by a 45° wide wedge oriented with its point in the centroid of the site and the angle towards the nearest pipeline.

3.7.7 Bungee Workflow

Bungee is software that combines variations of many well known optimization techniques to plan locations of well pads, access roads, and gathering pipelines in a way to avoid some environmental impacts while explicitly accounting for and limiting construction costs. Due to its size, we cannot fully document Bungee here. Instead, we describe the optimization components of Bungee and how these components fit together. We refer the reader to the Bungee user guide and technical documentation for a more complete understanding of its methods.

The Bungee code is a set of Python scripts written almost exclusively in Python 2.7.2, with some optimized code written in Cython 0.19.1. Bungee consists of two major parts: a) the Python module, which can be imported and used like any other Python module, and b) Bungee GIS, which is a Python toolbox for ArcGIS. Both require ArcGIS 10.1+ with the Spatial Analyst Extension. The module structure of Bungee is similar to ArcGIS's arcpy module, with major tools composing the main steps of the workflow as well as some minor but useful tools all directly accessible within the Python environment. Using Bungee GIS with default settings, analyses take on the order of hours to days, where the scaling of runtime increases nonlinearly with the size of the analysis area. Runtime also increases with the number of well pads, but less closely. For instance, Bungee placed infrastructure in a 12 km^2 area in 45 min on a typical desktop computer. By comparison, Bungee required multiple days but less than one week to plan infrastructure in the largest of the sites in this analysis (area=72 km^2). At the time of writing, access to Bungee can be obtained by emailing the corresponding author.

The Bungee optimization workflow for a single site consists of six major steps:

3.7.7.1 Place production units

Production units (or Drainage units) – the area being drained by a single pad – are placed in the lease-hold to maximally drain gas while minimizing the number of pads being developed. The output of this step is a set of 1) production unit polygons and 2) pad envelopes, which show the restricted area within each drainage unit where a well pad may be placed. See 3.7.8 Production Unit Packing Methods.

3.7.7.2 Set up infrastructure restrictions

Constraints on the locations of well pads, access roads, and gathering pipelines are set up. In this analysis, we follow those constraints described in 3.7.4 Setbacks and Other Restrictions. This step produces rasters which denote allowable areas for each type of infrastructure.

3.7.7.3 Determine construction budget

Bungee creates an infrastructure layout that attempts to minimize construction costs. For this analysis, Bungee then enforced a maximum budget that was twice as large as the least-cost layout. During the optimization, Bungee actually incrementally increases the budget, from the least-cost amount up to the maximum budget and creates layouts at each incremental budget. In 3.7.9 Additive Layout Creation Methodswe describe the algorithm used to create layouts. To get the least-cost layout, Bungee uses the cost surfaces in place of impact surfaces and uses an unlimited budget.

3.7.7.4 Determine normalization constants

As described in Impact Score calculation, each impact metric is normalized using a normalization constant. To determine this constant for a single impact metric and site, Bungee creates many layouts (n=1000 in this analysis) that attempt to minimize the value of the current impact metric. The minimum value of the metric across those layouts is the normalization constant. In 3.7.9 Additive Layout Creation Methods we describe the algorithm used to create layouts. In addition, the average construction cost across metrics from this step is used as a normalization constant for costs.

3.7.7.5 Estimate additive impact surfaces

After step 4, Bungee can calculate the Impact Score. At this point, Bungee estimates spatially additive surfaces of the Impact Score, one for each type of infrastructure, and which are important inputs into the final step. In each pixel of one such surface for a single infrastructure type is the Impact Score if only that infrastructure were developed only in that pixel. As such, the true aggregate Impact Score is underestimated in each pixel, since non-additivity among multiple pixels of one infrastructure or multiple infrastructures is not taken into account.

3.7.7.6 Create final layouts

The outputs of the previous steps enable Bungee to create feasible layouts, which adhere to setbacks and construction budgets. At this point, Bungee creates layouts that attempt to minimize the aggregate Impact Score. To do so, Bungee incrementally increases the construction budget from the least-cost-layout's cost up to the maximum cost. For this analysis, we used 40 increments. At each increment, the additive impact surfaces and normalized cost surfaces are weighted and added together to form a hybrid objective surface. This hybrid surface encourages Bungee to balance impacts and costs. The impact weight in each increment is given by

$$w_{Impact} = \frac{b_i - B_{min}}{B_{max} - B_{min}}$$

where b_i is the construction budget at the current increment, B_{min} is the cost of the least-cost layout, and B_{max} is the maximum construction budget, which here was $2B_{min}$. The weight of the normalized cost surface is $1 - w_{Impact}$.

Bungee plans layouts according to the hybrid objective surfaces (see 3.7.9 Additive Layout Creation Methods). At the same time, it uses the aggregate Impact Score to judge each layout once the full layout has been proposed. It also limits the layout's budget and calculates construction costs based on the non-normalized cost surfaces. Finally, Bungee keeps only those layouts that are not simultaneously more impacting and more costly than any other layouts, i.e. are Pareto-improvements. This forms the set of final layouts for each site. In our analysis, we ran 3.7.7.6 Create final layouts five times and repeating the Pareto-optimality filtering across runs to increase the chances Bungee actually found Pareto-optimal solutions.

3.7.8 Production Unit Packing Methods

Production units are sited as a means to determine the number and approximate locations of well pads within a site. A production unit is the potential area drained by a well pad once all its wells have been drilled. Since real well pads have variable numbers of wells in variable configurations, production units must also be various shapes. Bungee is able to accommodate a range of number of wells per pad and optimizes the number of wells within this range, their configurations, and locations of the production units accordingly. In this analysis, we took a simplified approach and had Bungee place only 6-well production units 3352.8 m tall by 914.4 m wide (3000 by 11000 ft) and rotated 27° counter-clockwise. As such, we also simplify the explanation of the production unit packing algorithm for brevity and clarity.

Production units are packed into the site by iteratively adding production units to the site and then using simulated annealing to shuffle them and free up space. The objective of the optimization is to minimize the un-drained area of the site. The optimization acts on a raster/array basis, such that production units, the site, production unit locations, and the objective function all operate on a pixel-by-pixel basis. The algorithm proceeds as follows:

1. Add a 6-well production unit in the location that minimizes its bad overlap. 'bad overlap' for a single production unit is its area of overlap with a) other production units and b) areas outside the site.
2. If any present production units exceed the overlap threshold (v=10% of production unit area), proceed to step 3. Otherwise, return to step 1.
3. Perform simulated annealing. Shuffle production units to free up space in the site. For a fixed number of iterations (n=5000):
    a. For each production unit:
        i. Move it one pixel in a random direction and calculate its bad overlap.
        ii. If the bad overlap is reduced from its previous position, keep the new position. Otherwise, keep the new position with some probability (p_{start}=0.1).
    b. Calculate the un-drained area of the site. If the un-drained area decreased, keep this new configuration of production units. Otherwise, keep the new configuration with some probability (same as 3.a.ii.).
    c. Decrease the probability that worse solutions are kept such that by the end of step 3, the probability of keeping worse solutions is zero.

4. Repeat the check in step 2. If it still fails, remove the last production unit and return the set of production units placed. Otherwise, return to step 1.

The optimization determines the number of production units and their locations. To get the pad envelopes, which define the regions where a pad may be placed before setbacks are enforced, we center a 304.8 m wide by 1524 m tall (1000 by 5000 ft) rectangle in each production unit. Note that, like the production units, the shapes of pad envelopes vary with the number and configurations of wells per pad.

3.7.9 Additive Layout Creation Methods

3.7.9.1 Simplifying assumptions

Bungee simplifies the infrastructure planning problem so that impact-avoiding layouts can be proposed on a desktop computer in a reasonable amount of time (one to several hours or days). There are two important characteristics of the planning problem which Bungee's algorithm partially avoids. An explanation of these is necessary to understand the approach we took in deriving Bungee's layout planning methods. First, the Impact Score is a spatially non-additive function of infrastructure layouts, such that it cannot be fully assessed without the entire layout. This poses a difficulty due to the large solution space of layouts and few ways to narrow down proposed layouts to avoid searching the whole solution space. Second, the optimal (i.e. Impact Score minimizing) positions of well pads, access roads, and gathering pipelines are all interdependent since each type of infrastructure can affect the locations of others. This difficulty occurs even if the Impact Score were spatially additive. We programmed Bungee to deal with these difficulties by simplifying the infrastructure planning algorithm. We make the following assumptions

- The Impact Score can be estimated on a pixel-by-pixel basis by treating each pixel as if it were developed independently of others.
- The optimal route of an access road or gathering pipeline depends only on infrastructure that has already been proposed and on the pixel-by-pixel estimate of the Impact Score.

Since layouts are based on the spatially additive approximation of the Impact Score, we call them additive layouts.

3.7.9.2 Additive layout attribute optimization

With these assumptions in mind, we developed a hierarchical algorithm for planning infrastructure layouts. In the higher level, a genetic algorithm is used to optimize the attributes of a layout - well pad locations and the order of planning of access roads and gathering pipelines. For instance, in a site with two well pads, the attributes might look like

Attribute	Y_a	X_a	Y_b	X_b	L_1	L_2	L_3	L_4
Example	41° N	80° W	42° N	81° W	$Road_b$	$Pipe_a$	$Pipe_b$	$Road_a$

where *Y* and *X* are the vertical and horizontal positions of well pads *a* and *b*, and L_i denotes the order of development of linear infrastructure corresponding to each well pad. Note that the order of linear infrastructure is important because Bungee "fixes" each planned route – treats it as already developed -such that subsequent linear infrastructure can be co-located with or terminated at that infrastructure (see Bungee documentation for more).

The objective of the optimization is to minimize the Impact Score given a budget constraint. Bungee uses a least-cost path algorithm in the lower level to plan the routes of access roads and gathering pipelines, which provides the information needed to evaluate the layout for its Impact Score. We describe the higher level of the algorithm in the paragraphs below and the lower level in 3.7.9.3 Linear infrastructure route optimization. We use an asterisk (*) to denote details further described in the Bungee documentation.

1. Propose parent population (n=20) of layout attributes. Each layout is proposed by:
    a. Decide pad centroid locations. Bias pads towards lower-impact areas.*
    b. Decide the order of linear infrastructure planning by random draw.
    c. Evaluate layout:
        i. Overlay pad footprint on pad centroids.
        ii. Plan routes of access roads and gathering pipelines (see 3.7.9.3 Linear infrastructure route optimization).
        iii. Calculate Impact Score and construction cost. If too expensive, discard the layout.

2. Add lowest-impact layouts (n=3) to offspring population without altering them.
3. Choose two parent layouts to crossover/mate. Bias choices to those with lower Impact Scores.
4. Crossover parent layouts with some probability (p=0.7) to form an offspring layout. If not crossed-over, the offspring is an identical copy of the first parent. Otherwise:
   a. Randomly draw each pad location from one parent or the other.
   b. Decide offspring's infrastructure order. For each order slot L_i:
      i. Choose at random one parent to take from.
      ii. Fill in the slot with the next un-used infrastructure (e.g. $Road_b$) from the parent chosen in 4.b.i.
5. Mutate/alter new offspring layout created in 4:
   a. Move each well pad some number of times* in one-pixel steps.
   b. Shuffle infrastructure order. Loop through each slot and with some probability (p=0.1) swap the infrastructure in this slot with the next slot.
6. Evaluate offspring layout as in 1.c.
7. If offspring layout has lower Impact Score than both parents, add it to the offspring population. Otherwise, add it anyway with some probability (p=0.05).
8. If some number (n=5,000) of layouts have been proposed and kept *or* discarded, return the lowest-impact layout from the offspring population. Otherwise, go to 9.
9. If the offspring population is full (n=20 layout attributes), go to 10. Otherwise, go to 3.
10. Check for convergence. If the lowest-impact layout in the offspring population has not changed in some number of generations (n=7), propose a new set of layout attributes (n=17) as in 1, but keep the current lowest-impact layouts from 2. Otherwise, replace the current parent population with the offspring population and go to 2.

3.7.9.3 <u>Linear infrastructure route optimization</u>

In 3.7.9.1 Simplifying assumptions we describe assumptions that allowed us to entirely represent a layout by the locations of well pad centroids and order of development of linear infrastructure. To go from its attributes to the binary rasters necessary to calculate the Impact Score and construction cost, we need only to overlay a well pad footprint on the well pad centroids and plan the routes of linear infrastructure. We can plan the route of a piece of infrastructure using a least-cost-path algorithm.

The least-cost-path algorithm we use is a modified version of Dijkstra's algorithm[2] and guarantees that the route found is globally optimum given the assumptions we outlined. Due to its complexity, we do not list the algorithm steps here, but instead describe how our algorithm builds on Dijkstra's algorithm. We fully describe the optimization in the Bungee documentation.

Dijkstra's algorithm works on raster (or array) data to find the least-cost route from a set of source/starting pixels to a set of destination pixels. It assumes that the total cost of traversing a complete or incomplete route is the sum of the costs of each pixel along the route, i.e. the cumulative cost is spatially additive. In this context, the source pixels are well pads including their footprints and the destination pixels are existing and already planned roads or pipelines, depending on the infrastructure being planned. The cost surface over which a route is planned is the hybrid impact-cost surface described in 3.7.7.6 Create final layouts.

Our least-cost-path algorithm adds two features to Dijkstra's algorithm. First, we add a binary "traversability" raster which denotes areas that are off-limits for the infrastructure being planned. This raster includes, for instance, those pixels in high-slope areas. Those off-limit areas are not considered when planning routes. Second, we add a secondary cost surface which is used to invalidate routes. The secondary cost surface is the construction cost of either road or pipeline in each pixel. As the updating procedure central to Dijkstra's algorithm proceeds, the cumulative cost of the route passing through each pixel is stored. If a pixel's cumulative cost exceeds the cost budget, then the route passing through that pixel is invalidated, *even if that route is the lowest-impact route*. In this way, routes that are too expensive are not proposed.

[2] Dijkstra EW (1959) A Note on Two Problems in Connexion with Graphs. *Numer Math* 1(1):269–271.

Chapter 4: Comparing policies for the reduction of potential environmental impacts from shale gas surface infrastructure

A version of this chapter will be submitted for peer-reviewed publication.
Austin W. Milt, Paul R. Armsworth. "Comparing policies for the reduction of potential environmental impacts from shale gas surface infrastructure."

4.1 Abstract

Governments across the globe at multiple levels have shown interest in avoiding the environmental and human health impacts created by shale energy production. In areas where shale energy production is currently allowed, regulations restricting environmental impacts tend to be limited in scope and flexibility. We present a study looking at the cost effectiveness of implementing a new regulation that affects the spatial locations of well pads, access roads, and gathering pipelines with an aim at reducing aggregate environmental impacts from shale gas development. Specifically, we compare the outcomes of two policies: (1) a uniform and inflexible cap on site-level impacts, and (2) a cap and trade system which allows developers to trade permits for impacts. Both of these are also compared to optimum outcomes produced by planning by an omniscient social planner. We measure the total cost and impact of the system under each scenario. We analyze a case study of 56 sites in Pennsylvania, U.S.A., a location which has experienced extensive gas development already. We find that under ideal conditions cap and trade performs as well as an omniscient social planner, producing lower impact outcomes much less expensively than the uniform inflexible policy. Cap and trade could reduce impacts by as much as ~36% at an increased cost of 0.05% of not developing while still allowing all development to proceed. Having found large potential gains from trade, we explore how the cost effectiveness of a cap and trade system depends on the ability of the regulator to estimate impacts in the absence of additional regulation. In extreme cases, error in that estimate could make cap and trade less cost effective than the uniform inflexible approach. Our results clearly indicate that for intermediate levels of impact avoidance, cap and trade is a highly cost-effective alternative to a more traditional approach provided that the regulator is able to accurately estimate impacts in the absence of additional regulation.

4.2 Introduction

Shale gas development is an increasingly global issue due to energy and environmental concerns. In the United States, shale gas production has increased steadily over the past decades and now makes up ~40% of gas production (*Annual Energy Outlook 2014* 2014). Concerns have been raised about the environmental (Gillen & Kiviat 2012; Kiviat 2013; Olmstead *et al.* 2013; Jones *et al.* 2014) and human health (Perry 2012) effects of shale energy production (see dissertation Introduction and Chapter 2), leading to careful consideration of how to protect society and nature from those effects (Howarth, Ingraffea & Engelder 2011; Hays *et al.* 2015) and at times outright bans on development. Though policies and regulations in regions proceeding with development do exist, new regulations can expand their environmental scope to include priorities currently unregulated. These new policies and regulations will differ in their acceptability and cost effectiveness.

Shale gas production takes place in many stages (Burton *et al.* 2014) at multiple spatial scales. Throughout Chapters 2-4 we focus on one stage and one scale. Specifically, we focus on the construction of surface infrastructure at the least-hold scale. Lease-holds ("sites" hereafter) are boundaries of development that aggregate multiple gas leases and tend to range in size from several hundred hectares to many thousands of hectares. Shale gas extraction requires significant below-ground infrastructure which is often the focus of environmental studies (Hays *et al.* 2015). However, extraction requires significant surface infrastructure to access drilling sites, process gas, and transport it to market. We focus on well pads, access roads, and gathering pipelines, infrastructure which is common at all gas extraction sites and which has measurable environmental effects. The spatial planning of these three types of infrastructure is a complex process from a cost-minimization perspective. The cost-minimizing configuration of infrastructure relies on the simultaneous consideration of interactions among infrastructure locations. For instance, well pads form a terminus for access roads – wells affect roads, but roads cannot be built on very steep slopes up to a plateau where a well pad might be located – roads affect well pads.

Potential environmental damage ("impacts" hereafter) from shale gas surface infrastructure can be partially avoided by informed spatial planning (as discussed in the dissertation Introduction and Chapters 2-3). Many environmental features are impacted by shale gas surface infrastructure (Gillen & Kiviat 2012). Roads and pipelines fragment habitats, which increases habitat edges, produces dispersal barriers, and reduces core habitats. Construction exposes and mobilizes surface soils, potentially leading to erosion and subsequent sedimentation in water bodies. Stream-crossing infrastructure reduces freshwater connectivity by limiting upstream and downstream dispersal. These are a few of the common and pervasive impacts from surface infrastructure, all of which depend on the spatial configuration of infrastructure – "layout" hereafter – and which can be partially avoided by changing the layout. In the simplest case, reducing the amount of infrastructure reduces the area of land disturbed and thus some impacts. However, the shortest route for a road may be more fragmenting than an alternative route which circumnavigates an important habitat. And the least-expensive layout may be largely unrelated to the resulting impacts. Thus, impact avoidance through spatial planning of infrastructure can be a complex, many-dimensional decision process.

Current environmental regulations for shale gas surface infrastructure tend to be limited in their type and scope. In many places globally, there are either moratoriums or outright bans on shale gas development (http://keeptapwatersafe.org/global-bans-on-fracking/ visited 23 April, 2015). While these prevent environmental damage in the short term, it is possible that development will proceed as fossil fuels become more limited. The main form of regulation for sanctioned development is uniform command-and-control (Richardson *et al.* 2013), in which restrictions are uniform and absolute. However, the uniform strictness of such regulations is often offset by an ability to avoid them. For instance, in Pennsylvania well pads cannot be placed within 30 m (100 ft) of wetlands larger than 0.4 ha (1 acre), but exceptions can be granted at the site level when sufficient protective measures are proposed (Pennsylvania legislation Title 58 3215(b)(3)&(4)). Other environmental policies include performance practices which are usually not enforced but are encouraged (Richardson *et al.* 2013). These current regulations tend to focus on water features, an important part of environmental and human health concerns but of limited scope.

Market-based regulations can solve some issues with uniform and inflexible command-and-control approaches. First, market-based mechanisms exploit heterogeneity across regulated participants to achieve optimal outcomes (Hartwick & Olewiler 1998). Contrarily, a uniform command-and-control approach ignores this heterogeneity and leads to situations where development cannot occur or exceptions are granted, thereby achieving no effect. For instance, regulating forest loss by limiting it to 10 ha in each site might preclude development in fully-forested areas, which may be a desirable or undesirable outcome. Alternatively, a market-based mechanism might incentivize developers at fully-forested sites to compensate developers at moderately-forested sites to reduce their forest destruction even further than 10 ha. Second, market-based mechanisms could theoretically reduce some administrative costs of environmental regulation by reducing information and analysis requirements. In the shale gas context, a command-and-control attempt to tailor restrictions at a site might require much information about that site as well as the development process to decide on an optimal infrastructure layout. Contrarily, a market-based approach could incentivize developers to explore development alternatives using information they already have and without any need for the regulator to be directly involved in the planning process at a site.

At the same time, regulations differ in how much information is needed and how that information is used in implementation. First, uniform regulations may have low information requirements. Here, for instance, we distinguish between a scenario where caps are set uniformly across sites and another scenario where caps are site specific. Setting site-specific caps requires some additional information. Second, market-based regulations may have relatively high implementation costs because they require the regulator to institute, monitor, and regulate a market. Most regulations will have lower information requirements and implementation costs than an omniscient social planner. This omniscient social planner is a hypothetical construct that optimizes the system as a whole rather than individual sites and can be used as a benchmark for evaluating alternative policy designs. In order to affect optimum outcomes, the planner is assumed to have perfect information about the system and an ability to affect all decisions. In some cases this is not far from reality, e.g. when a sole owner holds all development rights.

Several aspects of regulating impacts in this context are of general interest. First, we are interested in the regulation of aggregate impacts rather than treating impacts individually. Regulating a single aggregate metric across multiple sites requires that the metric be comparable across sites, which may necessitate a sacrifice of the site-specific nature of some impacts. Second, regulating an aggregate metric also permits flexibility in which impacts are reduced and how they are reduced, e.g. by not necessarily penalizing sites that cannot perform well in one metric if they can offset that assumed performance with another. Note however that our analysis would apply readily to single-impact contexts. Third, options for how much impacts are produced in a site are discrete because the choices of layouts are limited, with several layouts being similar to one another followed by large changes to those layouts. Fourth, impacts in this context are one-off since changes to the land surface are irreversible and long-lived. As such, impacts at a site are not regulated over time.

In a recent special issue of *Ecological Economics* concerned with market-based instruments for ecosystem services, Gómez-Baggethun & Muradian (2015) summarize some of the strengths, weaknesses, and controversies surrounding market-based-instruments for environmental purposes. They point out that market-based-instruments have increased in popularity for policymakers and scientists (Pagiola & Platais 2002; Engel, Pagiola & Wunder 2008; Miles & Kapos 2008; Lockie 2013; Lapeyre, Froger & Hrabanski 2015) at the same time that trust in the power of markets has fallen (Sandel 2012; Gómez-Baggethun & Muradian 2015). Many of the issues surrounding market-based-instruments have to do with how purported market-based-instruments are structured and implemented and in what situations they are applied (Gómez-Baggethun & Muradian 2015). Even when properly structured and applied, the theoretical cost effectiveness of market-based-instruments (Foster & Hahn 1995; Goulder *et al.* 1999) depends on the ability of regulators to set optimal conditions for the market (United Nations Development Program 2011; ten Brink *et al.* 2012), which itself may require accurate estimates of benefits and costs (Salzman & Ruhl 2000; Kroeger & Casey 2007). Indeed we find that to be the case here. Regardless, market-based-instruments can have many strengths over rigid command-and-control approaches (Gómez-Baggethun & Muradian 2015).

While other peer-reviewed studies have looked at the current regulatory framework for shale gas development, to our knowledge none have quantitatively analyzed the environmental and monetary effects of implementing new regulations. Konschnik and Boling (2014) describe the current regulatory framework for shale gas in the U.S and go on to propose a framework for further governance of shale gas and how that could be applied for environmental or sustainability goals. Most other studies focus on a review of current regulations (Rahm 2011; Clark *et al.* 2012; Wiseman 2014) or on the assessment of risks or damages for future regulations (Clark *et al.* 2012; Hays *et al.* 2015). We draw on the foundational knowledge of these studies, which point to the limitations of existing regulations, and combine that knowledge with spatial planning of infrastructure for multiple environmental impacts at the site scale to address the implications of an additional regulation.

In this paper we explore the cost effectiveness of different environmental regulations for shale gas surface infrastructure, especially how regulations compare to one another and to an idealized benchmark. Specifically, we explore three scenarios. First, we explore a regulation that reflects the most common type of environmental regulation, which is a uniform - ignores site characteristics - restriction on impacts and is inflexible in how that restriction is met. Second, we explore and focus on how tradable permits in a cap and trade system reduces the cost of reducing impacts compared to the first scenario while leaving development decisions in the hands of developers. Third, we evaluate whether and idealized cap and trade system performs as well as an omniscient social planner, something expected in theory (Hartwick & Olewiler 1998). As discussed earlier, the cost effectiveness of market-based instruments depends on the market context created by the regulator. Consequently, we then discuss how error in the ability of the regulator to estimate impacts in the absence of an additional regulation affects the cost effectiveness of cap and trade. We then analyze a case study of development in Pennsylvania and discuss its implications for cap and trade in similar and broader contexts

4.3 Methods and Materials

4.3.1 Overview

In the following sections we describe our methods for analyzing the costs and impacts associated with regulating shale gas surface infrastructure using different regulations. In §4.3.2-4.3.5, we describe our general methods for analysis of two policy scenarios and an idealized benchmark, followed by an application of these methods to a case study described in §4.3.5. In §4.3.2 we describe the regulatory context, including our assumptions about the system, the goals and decisions made by a regulator committed to reducing environmental impacts from surface infrastructure, and the goals and decisions made by developers attempting to make profits from extracting gas from their sites. In §4.3.3 we describe the mathematical formulation and solution methods for the two policy scenarios and idealized benchmark. In §4.3.4 we expand our analysis to consider the counterfactual situation where the regulator does not perfectly know how large impacts will be in the absence of additional regulation. This introduces error in setting the site-specific cap for cap and trade. We discuss how three directions of error affect outcomes of the system. Finally, in §4.3.5 we describe the application of our methods to a case study set of 56 sites in Pennsylvania, USA. This application requires some additional data and analyses, some of which come from a previous study (Chapter 3); the rest is described in §4.3.5.

4.3.2 Regulatory Context

We follow several assumptions about the development context that affect how we analyze new regulations. First, the development rights at a site belong to only one developer and each developer has development rights to exactly one site. Thus, decisions about how to develop a site are site/developer specific. We change this assumption when exploring a sole-ownership scenario. Second, every layout option – configuration of well pads, access roads, and gathering pipelines – for a site has the same number of wells, all wells drain the same amount of gas, and all wells cost the same to drill. Thus, layouts for a site differ only in the cost of developing surface infrastructure. Third, the construction of infrastructure produces many environmental externalities (impacts), which it is the task of the regulator to internalize to the gas industry through a new regulation. Fourth, impacts incurred at a site are independent such that the aggregate impact of development of the system is just the sum of site-level impacts. Finally, all sites are developed simultaneously such that delays in gas production do not occur and the costs and profits from developing sites are independent of the start of production. We recognize these are simplifying assumptions of the system which limit our ability to fully predict outcomes of different regulations. However we feel this study is still an important first step toward understanding the implications of new shale gas regulations.

The regulator is responsible for creating a new regulation that forces developers to internalize environmental impacts created by surface infrastructure. The regulator would like to maximize social welfare by minimizing both environmental impacts and the monetary cost of internalizing those impacts. However, the regulator has a limited ability to do so for two reasons. First, the regulator does not know the social value of environmental impacts and so cannot directly maximize social welfare. Instead, the regulator can only choose a level of impact to achieve, which will result in some cost to the gas industry. Second, in the cap and trade system the regulator sets an individual cap for each site based on an estimate of the impact of the least-expensive layout at that site, which is the layout that would be developed without the regulation. There is error associated with that estimate, which prevents the regulator from knowing whether the choice of cap will lead to a larger or smaller total impact than estimated (described more in §4.3.4).

Each developer wants to maximize the net present value of his site, which is dependent on several factors. A site contains some amount of gas, the present value of which depends on the flow rate of gas from each well, the number of wells, the price of gas, and the monetary discount rate (p. 140 in Appendix for methods). To get profits from the gas, the developer must construct infrastructure to access the site, extract the gas, and pipe it to the market. There are many infrastructure layouts for a site, and each layout has an associated construction cost and environmental impact. We denote the discrete cost and impact functions for a site by $C_i(j)$ and $I_i(j)$, respectively, where $C_i(j)$ is the cost of constructing layout j at site i. These functions are monotonically increasing and decreasing respectively and thus their combination adheres to one important Pareto-efficiency condition (Varian 2003). We set up our analysis in such a way that there is no incentive to develop a layout that is simultaneously more impacting and more costly than any other layout. Because of the setup described here, a developer can maximize the net present value of his site by minimizing the cost of construction plus any additional costs from the new regulation. The specific form changes with each scenario (Table 4.1).

4.3.3 Scenarios and Solutions

We explore the two policy scenarios and one idealized benchmark already described above. In every scenario, the regulator puts a cap on impacts from development and the developer(s) choose layouts at each site to adhere to the cap while minimizing development costs. We show the optimization problem for developers in each scenario in Table 4.1. In each scenario, we find solutions for a range of caps and record the total cost – sum of costs across all sites – and total impact – sum of impacts across all sites.

We define the layout choice j with several characteristics to facilitate solving the planning problems. First, $j = 0$ represents the least-expensive and most impacting layout in a site, while $j = j_{max,i}$ is the layout with the highest cost and lowest impact where development still occurs. We define a special "dummy" layout for the decision to not develop a site. The choice to not develop occurs when the cost of development exceeds the profits from development. The layout that reflects this choice, denoted $j = \emptyset = j_{max,i} + 1$, has characteristics $I_i(\emptyset) = 0$ and $C_i(\emptyset) = V_i$, where V_i is the present value of gas in the site.

Solving the first and last problems shown in Table 4.1 is fairly simple. In *Uniform Cap without Trading*, each developer chooses the least-expensive layout that meets the cap. If the cap is lower than the impact of the $j_{max,i}$ layout then the site is not developed. In *Omniscient Social Planner*, the one decision maker chooses the cumulatively least-expensive combination of layouts across all sites that meets the cap. Because we assume impacts are additive across sites, this is a linear problem. To solve it, we start with all $j = 0$. We then calculate the return-on-investment (ROI) of switching each site's layout to the $j + 1$ layout, where. We iteratively swap the layout at the site with the highest ROI until the impact constraint in Eq. (3) is met. This method produced identical solutions to a global branch-and-bound solver.

Solving the outcome of the market in *Cap and Trade* is somewhat more difficult. Each developer can choose to be a supplier of permits if $I_i(j) - \alpha\hat{I}_i(0) < 0$ or a demander of permits if $I_i(j) - \alpha\hat{I}_i(0) > 0$, and the optimum choice depends on the price of permits in the market (Table 4.1). Because the impact production at each site is discrete, it is impossible in a small market to have supply exactly equal demand. Consequently, there will always be some excess supply or demand. Our earlier assumptions dictate that if there is excess demand, some sites are not developed.

To find the final market price of permits (P^*), the choice of layout for each site, and the total cost and impact of the system, we perform a two-level search for P^* (Fig. 4.1). We start with some additional assumptions: 1) the market is perfectly competitive, 2) the market has

Table 4.1. Developer optimization problems for policy scenarios, including the objective and constraint, which is set by the regulator.

Policy	Developer Objective	Impact Constraint	Eq.
Uniform Cap without Trading	$\min_{j} C_i(j)$	$I_i(j) \leq A \; \forall i$	(1)
Cap and Trade	$\min_{j} C_i(j) + P\left(I_i(j) - \alpha \hat{I}_i(0)\right)$	$I_i(j) - \max(0, \Phi_i) \leq \alpha \hat{I}_i(0)$	(2)
Omniscient Social Planner	$\min_{j \in J_i \in \boldsymbol{J}} \sum_{\forall i} C_i(j)$	$\sum_{\forall i} I_i(j) \leq \alpha \sum_{\forall i} I_i(0)$	(3)

i = site/developer index
j = layout index
J_i = set of layouts at site i
$\boldsymbol{J}$ = set of all layouts
$C_i(j)$ = cost of developing layout j at site i
$I_i(j)$ = impact of developing layout j at site i
A = uniform site-level cap
α = proportional cap on impacts
P = price of permits in market
$\hat{I}_i(0)$ = regulator's estimate of impact of least-cost layout
Φ_i = permits bought for site i

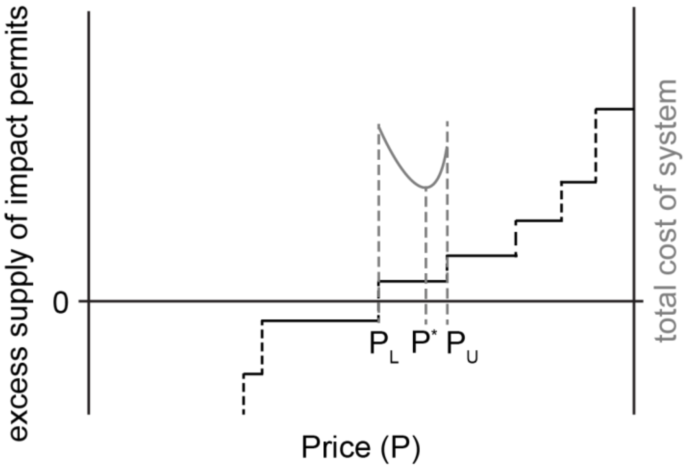


Fig. 4.1. Stylized illustration of the two-part process we use to estimate the final market price (P^*) of impact permits in *Cap and Trade*. (stepping black line) difference between supply and demand at market price P. First part of the search for P^* occurs along this curve. (gray curve) total cost to all developers in the market, which may be concave or monotonic within the range P_L-P_U. P^* is the price at the minimum of this second curve.

been established long enough to reach an equilibrium price, and 3) all developers simultaneously enter the market and trade. At the first level, we find the range of P where excess supply is minimized and there is no excess demand (P_L-P_U in Fig. 4.1) – to avoid forcing some sites out of development. There is a range of P that meets this condition because layout choices are discrete and as such there will be a range of P within which changes to P do not change the set of layouts chosen. At the second level, we find the P within this range that minimizes the total cost of the system (P^* in Fig. 4.1). Finally, because there will still be excess supply at the final P^*, suppliers will lose some potential profits and we add these lost profits (equal to excess demand times P) to the total cost of the system. Because of the way we analyze the market, our analysis is an optimistic estimate of the outcomes of using cap and trade.

4.3.4 Regulator's Error in Estimating $\hat{I}_i(0)$

The total cost and impact of the cap and trade system depends on the cap set for each site. To illustrate, take one site in isolation. At the extremes, the cap may be so low or so high that the developer cannot develop or does not reduce potential impacts, respectively. Within the range $\left[I_i\left(j_{max,i}\right), I_i(0)\right)$ the cap has some effect on the developer's choice of layout while still allowing development. In *Uniform Cap without Trade*, we assume the regulator ignores this effect because the most common environmental regulations currently do not tailor restrictions to each site. However, it is reasonable that the regulator could estimate impacts in the absence of the new regulation, e.g. by examining existing development, and thus increase the chance that restrictions on impact lead to development choices within the above range.

In *Cap and Trade,* we assume that the regulator has some ability to estimate site-level impacts in the absence of additional regulation, denoted $\hat{I}_i(0)$. This estimate has some error associated with it due to the regulator's lack of perfect information. In the case study below, we start with the case where the regulator can perfectly estimate impacts in the absence of additional regulation ($\hat{I}_i(0) = I_i(0)$) and then perform several sensitivity tests, including the regulator's estimate is 1) systematically high ($\hat{I}_i(0) > I_i(0)$), 2) systematically low ($\hat{I}_i(0) < I_i(0)$), and 3) incorrect but without bias. We will show that error in $\hat{I}_i(0)$ does not change the possible outcomes of the system, but affects which outcomes are revealed and how the regulator's choice (mis)matches with the outcome produced.

4.3.5 Case Study

We applied our framework to the Marcellus shale play in Pennsylvania, a place where enough development has occurred and enough knowledge about the development context exists to infer with some confidence the cost effectiveness of a cap and trade system. Over 9,000 horizontal wells have been drilled in the Marcellus region of Pennsylvania since 2008 – Pennsylvania Department of Environmental Protection's permit reporting database – and many more are likely to come. The construction of well pads, access roads, and gathering pipelines is occurring in areas of high conservation priority (Johnson *et al.* 2010) resulting in degradation and destruction of many environmental features including forests, wetlands, streams, and other features important for biodiversity and recreation in the area (Johnson *et al.* 2010) (also see dissertation Introduction).

We used the results of the previous chapter for our analysis here (Chapter 3). In our previous study, we created a spatial planning software call Bungee to place well pads, access roads, and gathering pipelines at 85 sites in Pennsylvania. Site boundaries were derived by overlaying production units on existing well locations and then joining adjacent land parcels to fully contain those production units. In that study, production units were 914 × 3353 m (3000 × 11000 ft) rectangles rotated 27° counter-clockwise and which represent the area of gas extracted by a well pad with 6 wells. Bungee uses a complex spatial optimization algorithm to find many infrastructure layouts within a site (Chapter 3 Appendix). The first such layout is a cost-minimizing layout that ignores environmental impacts other than those already imposed by regulation. Subsequent layouts reduce impacts at increasing cost, such that no final layout is simultaneously more impacting and more costly than any other.

We were forced to adjust the Impact Scores associated with layouts produced by Bungee in order to use them in our analysis here. The Impact Score aggregates across several metrics of environmental impact to represent the total impact of a layout in a site (Chapter 3 Appendix). It is formulated in such a way that Impact Scores at one site cannot be directly compared to Impact Scores at another site, which violates two conditions necessary for this analysis, including that impacts can be added across sites and that impacts can be traded (or offset) from one site to another. To get around this, we recalculated the Impact Scores associated with each of our layouts to make them comparable (p. 139). After transformation 28 sites had layouts that violated the Pareto conditions necessary for the analysis. We chose to exclude those 28 sites, leaving us with 56 sites in total (§4.6.1 in Appendix). Those 56 sites range in size (1-14 well pads or 6-84 wells) and number of layouts (2-16).

We also adjusted the construction costs estimated by Bungee to fit with this analysis. Bungee already estimates the construction cost of surface infrastructure, but this excludes many other costs associated with developing a site, including acquisition and leasing, below-ground infrastructure, and processing of gas (Hefley & Seydor 2015). We used the costs calculated for a single well and summarized in Table 8 of Hefley and Seydor (2015), excluding 'Permitting' and 'Site Preparation', which Bungee already includes. For most costs, we multiplied these single well costs by the number of wells in a site and added it to the surface infrastructure costs. For acquisition costs, we multiplied by the number of well pads in a site since Hefley and Seydor (2015) base acquisition cost on a single drilling unit corresponding to one pad. This approach likely overestimates the costs of developing a site. To calculate the present value of gas in a site which is used to create the "dummy" layout described in §4.3.3, we used a linear estimate of the flow rate of gas from wells in our study area and combined that with a constant market price of gas and monetary discount rate (p. 140). Gas was assumed to flow until the rate became zero, i.e. the gas ran out. We assumed every well would produce the same amount of gas at the same rate and simply multiplied the number of wells in a site by the present value of gas in a well to get the present value of gas in the site.

We analyzed the total cost and total impact across our 56 case study sites for various caps on impact and many sensitivity tests of the error in $\hat{I}_i(0)$. For *Uniform Cap without Trade*, we analyzed the system for 40 values of A between 0 and 5. The lower bound was chosen to show where zero impact was allowed, while the upper bound ensured that the cap would exceed any single site's maximum impact. For *Omniscient Social Planner*, we analyzed the system for 40 values of α between 0 and 1 to look at the full range of impacts. For *Cap and Trade*, we analyzed the system for combinations of α and error in $\hat{I}_i(0)$. As before, we looked at 40 values of α between 0 and 1. To look at the scenario where $\hat{I}_i(0)$ is high or low systematically, we added or subtracted, respectively, some portion $I_i(0)$ (Table 4.2). When looking at

Table 4.2. Analysis parameters showing various caps on impact set by the regulator as well as error in the regulator's estimate of impacts in the absence of additional regulation. Cap is absolute and at the site-level for *Uniform Cap without Trading* and relative to total impact and site-level impact for *Omniscient Social Planner* and *Cap and Trade*, respectively. Error is a proportion of the impact from the least-cost layout (j = 0) added to that impact.

Scenario	**Cap (A or α), n=40**
Uniform Cap without Trading	0, 0.13, 0.26, ..., 5
Cap and Trade	0, 0.03, 0.05, ..., 1
Omniscient Social Planner	0, 0.03, 0.05, ..., 1

Cap and Trade Error Direction	**Error Level (ε)**
Uniform Unbiased	0.1, 0.25, 0.5, 0.75, 1
Systematic Overestimate	0.1, 0.25, 0.5, 0.75, 1
Systematic Underestimate	-0.1, -0.25, -0.5, -0.75, -1

the effects of random error, we added $I_i(0)$ by a uniformly drawn random portion between $-\varepsilon$ and ε, the maximum amount of error. In other words, some sites received a positive error while others a negative error. We repeated this process 100 times for each ε and summarize the range of results. For instance, an error of ε = 0.5 would result in $\hat{I}_i(0) = 1.5I_i(0)$ for the systematic case and $-1.5I_i(0) \leq \hat{I}_i(0) \leq 1.5I_i(0)$ for the random case. We summarize these scenarios in Table 4.2.

4.4 Results

4.4.1 General

Fig. 4.2 summarizes our results conditioned on 1) the regulator does not know or use site-specific information to set the cap for *Uniform Cap* *without Trading* , 2) the regulator knows $I_i(0)$ perfectly for *Cap and Trade*, which is benchmarked against an *Omniscient Social Planner* with perfect information. Later we relax the assumption the regulator knows $I_i(0)$ exactly. In the figure, outcomes in the lower-left corner represent the business-as usual situation where no attempt is made to regulate impacts and all sites develop their least-cost, highest-impact layout. The total impact avoided is zero while the total cost is ~0.05% of the situation where no sites are developed. In the upper-right of Fig. 4.2a is the outcome where no impacts are allowed and as a result no sites are developed. The total impact avoided is 100% while the total cost is the cost of foregone profits from all sites (100%). Between these extremes developers vary the choice of layout in their site or choose not to develop such that some (black) or all (gray) sites are developed.

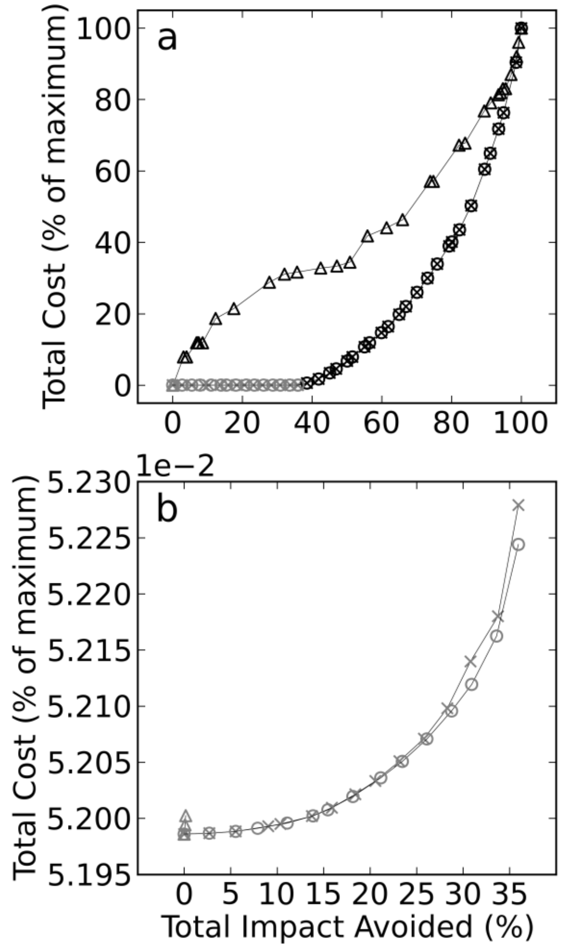


Fig. 4.2. Outcomes of implementing various policies as regulator's commitment to reducing impacts increases over a) entire range of possible outcomes and b) over range of outcomes where all sites develop. Triangles (△) are *Uniform Cap without Trading*, crosses (×) are *Omniscient Social Planner*, and circles (○) are *Cap and Trade* when regulator is perfectly able to estimate impacts in the absence of additional regulation. Vertical axis is percent of outcome where no sites develop. Gray symbols show where all sites are developed. Black symbols show where at least one site is not developed. There are three outcomes where all sites are developed in *Uniform Cap without Trading*.

Uniform Cap without Trading

There are several interesting characteristics of outcomes from implementing a *Uniform Cap without Trading* regulation. First, outcomes sit near an imaginary one-to-one line up until about 35% of potential impacts are avoided (Fig. 4.2a, triangles left of 35% Impact Avoided). For instance, avoiding 32% of potential impacts would require 31% of the cost of developing no sites. After 28%, there is a large amount of potential impacts that could be avoided for relatively little additional cost. For instance, reducing potential impacts from ~30% to ~50% would be ~4% more costly (Fig. 4.2a, middle plateau of triangles). After avoiding 61% of impacts, further avoidance becomes quickly more expensive up to 100% avoidance (Fig. 4.2a, triangles right of 61% Impact Avoided). Moving down to Fig. 4.2b reveals that there are three options for *Uniform Cap without Trading* where all sites are developed and these reduce potential impacts very little.

Cap and Trade and *Omniscient Social Planner*

Both *Cap and Trade* and *Omniscient Social Planner* start with very low costs to avoiding impacts (Fig. 4.2a circles and crosses). Up to 36% impact avoidance, the cost of avoiding impacts appears to be almost zero, though inspection of Fig. 4.2b shows that these costs are very low (<0.06% of maximum cost), but nonzero. Further impact avoidance greater than 40% incurs quickly escalating costs. For instance, avoiding impacts up to 80% requires about 40% of maximum costs (Fig. 4.2a, circles and crosses). Nearing 100% impact avoidance greatly escalates costs (Fig. 4.2a circles and crosses right of 80%).

Does *Cap and Trade* achieve minimum costs?

As expected by theory (Hartwick & Olewiler 1998), implementing *Cap and Trade* results in outcomes that are nearly identical to *Omniscient Social Planner*. As can be seen in Fig. 4.2b, we estimate even lower costs at the same level of impact reduction when implementing *Cap and Trade*, but this is due solely to the way we estimate the outcomes of a cap and trade system. A more accurate analysis would show that *Cap and Trade* performs no better than *Omniscient Social Planner*. We focus on the comparison of *Cap and Trade* and *Uniform Cap without Trading* hereafter.

How do policy scenarios compare when the regulator perfectly estimates impacts in the absence of additional regulation?

There are important differences in the total cost and impact of implementing different regulations. *Cap and Trade* achieves a lower cost at a given level of impact avoidance than *Uniform Cap without Trading*. This improvement is not small over most of the range of possible outcomes, and is as large as ~30% when avoiding ~36% of total impacts (Fig. 4.2). The two most likely regulatory scenarios differ wildly in how much potential impacts can be avoided (Fig. 4.2b). If *Cap and Trade* is implemented, up to ~36% reduction in potential impacts can be achieved for only 0.05% of the maximum cost while still allowing all sites to be developed. Compare this to the *Uniform Cap without Trading* scenario, for which very little (~0.1%) impact avoidance is possible while allowing all development. In addition, there are many more options for how much impacts are avoided while allowing all development under the *Cap and Trade* scenario.

The distribution of outcomes along the horizontal axis in Fig. 4.2a is also interesting. First, *Uniform Cap without Trading* exhibits a less smooth spacing of outcomes, which is a result of the way the regulation is implemented. Outcomes that are close together are similar in that the set of sites developed does not change from one outcome to the next, but only the set of layouts chosen for development. Large jumps between clusters of outcomes are due to one or more sites being pushed out of development by a reduction in the site-level cap. This discontinuity in outcomes means that small regulatory adjustments may have little effect on resulting impacts and costs. Contrarily, the other policy scenario has a smoother distribution of outcomes because the regulation allows more flexibility in how sites are developed. As a result, small policy adjustments more often affect the system.

4.4.2 How is the cost effectiveness of Cap and Trade affected by the regulator's ability to estimate $I_i(0)$?

<u>Idealized Scenarios</u>

Having identified *Cap and Trade* as a potentially very cost effective policy choice, we now explore how sensitive that finding is to our underlying assumptions. Specifically, we focus once again on our information assumptions. This time, though, we emphasize $I_i(0)$ used to set the site-specific cap, which is an obvious target for a sensitivity test because it is a quantity which developers will know much better than the regulator and which developers will have an incentive to hide. Note that in *Uniform Cap without Trading*, we hold to the previous assumption that the regulator does not know or chooses not to implement a site-specific cap.

In Fig. 4.3, we summarize the results of our sensitivity tests to explore how the cost effectiveness of *Cap and Trade* is affected by error in the regulator's ability to estimate impacts in the absence of additional regulation, denoted $\hat{I}_i(0)$. We calculated cost effectiveness as the ratio of system total absolute impact avoidance to total cost in billion USD. To plot all policy scenarios on the same horizontal axis, we transformed the cap (A) for *Uniform Cap without Trading* to a relative scale by dividing by the maximum value.

Before discussing the results of the error analysis, we first orient the reader to the idealized cost effectiveness curves (black open symbols in Fig. 4.3a-c). At regulator choices below ~0.15, there are no outcomes for *Uniform Cap without Trading* where sites are developed. The next two outcomes are two of the most cost effective for *Uniform Cap without Trading* because they reduce impacts a little while still allowing all development to proceed. At further commitments to reducing impacts, *Uniform Cap without Trading* outcomes have a low but increasing cost effectiveness, with a peak near (0.55, 0.01).

Cap and Trade and the idealized benchmark *Omniscient Social Planner* both exhibit almost identical outcomes. Increasing commitment to reducing impacts lead to increasingly more cost effective outcomes up to the point just before one site is pushed out of production. Peak cost effectiveness where all sites are developed occurs at (0.64, 8.5) in Fig. 4.3.

In the ideal case, *Cap and Trade* is generally more cost effective than *Uniform Cap without Trading*; only at the extreme regulator choices do the two scenarios converge, which is a necessary result. Although both scenarios have peak effectiveness up to the point just before one site is pushed out of production (highest triangle and circle in Fig. 4.3), the cost effectiveness of *Cap and Trade* at its peak is more than two orders of magnitude more cost effective than *Uniform Cap without Trading* at its peak. At higher commitments to reducing impacts (Fig. 4.3 right of ~0.4), the two scenarios have more similar cost effectiveness, but *Cap and Trade* is still five times more cost effective than *Uniform Cap without Trading* on average.

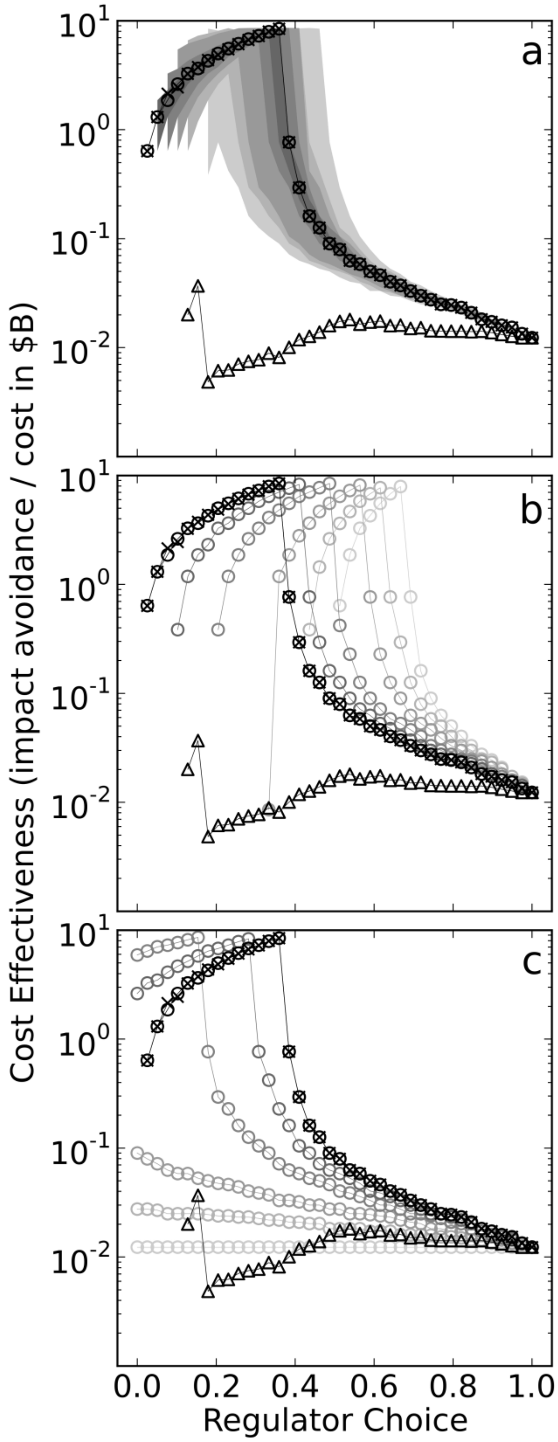
Error in $\hat{I}_i(0)$ affects the outcomes of implementing *Cap and Trade*.

We explored three types of error in $\hat{I}_i(0)$: uniformly distributed (Fig. 4.3a), systematically high (Fig. 4.3b), and systematically low (Fig. 4.3c). Because outcomes are based on the choice of layouts at sites and these choices are highly discrete, error in $\hat{I}_i(0)$ serves mainly to stretch or compress the distribution of outcomes as the regulator's choice changes rather than reveal entirely different outcomes.

Uniformly random error in $\hat{I}_i(0)$

When error in $\hat{I}_i(0)$ is uniformly random across sites, outcomes may be more or less cost effective at a particular regulator choice. For instance, when $\hat{I}_i(0)$ is up to 100% different from $I_i(0)$ (lightest gray region in Fig. 4.3a), the regulator's decision to reduce potential impacts by 30% may lead to a cost-effectiveness a full order of magnitude lower than if the regulator can perfectly estimate $I_i(0)$. Uniformly unbiased error tends to lead to lower cost-efficiency outcomes rather than higher as reflected by the wider range of outcomes below/left of the perfect-estimate outcomes in Fig. 4.3a.

Fig. 4.3. Effect of error in regulator's estimate of impacts in the absence of additional regulation ($\hat{I}_i(0)$) for Cap and Trade when error is a) uniform but unbiased, b) systematically high, and c) systematically low. Horizontal axis ranges from no commitment to reducing impacts (0.0) up to no allowance of impact (1.0). In all panels, outcomes from a zero-error estimate are shown in black, while increasingly lighter gray shows outcomes with increasing error. Triangles (△) are Uniform Cap without Trading, crosses (✕) are Omniscient Social Planner, and circles (○) are Cap and Trade. Error levels are summarized in Table 4.2. In a) shaded regions show a range of outcomes over 100 trials of uniformly distributed error in $\hat{I}_i(0)$.


a
b
c
Cost Effectiveness (impact avoidance / cost in $B)
Regulator Choice
0.0
0.2
0.4
0.6
0.8
1.0

Systematically overestimating $I_i(0)$

Systematically overestimating $I_i(0)$ compresses the possible outcomes from *Cap and Trade*, which has several effects on regulation (Fig. 4.3b). Low levels of commitment to reducing impacts may not reduce impacts at all, since developers will not have to change their choice of layout to meet the impact cap (leftmost points for *Cap and Trade* in Fig. 4.3b). Once the cap is high enough to affect developer's choices, a systematic overestimate of $I_i(0)$ will lead to lower cost efficiency of outcomes up to the point where all sites are developed in the $I_i(0) = \hat{I}_i(0)$ case. After this, a systematic overestimate of $I_i(0)$ leads to higher cost efficiency of outcomes. Again, this is due to the fact that it takes larger commitments to reducing impacts to achieve the same outcomes as when $I_i(0) = \hat{I}_i(0)$. At larger error levels, higher cost-efficiency outcomes are more likely, but at a much increased risk of having no effect on development at lower commitments to reducing impact.

Systematically underestimating $I_i(0)$

Systematically underestimating $I_i(0)$ stretches the possible outcomes from *Cap and Trade*, which has several effects on regulation (Fig. 4.3c). When error is low, lower commitments to reducing impacts lead to higher efficiency outcomes. However, at error levels larger than 25%, any commitment to reducing impacts will lead to a lower efficiency outcome. At very high error levels, the cost efficiency of *Cap and Trade* may even be lower than *Uniform Cap without Trade* (lightest gray circles are below some triangles in Fig. 4.3c). Since the regulator is underestimating impacts at sites, caps on impact will be almost guaranteed to affect developers' choices of layouts and consequently lead to lower-impact outcomes, yet this comes with a risk of lower-efficiency outcomes and increased probability of pushing sites out of production.

4.5 Discussion

Ongoing shale gas development creates environmental externalities which may be internalized and reduced at reasonable costs through cap and trade. We have analyzed two policy scenarios that may be implemented and how these compare to one another and a best-case scenario in terms of their total resulting impact and monetary cost. We found that the policy scenario most reflective of current regulations (*Uniform Cap without Trading*), which forces developers to reduce impacts in a uniform fashion or not develop, may lead to expensive outcomes with few options to reduce impacts while still allowing all development to proceed (Fig. 4.2a). In contrast, a cap and trade scenario (*Cap and Trade*) could perform as well as would an omniscient social planner by avoiding impacts across 56 sites in Pennsylvania by ~36% for 0.05% of the cost of not developing any sites and while still allowing all development to proceed (Fig. 4.2b). A similar level of impact avoidance in the *Uniform Cap without Trading* scenario would be close to 35 percentage points more costly. The relative costs of *Cap and Trade* versus *Uniform Cap without Trading* converge at higher or lower levels of avoidance. For instance, at 20% and 80% impact avoidance the difference between the two scenarios is 20 and 25 percentage points, respectively. However, we also determined that the ability of a regulator to match its commitment to reducing impacts to actual outcomes depends on the regulator's ability to estimate impacts in the absence of the new regulation. For instance, *Cap and Trade* could be totally ineffective if the regulator systematically overestimated those impacts and had a low commitment to reducing impacts (Fig. 4.3b). Similarly, the cost effectiveness of *Cap and Trade* could be almost three orders of magnitude lower than ideal if the regulator systematically underestimated those impacts by more than 25% (Fig. 4.3c).

Our results have several implications for policy design and implementation. Cap and trade can offer large savings over a more traditional uniform and inflexible approach, which agrees with theory (Hartwick & Olewiler 1998; Goulder & Parry 2008). Further, we find it can reduce impacts much more while allowing all development to proceed. At the same time, implementation efforts are not the same for the two approaches. In either scenario, the regulator needs to enable the gas industry to evaluate impacts produced by an infrastructure layout, which requires the regulator know what impacts are relevant, what priority they have, and how they are calculated. Additionally, both scenarios require the monitoring of surface development, which could be attached to current drilling permitting processes. Intentionally cost effective cap and trade as outlined here requires that the regulator must be able to estimate impacts in the absence of the new regulation. This requires some knowledge of the development process in the regulatory region. At minimum, a regulator could evaluate existing development. A more detailed method that models development could ensure higher cost effectiveness. Cap and trade also relies on the distribution, tracking, and enforcement of tradable permits, which could be accomplished with an online market system. We expect the total cost of implementing cap and trade at intermediate levels of impact avoidance would be compensated by the long-term savings over the other inflexible approach we explored (see Fig. 4.2).

Although the level of commitment to reducing impacts would ideally be determined by society's value of impact avoidance, our results do suggest that moderately low commitments will be most cost effective (Fig. 4.3) with cap and trade. In our case study, the highest cost-effectiveness was achieved by committing to impact reduction around 36%, which is contingent on being perfectly able to estimate impacts in the absence of the cap and trade system. At commitments lower than 36%, *Cap and Trade* was still more cost effective than commitments larger than 36%. Interestingly, the hump shape of the cost effectiveness curve which leads to this outcome is due to a combination of two things. First, our estimates of the cost of developing a site is much lower than our estimates of the profits from gas extraction. As a result, not developing a site leads to large increases in the cost of the system. Second, there is large potential to reduce impacts of the system while still allowing all development to proceed (Fig. 4.2). Combined, large reductions in impacts can be achieved without increasing costs a lot relative to profits gained from development (Fig. 4.2 gray circles). When commitments to reducing impacts exceed a certain level (36% here), some sites are forced out of production leading to large increases in cost for relatively little change in overall impacts, which greatly reduces the cost effectiveness of the system (Fig. 4.3: switch happens where the slope of *Cap and Trade* becomes negative as regulator choice approaches 1).

One purpose of this study was to apply existing knowledge about the relative cost effectiveness of market-based policies to inflexible uniform policies in the shale gas context. We show clearly some of the potential gains from trade created by a cap and trade system that regulates an aggregate impact metric. Other approaches may also be effective in this and other contexts. For instance, cap and trade for individual metrics (e.g. forest clearing) might increase the transparency and understanding of the market and increase support, though at an increased implementation cost due to maintaining multiple markets. In addition, a bubble-offsets approach might obviate the need for a market, especially when there is large spatial heterogeneity in the cost of reducing impacts at individual sites. A bubble policy would treat a subset of sites that are close to one another or have the same developer as a single unit ("bubble"), putting a cap on total impacts within the bubble (Tietenberg 1985). Similarly, when development rights across all sites are held by just a few developers, enforcing a cumulative cap for each developer could be effective. To be effective this would require that each developer has development rights at sites with heterogeneous costs of impact avoidance. Many other alternatives exist. We took an approach that should be generally applicable across regions where many sites are ready for development, where developers have rights to one or a few sites, and where reducing aggregate impacts is the major goal.

More complete analyses could benefit from several adjustments to our methodology. First we assume that each developer has rights to only one site being developed. In Pennsylvania, there are many developers, but the distribution of development is skewed towards developers with many holdings (Pennsylvania Department of Environmental Protection's permit reporting database). When combined with assumptions about market dynamics, it is likely that those developers with many sites would exert a measurable effect on the market and could compromise the effectiveness of the market (Hartwick & Olewiler 1998). Second, we assume all sites are to be developed simultaneously and thus enter the market simultaneously. A more complete analysis on a small market would include the staggering of development over time and adjust developer's decisions about when to develop (i.e. enter the market). Third, we assume impacts are independent across sites and thus can be combined additively. One alternative approach would be to treat nearby or adjacent sites as having dependent impacts, e.g. by combining their development boundaries to treat them as one unit when evaluating impacts. This approach would require a more complex decision process as well as stricter assumptions about the simultaneity of development across sites. Finally, we chose to focus on the direct regulation of a single aggregate metric such that trading among individual metrics could occur at the site level. This choice ignores one alternative approach to regulating multiple impacts, which is to put a cap on each individual metric. While this approach would more directly enforce local priorities for each impact, it would limit development options within sites. Further, because many impacts are positively correlated and some are negatively correlated (Chapters 2 and 3), the link between an impact's cap and the resulting development choice could be confounded by choices driven by other impact caps (Bennear & Stavins 2007), and thus presents a challenge to matching environmental goals to policy outcomes. This is a unique characteristic of regulating multiple impacts through multiple, impact-specific caps.

We have applied existing cost effectiveness analysis methods to the novel context of regulation of environmental impacts from shale gas surface development and found that large gains from trade are possible. As shale gas development proceeds globally, governments at multiple levels should consider the environmental implications of shale gas extraction and design policies that properly internalize environmental externalities. In regions where development rights are centrally owned or distributed, significant environmental savings can potentially be achieved without the need for additional regulations. In other regions, our findings can be used to motivate regulations that do better than traditional command-and-control approaches. As such, we see large potential to develop shale gas more conscientiously in the coming decades.

4.6 Appendix

4.6.1 Reconfiguring Impact

Impact Scores output by Bungee are not comparable across sites. However, we need something comparable/tradable to be able to allocate impacts across sites. To accomplish this, we do the following:

1. Calculate individual impact metrics $f_{m,i}(X_i)$ where m is the metric name, i is the index of the site, and X is the layout.
2. Likewise, calculate the baselines b_i at each site. In our study, the baseline for all metrics is 0.
3. Find the maximum unit impact for each metric across layouts: $(f_m/w)_{max} = \max\big(f_{m,i}(X_i)/w_i \; \forall \, X_i\big)$.
4. Scale each metric for each layout using [3]:

$$g_{m,i}(X_i) = \frac{\left|f_{m,i}(X_i) - b_i\right|}{w_i(f_m/w)_{max}}$$

5. Finally, we calculate the new Impact Score for each layout as

$$\text{Impact Score}(X_i) = \sum_{\forall\, m} v_m \ln\big(g_{m,i}(X_i) + 1\big)$$

   where v_m is the relative priority of the metric. This gives us a way to trade impacts across sites since the scaling values across sites are the same.

This transformation keeps the general shape of tradeoff curves across sites, but fails in one important respect. Many tradeoff curves (28 of 84 sites representing 270 of 607 layouts) no longer satisfy the Pareto-improvement shape needed for our analysis (see Fig. S 4.1). The reason these curves are no long monotonically increasing is because the scaling values for each impact metric have changed relative to other impact metrics in the site. As such, the relative contribution of each metric to the Impact Score changes which causes some layouts to fall away from the newly defined Pareto frontier. To get around this, we simply omit from our analysis those sites whose tradeoff curves violate the Pareto-improvement conditions. We were left with the 56 sites (Fig. S 4.2) with black tradeoff curves in Fig. S 4.1.

4.6.2 Present Value of Gas Production, V_i

We calculate the profit of gas production in a site using the present value. We base our estimates on the flow rate of gas from wells in the counties from our sites in our previous study as well as recent national well-head gas prices.

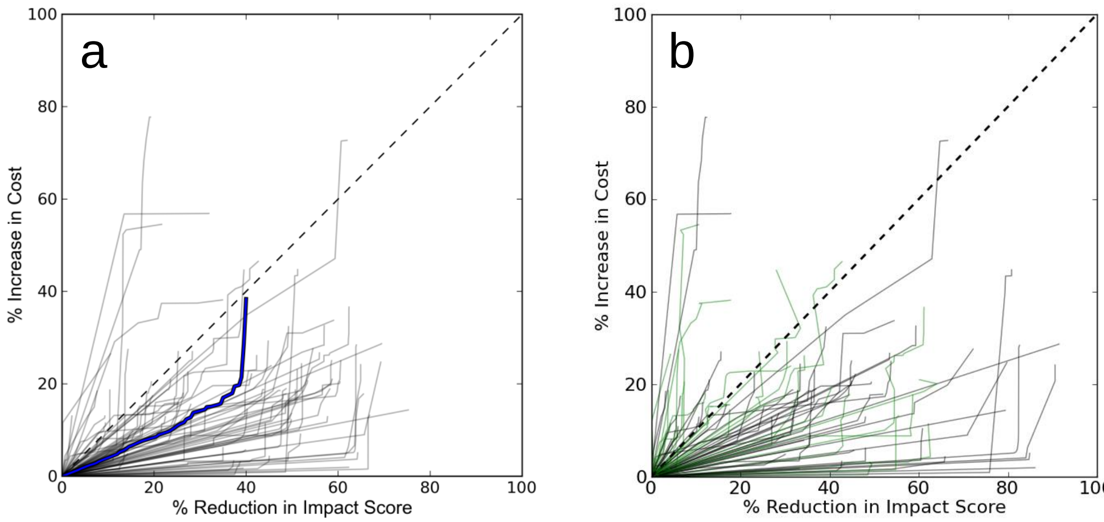


Fig. S 4.1. (a) results from our previous study, where Impact Scores are calculated using site-specific normalization constants for each impact metric and tradeoff curves are monotonically increasing. (b) results of re-scaling impact metrics to calculate tradable impacts, the result of which is that many tradeoff curves (green, n=28 of 84) are no longer monotonically increasing, which violates the conditions necessary for optimal allocation of impacts.

The present value V_i in a site depends on several things:

1. q_0, the initial rate of flow of gas from when production begins (Mcf/day)
2. λ, the change in the flow rate from a well per unit time (Mcf/day^2)
3. d, the monetary discount rate
4. w_i, the number of wells in the site

The flow of gas from a well at a particular time t is approximated by the linearly declining curve

$$q(t) = q_0 - \lambda t \qquad (4)$$

with units of <amount of gas> per <unit of time>. Note that several other more complex and accurate estimates of gas flow exist that account for linear decrease in flow in the early portions of the well's life followed by declining marginal returns later (Al Ahmadi, Almarzooq & Wattenbarger 2010; Miller, Jenkins & Rai 2010). We chose this one because the present-value of gas is relatively easy to calculate with it and it provides estimates of the total amount of gas in a well that agree with other findings (Aucott & Melillo 2013). Further, many wells exhibit linear decreases in flow during the first decade of production (Al Ahmadi, Almarzooq & Wattenbarger 2010; Miller, Jenkins & Rai 2010). The simplest alternative is the exponential decay curve, but using this gave estimates of the total amount of gas in a well three orders of magnitude above (Aucott & Melillo 2013).

The value of gas at a particular time is just $p_{gas}q(t)$, where p_{gas} is the price of gas. The present value of gas from all wells in a site (assuming all wells begin production at the same time) is given by

$$V_i = w_i \int_{t=0}^{q_0/\lambda} (1+d)^{-t} p_{gas}(q_0 - \lambda t)$$

The integral goes to q_0/λ since this is the point at which $q(t) = 0$, where the estimate of gas flow breaks down.

$$V_i = \frac{w_i p_{gas}}{\ln^2(d+1)}\left[\lambda(d+1)^{-\frac{q_0}{\lambda}} + q_0 \ln(d+1) - \lambda\right] \qquad (5)$$

We estimated these values from existing production data in Pennsylvania and our sites from our previous study. We estimate q_0 and λ by fitting Eq. (4) to 827 production time-series across 5 counties in Pennsylvania (taken from https://www.paoilandgasreporting.state.pa.us/publicreports/Modules/Production/ProductionByCounty.aspx). For each well, we estimated q_0 and λ. If $\lambda \geq 0$ or there were less than 5 data points, we discarded the estimate, which reduced our count from 1,801 to 827 wells. We then took the average values across the wells. We found an average $\overline{q_0} = 3647\ \mathrm{Mcf} \cdot \mathrm{day}^{-1} (\sigma = 1952)$ and $\bar{\lambda} = 2.62\ \mathrm{Mcf\ day}^{-2}$ $(\sigma = 1.70)$. For the price of gas we used 4.5465 USD per 1000 ft3 - the average wellhead price reported between 2008 and 2012 (most recent year) from the EIA, which is the period over which development we study took place (http://www.eia.gov/dnav/ng/ng_pri_sum_dcu_nus_m.htm). These data from the EIA are at the national level but do not include Pennsylvania. We used a discount rate of $d = 0.05\ \mathrm{yr}^{-1} = 0.00001369863\ \mathrm{day}^{-1}$.

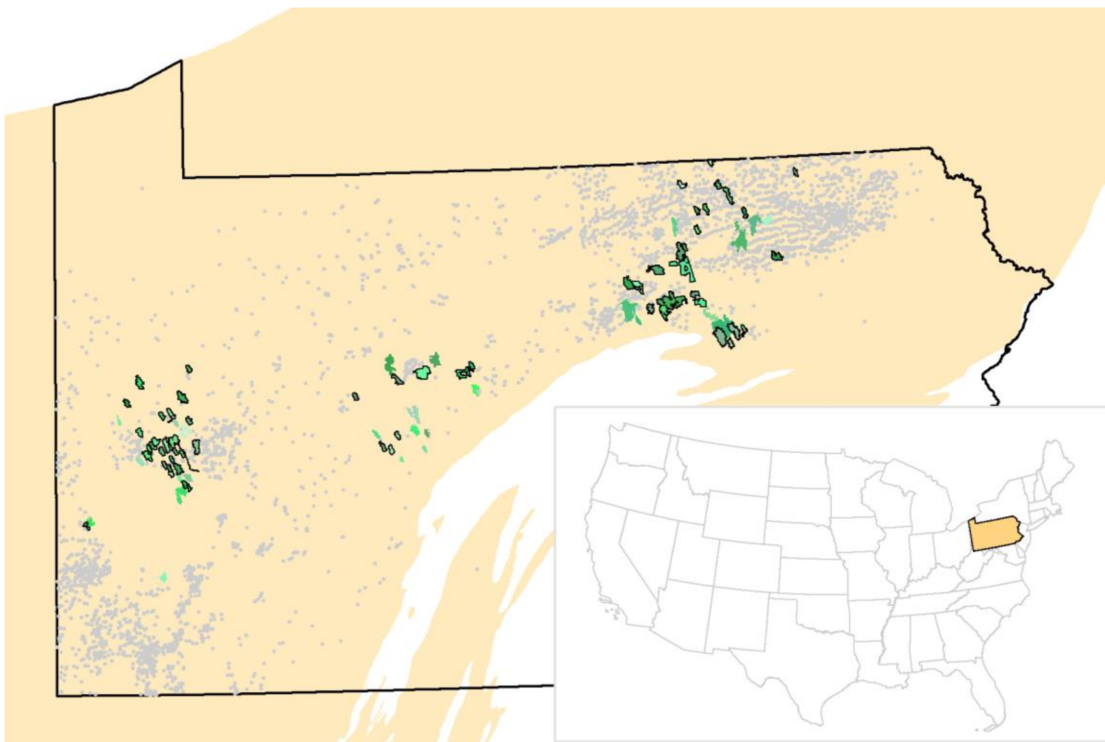


Fig. S 4.2. Development sites with grouped well permits (gray) and the Marcellus shale play (beige). Sites were derived by overlaying production units on existing well permits and taking contiguous land parcels under those production units by a single operator. Of the 84 candidate sites (green polygons), 56 (black outlines) were analyzed after transformation of Impact Scores.

In our previous study we assumed every pad would have six wells, but they varied by the number of pads. In Fig. S 4.3 we show the distribution for the present-value of gas across sites. We calculated a total volume of gas for each well at 1.4 Bcf, which is on the lower end of the best estimates from (Aucott & Melillo 2013).

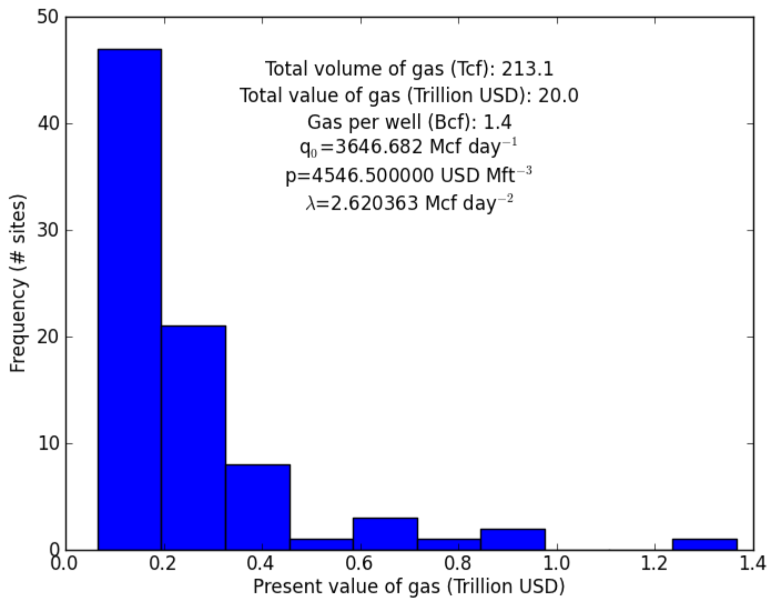


Fig. S 4.3. Distribution of gas values across our 56 study sites.

Conclusion

As the world changes, it is critical for conservation planning as a practice and science to change in accordance. In my dissertation I created methods and performed analyses to advance the science and practice of applied conservation planning. In Chapter 1 I assessed how new records of species observations change spatial priorities in Tennessee and found that when planning for complementary richness new observations will continue to affect priorities in the coming years. In Chapter 2 I evaluated the relative performance of simple planning guidelines to reduce environmental impacts of surface infrastructure (“impacts”) for shale gas and how such planning produces tradeoffs between impacts. I found that while not pervasive, there are tradeoffs between impacts and that single guidelines do not perform best across all impacts. In Chapter 3 I explored site-level costs of reducing impacts and found heterogeneity across sites in both the ability to reduce impacts and the relative costs of doing so. In Chapter 4 I analyzed the cost effectiveness of different regulations for reducing impacts across multiple sites and found large potential savings by using a market-based, cap and trade approach as opposed to a traditional uniform, inflexible approach.

My dissertation has many implications for conservation science and practice. For instance, the results of Chapter 1 suggest that biodiversity sampling is most important in the early stages of conservation effort, but even at later points continued data collection is useful for determining spatial priorities. The results of Chapter 2 justify the use simple planning guidelines in sites where major impacts and conservation priorities are few and obvious, but that tools that explicitly reduce multiple potential impacts are called for in many places. From Chapter 3 I can advocate that while many sites will be able to avoid impacts inexpensively when taking an aggregate approach, a uniform policy that requires developers to act at a median standard will push many sites out of production, perhaps unnecessarily. Finally, the findings of Chapter 4 endorse market-based tools for reducing impacts from shale gas surface infrastructure, especially when regulators make an intermediate commitment to reducing impacts.

Several important qualities of my dissertation are apparent when looking across multiple chapters. First, I find that the granularity of biodiversity data used to define benefits for spatial planning is influential for planning outcomes (Sutton & Armsworth 2014). Species are not distributed randomly on the landscape at large or small scales (Gaston 2000). Combined with the types of data – usually opportunistic, presence-only, element occurrence points of rare species of interest – and the result is very grainy biodiversity data. In Chapters 1 and 3 I find that grainy biodiversity data play a key role in my results. In Chapter 1, biodiversity data are concentrated in relatively few sites across Tennessee. This is partially due to the natural distribution of biodiversity as well as sampling bias inherent in an opportunistically collected dataset. The complementary richness of sets of sites for priority action responded to the granularity of the biodiversity data, and this led to qualitatively different outcomes than an approach that focused on site-specific species richness. In Chapter 3, large differences in the numbers of observed species between spatially adjacent habitat types produced a very grainy planning surface for the *rare spp.* impact metric. As a result, in some cases the impacts of expected infrastructure could be greatly reduced by relatively small changes to infrastructure placement. Further, *rare spp.* contributed significantly more to reductions in the *Impact Score* than the other four metrics included in that chapter.

A second important quality of my dissertation is that it crosses multiple scales in space and time. The scale of decision making is essential to consider when designing applied conservation planning research (Spies & Johnson 2003). In Chapters 2 and 3, I focus on site-level planning that takes place over a few months. At this scale, developers and conservation groups are making decisions with observable effects on individual organisms and populations that extend throughout the site and to adjacent areas. This is the scale at which development decisions most directly impact conservation priorities. In Chapters 1 and 4 I look at a scale of planning across many sites and longer times. At these larger scales, decisions are being made by individuals higher in organizations and in organizations that operate at larger scales. The information from decisions at smaller scales may play an important role in these larger scale decisions (Spies & Johnson 2003).

Third, my dissertation chapters are all motivated by a need to do science that is immediately useful for practice. For instance, In Chapter 1 I make recommendations for the immediate use of my results for large scale planning in Tennessee. I also outline strategies for long-term planning in other places, especially where conservation actions are just beginning. In Chapter 2 I evaluate simple practices for planning surface infrastructure to see in what situations simplified planning is sufficient for avoiding multiple impacts. This kind of "rule of thumb" research is often a goal of conservation groups working to create best management practices or other strategies that can be applied without huge data or analysis resources (*pers. obs.*). I go a step further in the later chapters by creating Bungee, a decision-support tool for planning surface infrastructure to avoid aggregate environmental impacts. Bungee is made most directly for use by the gas industry, but also has defined uses for conservation stakeholders and regulators (see Chapter 3 Discussion). At the time of writing, EnSitu – the ArcMap Python toolbox that uses

Bungee as the analysis engine – was either being used or was intended to be used by 27 groups spanning the conservation, government, industry, and research sectors. EnSitu is a streamlined front-end for Bungee. It combines the individual submodules of Bungee into a simpler two-step process in ArcMap. Bungee has its own visual interface, but running it entails more work while offering greater flexibility within ArcMap. Because of its streamlined structure, EnSitu is the product presented to primary users.

One challenge I faced in my dissertation was to balance the needs of science with the needs and limitations of conservation practice. In all my chapters I use almost exclusively publicly available datasets, datasets which are used in practice. These datasets are free and usually large, but are not always of a quality or information content that ideally suits the scientific questions being asked. However, in advancing science that advances practice, it is important to reflect the decision processes and inputs used by decision makers. Another notable place where my science met conservation practice is in the design of Bungee. The planning problem addressed by Bungee is one that, in order to solve globally for realistic scenarios, would take massive computational resources and/or a very long time. However, in designing Bungee as a tool for use outside of academia, we constrained ourselves to solutions that could be derived on the order of hours on a desktop personal computer. As a result, Bungee is actually useful for surface infrastructure planning, though at a cost of confidence in the overall (i.e. global) optimality of results.

There is no shortage of free spatial conservation planning tools, and as a result much of conservation is (sometimes inappropriately) performed with existing tools (*pers. obs.*). In my dissertation and with Bungee I took a different approach. I chose to design a tool for a specific conservation planning problem, namely the planning of shale gas surface infrastructure to avoid multiple environmental impacts. This decision created some distinct advantages, the largest of which is that I was able to tailor my solution to that problem rather than sacrifice on the relevance of the solution. As a result, my scientific findings address the problem more accurately and represent the actual decision process more directly. In addition, end users of Bungee benefit because they can work in an environment more similar to their experiences, which increases buy-in and encourages appropriate use of the software. At the same time, tailoring of software presents challenges relative to using existing tools, since programming, datasets, test scenarios, case studies, and documentation must all be created from scratch. The time and effort to tailor solutions to problems is worthwhile if it leads to those solutions being used in practice.

Bungee is not another Marxan (http://www.uq.edu.au/marxan/) or Zonation (http://cbig.it.helsinki.fi/software/zonation), two of the most widely used conservation planning programs, because it takes a fundamentally different approach to conservation. First, the perspective is an intentional balance between conservation and development. On the surface, Bungee is a decision support tool for placing shale gas surface infrastructure. Using EnSitu (the visual interface) can be as simple as inputting two development-related shapefiles and clicking 'Go'. However, the primary driver of infrastructure placement is a suite of conservation-based, ecologically relevant environmental impact metrics. It is because of this balance that I normally think of Bungee's planning task as 'conservation-oriented development'. Second, Bungee plans multiple explicitly interacting geometries. It is not uncommon for spatially non-additive – think complementary richness – objectives to be part of conservation planning tools (Sarkar *et al.* 2006). But how Bungee differs is by planning well pad (points/polygons), road (lines), and pipeline (lines) features together in a way that goes beyond picking pixels to instead pick spatial configurations that would not be derived by thinking of the geometries independently. Third, Bungee represents a move from 'where to protect' to 'where to condemn'. This is not the first time such an approach has been taken (Moilanen 2012), but it is a recognition and acceptance that threats to conservation priorities are increasingly best handled by moderating those threats rather than attempting to prevent them.

Bungee is not limited to planning in the Marcellus shale play or to shale gas, but neither are its applications limitless. Within the shale gas context, Bungee will be most useful in its current form in areas where 1) the shale is moderately uniform thickness, 2) multiple well pads with flexible locations are planned simultaneously, and 3) impacts are spatially variable over small (tens of meters) scale. The Marcellus play in Appalachia is by design the most natural place for using Bungee, but groups in other parts of the U.S. and distant countries see low-effort extensions of the tool. Bungee could also be used for planning outside shale gas. In Chapter 3 I draw an analogy to rural home development, which is a comfortable context in which Bungee could be applied. Bungee could also be used to plan wind or solar installations. All of these applications can be achieved with little or no modification to the software, provided the main planning and impacting features are roads, transmission lines, and point/polygon features. Bungee is limited, however, in the scale of analysis. At the moment, the tool plans at 30-m pixel resolution. At resolutions coarser than ~60 m or finer than ~10 m, some of the built in assumptions of the software will be violated. This precludes two natural uses of the tool, which are to do fine scale final planning of infrastructure and separately to simultaneously plan infrastructure over whole regions.

One future research direction is to apply my fairly general methods to similar analyses at multiple scales across the globe. Since all of my chapters aim at providing recommendations for conservation practice now, performing such analyses in new places would benefit practitioners in those regions immediately. The challenges of applying my methods elsewhere are largely data challenges. In some of the places where my methods would be most useful, data are sparse and/or poorly maintained, and the only remedy is to collect data and perform intermediate analyses. In particular, using Bungee in new regions would also require an evaluation of conservation priorities, formulation of metrics of impact, and confirmation of model assumptions, none of which are trivial tasks.

Other possible research extensions include longer-term and ongoing research, especially with regards to shale gas development, which would increase the certainty of my results. Especially useful would be to see how implementing some of my recommendations affect the outcomes of future conservation-relevant activities. For instance, a comparison of development using 1) standard practices versus 2) Bungee versus 3) rules of thumb could strengthen the case for informed spatial planning I undertook. In addition, examining long time series of the effects of additional data on spatial priorities would inform the generality of my Chapter 1 results. Testing my assumptions about market dynamics and developer behavior in Chapter 4 in the lab using human subjects would inform the cost effectiveness of cap and trade for shale gas (Cason & Gangadharan 2003).

One question my dissertation raises is what should be the balance of regulated versus voluntary reduction of environmental impacts from shale gas development. On one hand, the primary goal of creating Bungee is to provide a tool for the gas industry to incorporate environmental priorities in their planning at low or no cost and minimal effort to them. Currently there are few strong incentives for the industry to go above and beyond regulation using Bungee. Namely, the main incentives I have identified are 1) to promote an environmentally friendly image, 2) to adhere to a personal or mission-based conservation ethic, and 3) to reveal the financial bottom-line (*pers. obs.*). Those incentives are probably overwhelmed by more certain financial incentives. However, pressure from society to be green may change that. On the other hand, new regulations can create incentives, e.g. by creating a market for impacts. However, there are many issues with creating new regulations. New regulations require administration, research, and infrastructure, all of which can be costly (Stavins 1995; Falconer, Dupraz & Whitby 2001; Joshi, Krishnan & Lave 2001). Poorly designed regulations can create moral hazards in which developers are incentivized to harm the environment more (Ozanne, Hogan & Colman 2001). Lastly, new regulations that internalize environmental externalities will be politically controversial and thus harder to institute. Though my dissertation does not inform the feasibility of reducing impacts voluntarily versus through regulation, it does provide information and tools for both resolutions.

Bibliography

Al Ahmadi, H.A., Almarzooq, A.M. & Wattenbarger, R.A. (2010) Application of Linear Flow Analysis to Shale Gas Wells - Field Cases.

Ahrends, A., Rahbek, C., Bulling, M.T., Burgess, N.D., Platts, P.J., Lovett, J.C., Kindemba, V.W., Owen, N., Sallu, A.N., Marshall, A.R., Boniface, E.M., Fanning, E. & Marchant, R. (2011) Conservation and the botanist effect. *Biological Conservation*, **144**, 131–140.

Ando, A. (1998) Species Distributions, Land Values, and Efficient Conservation. *Science*, **279**, 2126–2128.

Annual Energy Outlook 2014. (2014) Washington, DC.

Araujo, M.B., Thuiller, W., Williams, P.H. & Reginster, I. (2005) Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Global Ecology and Biogeography*, **14**, 17–30.

Arponen, A., Lehtomäki, J., Leppänen, J., Tomppo, E. & Moilanen, A. (2012) Effects of connectivity and spatial resolution of analyses on conservation prioritization across large extents. *Conservation Biology*, **26**, 294–304.

Aucott, M.L. & Melillo, J.M. (2013) A Preliminary Energy Return on Investment Analysis of Natural Gas from the Marcellus Shale. *Journal of Industrial Ecology*, **17**, 668–679.

Babcock, B.A., Lakshminarayan, P.G., Wu, J. & Zilberman, D. (1997) Targeting Tools for the Purchase of Environmental Amenities. *Land Economics*, **73**, 325–339.

Banzhaf, H.S., Burtraw, D., Chung, S., Evans, D.A., Krupnick, A. & Siikamaki, J. (2014) Valuation of Ecosystem Services in the Southern Appalachian Mountains. *Resources For the Future, RFF DP*, 14–16.

Bennear, L. & Stavins, R. (2007) Second-best theory and the use of multiple policy instruments. *Environmental and Resource Economics*, **37**, 111–129.

Bennett, E.M., Peterson, G.D. & Gordon, L.J. (2009) Understanding relationships among multiple ecosystem services. *Ecology Letters*, **12**, 1394–1404.

Best, L.B. (1983) Bird Use of Fencerows: Implications of Contemporary Fencerow Management Practices. *Wildlife Society Bulletin (1973-2006)*, **11**, 343–347 CR – Copyright © 1983 Wiley.

Bode, M., Tulloch, A.I.T., Mills, M., Venter, O. & W. Ando, A. (2015) A conservation planning approach to mitigate the impacts of leakage from protected area networks. *Conservation Biology*, **29**, 765–774.

Bottrill, M.C., Joseph, L.N., Carwardine, J., Bode, M., Cook, C., Game, E.T., Grantham, H., Kark, S., Linke, S., McDonald-Madden, E., Pressey, R.L., Walker, S., Wilson, K.A. & Possingham, H.P. (2008) Is conservation triage just smart decision making? *Trends in ecology & evolution*, **23**, 649–54.

Boxall, A.B.A. & Maltby, L. (1995) The characterization and toxicity of sediment contaminated with road runoff. *Water Research*, **29**, 2043–2050.

Boyd, J., Epanchin-Niell, R. & Siikamäki, J. (2015) Conservation Planning: A Review of Return on Investment Analysis. *Review of Environmental Economics and Policy* .

Ten Brink, P., Mazza, L., Badura, M., Kettunen, M. & Withana, S. (2012) *Nature and Its Role in the Transition to a Green Economy*.

Brooks, T.M., Mittermeier, R.A., da Fonseca, G.A.B., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D. & Rodrigues, A.S.L. (2006) Global biodiversity conservation priorities. *Science*, **313**, 58–61.

Burke, D.M. & Nol, E. (1998) Influence of Food Abundance, Nest-Site Habitat, and Forest Fragmentation on Breeding Ovenbirds. , **115**, 96–104.

Burton, G.A., Basu, N., Ellis, B.R., Kapo, K.E., Entrekin, S. & Nadelhoffer, K. (2014) Hydraulic 'Fracking': Are surface water impacts an ecological concern? *Environmental Toxicology and Chemistry*, **33**, 1679–1689.

Cabeza, M. & Moilanen, A. (2001) Design of reserve networks and the persistence of biodiversity. *Trends in Ecology & Evolution*, **16**, 242–248.

Caddy, J.F. (1973) Underwater Observations on Tracks of Dredges and Trawls and Some Effects of Dredging on a Scallop Ground. *Journal of the Fisheries Research Board of Canada*, **30**, 173–180.

Cadenasso, M.L. & Pickett, S.T.A. (2001) Effect of Edge Structure on the Flux of Species into Forest Interiors. *Conservation Biology*, **15**, 91–97.

Carwardine, J., O'Connor, T., Legge, S., Mackey, B., Possingham, H.P. & Martin, T.G. (2012) Prioritizing threat management for biodiversity conservation. *Conservation Letters*, **5**, 196–204.

Cason, T. & Gangadharan, L. (2003) Transactions Costs in Tradable Permit Markets: An Experimental Study of Pollution Market Designs. *Journal of Regulatory Economics*, **23**, 145–165.

Cattaneo, A., Hellerstein, D., Nickerson, C. & Myers, C. (2006) *Balancing the Multiple Objectives of Conservation Programs*.

Clark, C.E., Burnham, A.J., Harto, C.B. & Horner, R.M. (2012) The Technology and Policy of Hydraulic Fracturing and Potential Environmental Impacts of Shale Gas Development. *Environmental Practice*, **14**, 249–261.

Coll, M., Piroddi, C., Albouy, C., Ben Rais Lasram, F., Cheung, W.W.L., Christensen, V., Karpouzi, V.S., Guilhaumon, F., Mouillot, D., Paleczny, M., Palomares, M.L., Steenbeek, J., Trujillo, P., Watson, R. & Pauly, D. (2012) The Mediterranean Sea under siege: spatial overlap between marine biodiversity, cumulative threats and marine reserves. *Global Ecology and Biogeography*, **21**, 465–480.

Cover, M.R., May, C.L., Dietrich, W.E. & Resh, V.H. (2008) Quantitative linkages among sediment supply, streambed fine sediment, and benthic macroinvertebrates in northern California streams. *Journal of the North American Benthological Society*, **27**, 135–149.

Cox, T.M., Lewison, R.L., Zydelis, R., Crowder, L.B., Safina, C. & Read, A.J. (2007) Comparing Effectiveness of Experimental and Implemented Bycatch Reduction Measures: the Ideal and the Real. *Conservation Biology*, **21**, 1155–1164.

Crawley, M.J. (2007) *The R Book*. John Wiley & Sons, Ltd.

Crouse, D.T., Crowder, L.B. & Caswell, H. (1987) A Stage-Based Population Model for Loggerhead Sea Turtles and Implications for Conservation. *Ecology*, **68**, 1412–1423.

Crowder, L.B., Crouse, D.T., Heppell, S.S. & Martin, T.H. (1994) Predicting the Impact of Turtle Excluder Devices on Loggerhead Sea Turtle Populations. *Ecological Applications*, **4**, 437–445.

Crowder, L.B. & Murawski, S.A. (1998) Fisheries Bycatch: Implications for Management. *Fisheries*, **23**, 8–17.

Csuti, B., Polasky, S., Williams, P.H., Pressey, R.L., Camm, J.D., Kershaw, M., Kiester, A.R., Downs, B., Hamilton, R., Huso, M. & Sahr, K. (1997) A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biological Conservation*, **80**, 83–97.

Curry, R.A. & MacNeill, W.S. (2004) Population-level responses to sediment during early life inbrook trout. *Journal of the North American Benthological Society*, **23**, 140–150.

Davies, K.F., Margules, C.R. & Lawrence, J.F. (2000) Which traits of species predict population declines in experiment forest fragments? *Ecology*, **81**, 1450–1461.

Davis, J.B. & Robinson, G.R. (2012) A Geographic Model to Assess and Limit Cumulative Ecological Degradation from Marcellus Shale Exploitation in New York , USA. *Ecology and Society*, **17**.

Detrow, S. (2012) Lycoming County Fracking Spill: Truck Belongs To Spill Response Company. *StateImpact*, 3–5.

Diamond, J.M. (1975) The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biological Conservation*, **7**, 129–146.

Didham, R.K., Hammond, P.M., Lawton, J.H., Eggleton, P. & Stork, N.E. (1998) Beetle species responses to tropical forest fragmentation. *Ecological Monographs*, **68**, 295–323.

Didier, K.A., Glennon, M.J., Novaro, A., Sanderson, E.W., Strindberg, S., Walker, S. & Di Martino, S. (2009) The Landscape Species Approach: spatially-explicit conservation planning applied in the Adirondacks, USA, and San Guillermo-Laguna Brava, Argentina, landscapes. *Oryx*, **43**, 476–487.

Dijkstra, E.W. (1959) A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, **1**, 269–271.

Dobson, A.P., Rodriguez, J.P., Roberts, W.M. & Wilcove, D.S. (1997) Geographic distribution of endangered species in the United States. *Science*, **275**, 550–553.

Drohan, P.J., Brittingham, M., Bishop, J. & Yoder, K. (2012) Early trends in landcover change and forest fragmentation due to shale-gas development in Pennsylvania: a potential outcome for the Northcentral Appalachians. *Environmental Management*, **49**, 1–15.

Engel, S., Pagiola, S. & Wunder, S. (2008) Designing payments for environmental services in theory and practice: An overview of the issues. *Ecological Economics*, **65**, 663–674.

Entrekin, S., Evans-White, M., Johnson, B. & Hagenbuch, E. (2011) Rapid expansion of natural gas development poses a threat to surface waters. *Frontiers in Ecology and the Environment*, **9**, 503–511.

ESRI. (2013) ArcMap.

Evans, J.S. & Kiesecker, J.M. (2014) Shale Gas, Wind and Water: Assessing the Potential Cumulative Impacts of Energy Development on Ecosystem Services within the Marcellus Play. *PLoS ONE*, **9**, e89210.

Fahrig, L. (2003) Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 487–515.

Falconer, K., Dupraz, P. & Whitby, M. (2001) An Investigation of Policy Administrative Costs Using Panel Data for the English Environmentally Sensitive Areas. *Journal of Agricultural Economics*, **52**, 83–103.

Felinks, B., Pardini, R., Dixo, M., Follner, K., Metzger, J.P. & Henle, K. (2011) Effects of species turnover on reserve site selection in a fragmented landscape. *Biodiversity and Conservation*, **20**, 1057–1072.

Ferraro, P.J. (2008) Asymmetric information and contract design for payments for environmental services. *Ecological Economics*, **65**, 810–821.

Forman, R.T.T. & Alexander, L.E. (1998) Roads and their major ecological effects. *Annual Review of Ecology and Systematics*, **29**, 207–231.

Foster, V. & Hahn, R.W. (1995) Designing more efficient markets: lessons from Los Angeles smog control. *Journal of Law and Economics*, 19–48.

Franklin, J. & Forman, R.T. (1987) Creating landscape patterns by forest cutting: Ecological consequences and principles. *Landscape Ecology*, **1**, 5–18.

Freitag, S. & van Jaarsveld, A.S. (1998) Sensitivity of selection procedures for priority conservation areas to survey extent, survey intensity and taxonomic knowledge. *Proceedings of the Royal Society of London B.*, **265**, 1475–1482.

García-Tejero, S., Taboada, Á., Tárrega, R. & Salgado, J.M. (2013) Land use changes and ground dwelling beetle conservation in extensive grazing dehesa systems of north-west Spain. *Biological Conservation*, **161**, 58–66.

Gaston, K.J. (2000) Global patterns in biodiversity. *Nature*, **405**, 220–227.

Gaston, K.J. & Rodrigues, A.S.L. (2003) Reserve selection in regions with poor biological data. *Conservation Biology*, **17**, 188–195.

Gibbs, J.P. (1998) Distribution of woodland amphibians along a forest fragmentation gradient. *Landscape Ecology*, **13**, 263–268.

Gillen, J.L. & Kiviat, E. (2012) Hydraulic Fracturing Threats to Species with Restricted Geographic Ranges in the Eastern United States. *Environmental Practice*, **14**, 320–331.

Gladstone, W. & Davis, J. (2003) Reduced survey intensity and its consequences for marine reserve selection. *Biodiversity & Conservation*, **12**, 1525–1536.

Gómez-Baggethun, E. & Muradian, R. (2015) In markets we trust? Setting the boundaries of Market-Based Instruments in ecosystem services governance. *Ecological Economics*, **In Press**.

Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, **4**, 379–391.

Goulder, L.H. & Parry, I.W.H. (2008) Instrument Choice in Environmental Policy. *Review of Environmental Economics and Policy* , **2** , 152–174.

Goulder, L.H., Parry, I.W.H., III, R.C.W. & Burtraw, D. (1999) The cost-effectiveness of alternative instruments for environmental protection in a second-best setting. *Journal of Public Economics*, **72**, 329–360.

Graham, M. (1955) Effect of trawling on animals of the sea bed. *Deep Sea Res*, **3**, 1–6.

Grantham, H.S., Moilanen, A., Wilson, K.A., Pressey, R.L., Rebelo, T.G. & Possingham, H.P. (2008) Diminishing return on investment for biodiversity data in conservation planning. *Conservation Letters*, **1**, 190–198.

Grantham, H.S., Pressey, R.L., Wells, J.A. & Beattie, A.J. (2010) Effectiveness of Biodiversity Surrogates for Conservation Planning: Different Measures of Effectiveness Generate a Kaleidoscope of Variation. *PLoS ONE*, **5**, e11430.

Grantham, H.S., Wilson, K.A., Moilanen, A., Rebelo, T. & Possingham, H.P. (2009) Delaying conservation actions for improved knowledge: how long should we wait? *Ecology Letters*, **12**, 293–301.

Groves, C.R., Game, E.T., Anderson, M.G., Cross, M., Enquist, C., Ferdaña, Z., Girvetz, E., Gondor, A., Hall, K.R., Higgins, J., Marshall, R., Popper, K., Schill, S. & Shafer, S.L. (2012) Incorporating climate change into systematic conservation planning. *Biodiversity and Conservation*, **21**, 1651–1671.

Hampton, S.E., Strasser, C.A., Tewksbury, J.J., Gram, W.K., Budden, A.E., Batcheller, A.L., Duke, C.S. & Porter, J.H. (2013) Big data and the future of ecology. *Frontiers in Ecology and the Environment*, **11**, 156–162.

Hansen, A.J., Spies, T.A., Swanson, F.J. & Ohmann, J.L. (1991) Conserving Biodiversity in Managed Forests. *BioScience*, **41**, 382–392.

Hartwick, J.M. & Olewiler, N.D. (1998) *The Economics of Natural Resource Use*, 2nd ed (ed D Clinton). Addison-Wesley Educational Publishers, Inc., Reading, Massachusetts, Massachusetts.

Haskell, D.G. (2000) Effects of Forest Roads on Macroinvertebrate Soil Fauna of the Southern Appalachian Mountains. *Conservation Biology*, **14**, 57–63.

Hays, J., Finkel, M.L., Depledge, M., Law, A. & Shonkoff, S.B.C. (2015) Considerations for the development of shale gas in the United Kingdom. *The Science of the total environment*, **512-513C**, 36–42.

Hefley, W. & Seydor, S. (2015) Direct Economic Impact of the Value Chain of a Marcellus Shale Gas Well. *Economics of Unconventional Shale Gas Development SE - 2* Natural Resource Management and Policy. (eds W.E. Hefley & Y. Wang), pp. 15–45. Springer International Publishing.

Henson, A., Williams, D., Dupain, J., Gichohi, H. & Muruthi, P. (2009) The Heartland Conservation Process: enhancing biodiversity conservation and livelihoods through landscape-scale conservation planning in Africa. *Oryx*, **43**, 508–519.

Howarth, R.W., Ingraffea, A. & Engelder, T. (2011) Natural gas: Should fracking stop? *Nature*, **477**, 271–275.

Howe, C., Suich, H., Vira, B. & Mace, G.M. (2014) Creating win-wins from trade-offs? Ecosystem services for human well-being: A meta-analysis of ecosystem service trade-offs and synergies in the real world. *Global Environmental Change*, **28**, 263–275.

Jaeger, J.A.G. (2000) Landscape division, splitting index, and effective mesh size : new measures of landscape fragmentation. *Landscape Ecology*, **15**, 115–130.

Johnson, N., Gagnolet, T., Ralls, R., Zimmerman, E., Eichelberger, B., Tracey, C., Kreitler, G., Orndorff, S., Tomlinson, J., Bearer, S. & Sargent, S. (2010) *Pennsylvania Energy Impacts Assessment*. Harrisburg, PA.

Jones, J.B. (1992) Environmental impact of trawling on the seabed: A review. *New Zealand Journal of Marine and Freshwater Research*, **26**, 59–67.

Jones, I.L., Bull, J.W., Milner-Gulland, E.J., Esipov, A. V & Suttle, K.B. (2014) Quantifying habitat impacts of natural gas infrastructure to facilitate biodiversity offsetting. *Ecology and Evolution*, **4**, 79–90.

Joseph, L.N., Maloney, R.F. & Possingham, H.P. (2009) Optimal allocation of resources among threatened species: a project prioritization protocol. *Conservation Biology*, **23**, 328–338.

Joshi, S., Krishnan, R. & Lave, L. (2001) Estimating the Hidden Costs of Environmental Regulation. *The Accounting Review*, **76**, 171–198.

Kargbo, D.M., Wilhelm, R.G. & Campbell, D.J. (2010) Natural gas plays in the Marcellus Shale: challenges and potential opportunities. *Environmental science & technology*, **44**, 5679–84.

Kassotis, C.D., Tillitt, D.E., Davis, J.W., Hormann, A.M. & Nagel, S.C. (2013) Estrogen and Androgen Receptor Activities of Hydraulic Fracturing Chemicals and Surface and Ground Water in a Drilling-Dense Region. *Endocrinology*, **155**, 897–907.

Kiviat, E. (2013) Risks to biodiversity from hydraulic fracturing for natural gas in the Marcellus and Utica shales. *Annals of the New York Academy of Sciences*, **1286**, 1–14.

Konschnik, K.E. & Boling, M.K. (2014) Shale Gas Development: A Smart Regulation Framework. *Environmental Science & Technology*, **48**, 8404–8416.

Kroeger, T. & Casey, F. (2007) An assessment of market-based approaches to providing ecosystem services on agricultural lands. *Ecological Economics*, **64**, 321–332.

Kukkala, A.S. & Moilanen, A. (2013) Core concepts of spatial prioritisation in systematic conservation planning. *Biological Reviews*, **88**, 443–464.

Lane, P.N.J. & Sheridan, G.J. (2002) Impact of an unsealed forest road stream crossing: water quality and sediment sources. *Hydrological Processes*, **16**, 2599–2612.

Lapeyre, R., Froger, G. & Hrabanski, M. (2015) Biodiversity offsets as market-based instruments for ecosystem services? From discourses to practices. *Ecosystem Services*, **In Press**.

Lawler, J.J., Ruesch, A.S., Olden, J.D. & McRae, B.H. (2013) Projected climate-driven faunal movement routes. *Ecology Letters*, **16**, 1014–1022.

Lehtinen, R.M., Galatowitsch, S.M. & Tester, J.R. (1999) Consequences of habitat loss and fragmentation for wetland amphibian assemblages. *Wetlands*, **19**, 1–12.

Lentini, P.E. & Wintle, B.A. (2015) Spatial conservation priorities are highly sensitive to choice of biodiversity surrogates and species distribution model type. *Ecography*, **In Press**.

Lewison, R.L., Freeman, S.A. & Crowder, L.B. (2004) Quantifying the effects of fisheries on threatened species: the impact of pelagic longlines on loggerhead and leatherback sea turtles. *Ecology Letters*, **7**, 221–231.

Lindenmayer, D.B., Franklin, J.F. & Fischer, J. (2006) General management principles and a checklist of strategies to guide forest biodiversity conservation. *Biological Conservation*, **131**, 433–445.

Lindenmayer, D.B., Margules, C.R. & Botkin, D.B. (2000) Indicators of Biodiversity for Ecologically Sustainable Forest Management. *Conservation Biology*, **14**, 941–950.

Lockie, S. (2013) Market instruments, ecosystem services, and property rights: Assumptions and conditions for sustained social and ecological benefits. *Land Use Policy*, **31**, 90–98.

Lustgarten, A. (2009) Frack Fluid Spill in Dimock Contaminates Stream , Killing Fish. *ProPublica*, 3–4.

Magurran, A.E., Baillie, S.R., Buckland, S.T., Dick, J.M., Elston, D.A., Scott, E.M., Smith, R.I., Somerfield, P.J. & Watt, A.D. (2010) Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in Ecology & Evolution& evolution*, **25**, 574–582.

Margules, C.R., Milkovits, G.A. & Smith, G.T. (1994) Constrasting Effects of Habitat Fragmentation on the Scorpion Cercophonius Squama and an Amphipod. *Ecology*, **75**, 2033–2042.

Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature*, **405**, 243–53.

Martin, L.J., Blossey, B. & Ellis, E. (2012) Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. *Frontiers in Ecology and the Environment*, **10**, 195–201.

Maskell, L.C., Crowe, A., Dunbar, M.J., Emmett, B., Henrys, P., Keith, A.M., Norton, L.R., Scholefield, P., Clark, D.B., Simpson, I.C. & Smart, S.M. (2013) Exploring the ecological constraints to multiple ecosystem service delivery and biodiversity. *Journal of Applied Ecology*, **50**, 561–571.

Mateo, R.G., de la Estrella, M., Felicísimo, Á.M., Muñoz, J. & Guisan, A. (2013) A new spin on a compositionalist predictive modelling framework for conservation planning: a tropical case study in Ecuador. *Biological Conservation*, **160**, 150–161.

Matlack, G.R. (1993) Microenvironment variation within and among forest edge sites in the eastern United States. *Biological Conservation*, **66**, 185–194.

Metcalfe, K., Delavenne, J., Garcia, C., Foveau, A., Dauvin, J.C., Coggan, R., Vaz, S., Harrop, S.R. & Smith, R.J. (2013) Impacts of data quality on the setting of conservation planning targets using the species–area relationship. *Diversity and Distributions*, **19**, 1–13.

Miles, L. & Kapos, V. (2008) Reducing Greenhouse Gas Emissions from Deforestation and Forest Degradation: Global Land-Use Implications. *Science* , **320** , 1454–1455.

Miller, M.A., Jenkins, C.D. & Rai, R.R. (2010) Applying Innovative Production Modeling Techniques to Quantify Fracture Characteristics, Reservoir Properties, and Well Performance in Shale Gas Reservoirs.

Moilanen, A. (2012) Planning impact avoidance and biodiversity offsetting using software for spatial conservation prioritisation. *Wildlife Research*.

Moilanen, A., Anderson, B.J., Eigenbrod, F., Heinemeyer, A., Roy, D.B., Gillings, S., Armsworth, P.R., Gaston, K.J. & Thomas, C.D. (2011) Balancing alternative land uses in conservation prioritization. *Ecological Applications*, **21**, 1419–1426.

Muggeo, V.M.R. (2008) segmented: an R package to fit regression models with broken-line relationships. *R News*, **8**, 20–25.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Naidoo, R., Balmford, A., Costanza, R., Fisher, B., Green, R.E., Lehner, B., Malcolm, T.R. & Ricketts, T.H. (2008) Global mapping of ecosystem services and conservation priorities. *Proceedings of the National Academy of Sciences* , **105** , 9495–9500.

Ni, S.J. & Zhang, J.H. (2007) Variation of chemical properties as affected by soil erosion on hillslopes and terraces. *European Journal of Soil Science*, **58**, 1285–1292.

Olden, J.D., Jackson, D.A. & Peres-Neto, P.R. (2002) Predictive Models of Fish Species Distributions: A Note on Proper Validation and Chance Predictions. *Transactions of the American Fisheries Society*, **131**, 329–336.

Olmstead, S.M., Muehlenbachs, L.A., Shih, J.-S., Chu, Z. & Krupnick, A.J. (2013) Shale gas development impacts on surface water quality in Pennsylvania. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 4962–4967.

De Ornellas, P., Milner-Gulland, E.J. & Nicholson, E. (2011) The impact of data realities on conservation planning. *Biological Conservation*, **144**, 1980–1988.

Owens, R.A. & Myres, M.T. (1973) Effects of agriculture upon populations of native passerine birds of an Alberta fescue grassland. *Canadian Journal of Zoology*, **51**, 697–713.

Ozanne, A., Hogan, T. & Colman, D. (2001) Moral hazard, risk aversion and compliance monitoring in agri-environmental policy. *European Review of Agricultural Economics* , **28** , 329–348.

Pagiola, S. & Platais, G. (2002) Payments for environmental services. *Environment Strategy Notes*, **3**, 1–23.

Pauchard, A. & Alaback, P.B. (2006) Edge type defines alien plant species invasions along Pinus contorta burned, highway and clearcut forest edges. *Forest Ecology and Management*, **223**, 327–335.

Pépino, M., Rodríguez, M.A. & Magnan, P. (2012) Impacts of highway crossings on density of brook charr in streams. *Journal of Applied Ecology*, **49**, 395–403.

Perry, S.L. (2012) Addressing the Societal Costs of Unconventional Oil and Gas Exploration and Production : Cumulative Risks and Uncertainties of Hydrofracking. *Environmental Practice*, **14**, 352–365.

Phalan, B., Onial, M., Balmford, A. & Green, R.E. (2011) Reconciling Food Production and Biodiversity Conservation: Land Sharing and Land Sparing Compared. *Science* , **333** , 1289–1291.

Polasky, S. & Solow, A.R. (2001) The value of information in reserve site selection. *Biodiversity and Conservation*, **10**, 1051–1058.

Pressey, R.L., Cabeza, M., Watts, M.E., Cowling, R.M. & Wilson, K.A. (2007) Conservation planning in a changing world. *Trends in ecology & evolution*, **22**, 583–92.

Pressey, R.L., Johnson, I.R. & Wilson, P.D. (1994) Shades of irreplaceability: towards a measure of the contribution of sites to a reservation goal. *Biodiversity and Conservation*, **3**, 242–262.

Pryce, B., Iachetti, P., Wilhere, G., Ciruna, K., Floberg, J., Crawford, R., Dye, R., Fairbarns, M., Farone, S., Ford, S., Goering, M., Heiner, M., Kittel, G., Lewis, J., Nicolson, D. & Warner, N. (2006) *Okanagan Ecoregional Assessment, Volume 1 - Report*. Victoria, British Columbia.

Qiu, J. & Turner, M.G. (2013) Spatial interactions among ecosystem services in an urbanizing agricultural watershed. *Proceedings of the National Academy of Sciences* , **110** , 12149–12154.

Quinn, G.P. & Keough, M.J. (2002) *Experimental Design and Analysis for Biologists*. Cambridge University Press, New York.

Racicot, A., Babin-Roussel, V., Dauphinais, J.-F., Joly, J.-S., Noël, P. & Lavoie, C. (2014) A Framework to Predict the Impacts of Shale Gas Infrastructures on the Forest Fragmentation of an Agroforest Region. *Environmental management*, 1–11.

Rahm, D. (2011) Regulating hydraulic fracturing in shale gas plays: The case of Texas. *Energy Policy*, **39**, 2974–2981.

Rainho, A. & Palmeirim, J.M. (2013) Prioritizing conservation areas around multispecies bat colonies using spatial modeling. *Animal Conservation*, **16**, 438–448.

Rands, M.R.W., Adams, W.M., Bennun, L., Butchart, S.H.M., Clements, A., Coomes, D., Entwistle, A., Hodge, I., Kapos, V., Scharlemann, J.P.W., Sutherland, W.J. & Vira, B. (2010) Biodiversity Conservation: Challenges Beyond 2010. *Science* , **329** , 1298–1303.

Ranta, P., Blom, T., Niemela, J., Joensuu, E. & Siitonen, M. (1998) The fragmented Atlantic rain forest of Brazil: size, shape and distribution of forest fragments. *Biodiversity & Conservation*, **7**, 385–403.

Raudsepp-Hearne, C., Peterson, G.D. & Bennett, E.M. (2010) Ecosystem service bundles for analyzing tradeoffs in diverse landscapes. *Proceedings of the National Academy of Sciences* , **107** , 5242–5247.

Redford, K.H., Coppolillo, P., Sanderson, E.W., Da Fonseca, G.A.B., Dinerstein, E., Groves, C., Mace, G., Maginnis, S., Mittermeier, R.A., Noss, R., Olson, D., Robinson, J.G., Vedder, A. & Wright, M. (2003) Mapping the conservation landscape. *Conservation Biology*, **17**, 116–131.

Reid, L.M. & Dunne, T. (1984) Sediment production from forest road surfaces. *Water Resources Research*, **20**, 1753–1761.

Richardson, N.D., Gottlieb, M., Krupnick, A. & Wiseman, H. (2013) *The State of State Shale Gas Regulation*. Resources for the Future Washington, DC.

Rico, A., Kindlmann, P. & Sedláček, F. (2007) Barrier effects of roads on movements of small mammals. *Folia Zoologica*, **56**, 1–12.

Robles, M.D., Flather, C.H., Stein, S.M., Nelson, M.D. & Cutko, A. (2008) The geography of private forests that support at-risk species in the conterminous United States. *Frontiers in Ecology and the Environment*, **6**, 301–307.

Rondinini, C., Wilson, K.A., Boitani, L., Grantham, H. & Possingham, H.P. (2006) Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology letters*, **9**, 1136–45.

Ruijs, A., Wossink, A., Kortelainen, M., Alkemade, R. & Schulp, C.J.E. (2013) Trade-off analysis of ecosystem services in Eastern Europe. *Ecosystem Services*, **4**, 82–94.

Salafsky, N., Margoluis, R., Redford, K.H. & Robinson, J.G. (2002) Improving the Practice of Conservation: a Conceptual Framework and Research Agenda for Conservation Science. *Conservation Biology*, **16**, 1469–1479.

Salzman, J. & Ruhl, J.B. (2000) Currencies and the commodification of environmental law. *Stanford Law Review*, 607–694.

Sandel, M.J. (2012) *What Money Can't Buy: The Moral Limits of Markets*. Macmillan.

Sarkar, S., Pressey, R.L., Faith, D.P., Margules, C.R., Fuller, T., Stoms, D.M., Moffett, A., Wilson, K.A., Williams, K.J., Williams, P.H. & Andelman, S. (2006) Biodiversity Conservation Planning Tools: Present Status and Challenges for the Future. *Annual Review of Environment and Resources*, **31**, 123–159.

Saunders, D.A., Hobbs, R.J. & Margules, C.R. (1991) Biological Consequences of Ecosystem A Review Fragmentation. *Conservation Biology*, **5**, 18–32.

Sayer, J., Sunderland, T., Ghazoul, J., Pfund, J.-L., Sheil, D., Meijaard, E., Venter, M., Boedhihartono, A.K., Day, M., Garcia, C., van Oosten, C. & Buck, L.E. (2013) Ten principles for a landscape approach to reconciling agriculture, conservation, and other competing land uses. *Proceedings of the National Academy of Sciences* , **110** , 8349–8356.

Schloss, C.A., Lawler, J.J., Larson, E.R., Papendick, H.L., Case, M.J., Evans, D.M., DeLap, J.H., Langdon, J.G.R., Hall, S.A. & McRae, B.H. (2011) Systematic Conservation Planning in the Face of Climate Change: Bet-Hedging on the Columbia Plateau. *PLoS ONE*, **6**, e28788.

Shepard, D.B., Kuhns, A.R., Dreslik, M.J. & Phillips, C.A. (2008) Roads as barriers to animal movement in fragmented landscapes. *Animal Conservation*, **11**, 288–296.

Simaika, J.P., Samways, M.J., Kipping, J., Suhling, F., Dijkstra, K.D.B., Clausnitzer, V., Boudot, J.P. & Domisch, S. (2013) Continental-scale conservation prioritization of African dragonflies. *Biological Conservation*, **157**, 245–254.

Slonecker, E.T., Milheim, L.E., Roig-Silva, C.M., Malizia, A.R., Marr, D.A. & Fisher, G.B. (2012) *Landscape Consequences of Natural Gas Extraction in Bradford and Washington Counties , Pennsylvania , 2004 – 2010*.

Smith, R.K., Freeman, P.L., Higgins, J.V., Wheaton, K.S., FitzHugh, T.W., Ernstrom, K.J. & Das, A.A. (2002) *Priority Areas for Freshwater Conservation Action: A Biodiversity Assessment of the Southeastern United States*.

Smith, D.R., Snyder, C.D., Hitt, N.P., Young, J.A. & Faulkner, S.P. (2012) Shale Gas Development and Brook Trout: Scaling Best Management Practices to Anticipate Cumulative Effects. *Environmental Practice*, **14**, 366–381.

Sovacool, B.K. (2014) Cornucopia or curse? Reviewing the costs and benefits of shale gas hydraulic fracturing (fracking). *Renewable and Sustainable Energy Reviews*, **37**, 249–264.

Spies, T. & Johnson, K.N. (2003) The Importance of Scale in Assessing the Compatibility of Forest Commodities and Biodiversity. *Compatible Forest Management SE - 8* Managing Forest Ecosystems. (eds R. Monserud, R. Haynes & A. Johnson), pp. 211–235. Springer Netherlands.

Stavins, R.N. (1995) Transaction Costs and Tradeable Permits. *Journal of Environmental Economics and Management*, **29**, 133–148.

Stein, B.A. (2002) *States of the Union: Ranking America's Biodiversity*. Arlington, VA.

Stein, B.A., Kutner, L.S. & Adams, J.S. (eds). (2000) *Precious Heritage: The Status of Biodiversity in the United States.* Oxford University Press, New York.

Sunderland, T., Sunderland-Groves, J., Shanley, P. & Campbell, B. (2009) Bridging the Gap: How Can Information Access and Exchange Between Conservation Biologists and Field Practitioners be Improved for Better Conservation Outcomes? *Biotropica*, **41**, 549–554.

Sutherland, W.J., Adams, W.M., Aronson, R.B., Aveling, R., Blackburn, T.M., Broad, S., Ceballos, G., Côté, I.M., Cowling, R.M., Da Fonseca, G.A.B., Dinerstein, E., Ferraro, P.J., Fleishman, E., Gascon, C., Hunter Jr., M., Hutton, J., Kereiva, P., Kuria, A., MacDonald, D.W., Mackinnon, K., Madgwick, F.J., Mascia, M.B., McNeely, J., Milner-Gulland, E.J., Moon, S., Morley, C.G., Nelson, S., Osborn, D., Pai, M., Parsons, E.C.M., Peck, L.S., Possingham, H., Prior, S. V., Pullin, A.S., Rands, M.R.W., Ranganthan, J., Redford, K.H., Rodriguez, J.P., Seymour, F., Sobel, J., Sodhi, N.S., Stott, A., Vance-Borland, K. & Watkinson, A.R. (2009) One Hundred Questions of Importance to the Conservation of Global Biological Diversity. *Conservation Biology*, **23**, 557–567.

Sutton, N.J. & Armsworth, P.R. (2014) The Grain of Spatially Referenced Economic Cost and Biodiversity Benefit Data and the Effectiveness of a Cost Targeting Strategy. *Conservation Biology*, **28**, 1451–1461.

Tennessee Department of Environment and Conservation. (2014) *Year 2012 303(d) List: Final Version*. Nashville, TN.

Tennessee Wildlife Resources Agency. (2005) *Tennessee's Comprehensive Wildlife Conservation Strategy. TWRA Wildife Technical Report 05-08*. Nashville, TN.

The Nature Conservancy. (2006) *Standards for Ecoregional Assessments and Biodiversity Visions*. Arlington, VA.

Tietenberg, T.H. (1985) *Emissions Trading, an Exercise in Reforming Pollution Policy*. Resources for the Future.

U.S. Energy Information Administration. (2013) *Technically Recoverable Shale Oil and Shale Gas Resources: An Assessment of 137 Shale Formations in 41 Countries Outside the United States*.

United Nations Development Program. (2011) *Towards a Green Economy: Pathways to Sustainable Development and Poverty Eradication.* United Nations Environment Programme.

United States Department of Agriculture Farm Service Agency. (2011) Conservation Reserve Program Sign-Up 41 Environmental Benefits Index (EBI).

Vane-Wright, R.I., Humphries, C.J. & Williams, P.H. (1991) What to protect?—Systematics and the agony of choice. *Biological Conservation*, **55**, 235–254.

Varian, H.R. (2003) *Intermediate Microeconomics*, 6th ed (ed E Parsons). W.W. Norton & Company, New York, NY, NY.

Verity, G.E. & Anderson, D.W. (1990) Soil erosion effects on soil quality and yield. *Canadian Journal of Soil Science*, **70**, 471–484.

Waldron, A., Mooers, A.O., Miller, D.C., Nibbelink, N., Redding, D., Kuhn, T.S., Roberts, J.T. & Gittleman, J.L. (2013) Targeting global conservation funding to limit immediate biodiversity declines. *Proceedings of the National Academy of Sciences* , **110** , 12144–12148.

Walters, C.J. & Hilborn, R. (1978) Ecological optimization and adaptive management. *Annual Review of Ecology and Systematics*, **9**, 157–188.

Walther, B.A., Cotgreave, P., Price, R.D., Gregory, R.D. & Clayton, D.H. (1995) Sampling effort and parasite species richness. *Parasitology Today*, **11**, 306–310.

Watkins, R.Z., Chen, J., Pickens, J. & Brosofske, K.D. (2003) Effects of Forest Roads on Understory Plants in a Managed Hardwood Landscape. *Conservation Biology*, **17**, 411–419.

Watson, J.E.M., Grantham, H.S., Wilson, K.A. & Possingham, H.P. (2011) Systematic Conservation Planning: Past , Present and Future. *Conservation Biogeography* (eds R.J. Ladle & R.J. Whittaker), pp. 136–160. John Wiley & Sons, Ltd., Chichester, UK.

Wenner, C.A. (1983) Species associations and day-night variability of trawl-caught fishes from the inshore sponge-coral habitat, South Atlantic bight. *Fishery Bulletin*, **81**, 537–552.

West Virginia TAGIS Unit. (2014) Oil & Gas permit download, http://tagis.dep.wv.gov/home/OGPermitDownload

Whitacre, J. (2014) Carnegie Museum of Natural History Pennsylvania Unconventional Natural Gas Wells Geodatabase.

Williams, B.K. (2011) Passive and active adaptive management: approaches and an example. *Journal of Environmental Management*, **92**, 1371–1378.

Williams, P., Gibbons, D., Margules, C., Rebelo, A., Humphries, C. & Pressey, R. (1996) A comparison of richness hotspots, rarity hotspots, and complementary areas for conserving diversity of British birds. *Conservation Biology*, **10**, 155–174.

Wilson, C.R. (2011) *Documenting and Protecting Biodiversity on Land Trust Projects: An Introduction and Practical Guide* (ed K Mueller). Land Trust Alliance.

Wilson, K.A., Westphal, M.I., Possingham, H.P. & Elith, J. (2005) Sensitivity of conservation planning to different approaches to using predicted species distribution data. *Biological Conservation*, **122**, 99–112.

Wiseman, H.J. (2014) The Capacity of States to Govern Shale Gas Development Risks. *Environmental Science & Technology*, **48**, 8376–8387.

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A. & Group, N.P.S.D.W. (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763–773.

Zafra-Calvo, N., Cerro, R., Fuller, T., Lobo, J.M., Rodríguez, M.Á. & Sarkar, S. (2010) Prioritizing areas for conservation and vegetation restoration in post-agricultural landscapes: a Biosphere Reserve plan for Bioko, Equatorial Guinea. *Biological Conservation*, **143**, 787–794.

Vita

As both child of the woods and child of the chip, Austin has always searched for ways to apply his loves of math and programming to his passion for saving nature. As an undergraduate at the University of North Carolina at Chapel Hill, Austin majored in environmental sciences and minored in mathematics. His undergraduate research experience includes modeling the effects of biofuel substitution on ground-level ozone formation in Bangkok, Thailand, creating software for the identification of Red-cockaded Woodpecker habitat near military installations in North Carolina, USA, and running and analyzing paleoclimate models. Austin became one of three members of the inaugural Armsworthian Clan at the University of Tennessee at Knoxville in 2010. Starting with his dissertation, Austin worked closely with Tamara Gagnolet and others at The Nature Conservancy to develop software for the spatial planning of shale gas surface infrastructure, software which was used for his third chapter. This concurrent creation of tools for conservation science and practice are indicative of Austin's current and future career goals.