

INSIGHTS FROM A NEW STATISTICAL ANALYSIS OF NATURAL RESOURCE
DAMAGE SETTLEMENTS FROM OIL SPILLS IN THE UNITED STATES

by

Richard W. Dunford, Ph.D., Owner of Environmental Economics Services (EES), 3029 Sylvania
Drive, Raleigh, NC 27607 (RickDunford@EES-LLC.biz)

Greg E. Challenger, M.S., and Stephan Gmur, Ph.D., President and Environmental Scientist,
respectively, with Polaris Applied Sciences, 12525 131st Court NE, Kirkland, WA 98034

Melissa K. Lynes, Ph.D., Industry Economist with the U.S. Energy Information Administration
in the U.S. Department of Energy, 1000 Independence Ave., SW, Washington, DC 20585
(previously, Research Assistant with EES)

Michael A. Dunford, B.A., Contract Associate with PRA Health Sciences, 4130 Park Lake
Avenue, Suite 400, Raleigh, NC 27612 (previously, Research Assistant with EES)

2017-335 ABSTRACT

Major oil spills in the United States often result in some natural resource damages (NRD), which arise from injuries to natural resources and losses of their services. Other things being equal, larger spills lead to larger NRD. Previous research on NRD settlements using multiple-regression analysis identified several factors that predicted variations in damages, other than the number of gallons spilled (Dunford and Lynes 2014). Those factors included the geographic location of the spill, whether threatened and endangered species were injured, whether recreation closures occurred, and whether unvalued compensatory restoration was part of the settlement. The current paper extends the previous research by adding NRD settlements in the past three years, and by exploring a new multivariate statistical method. Specifically, the new method employs a binary division of sample observations to create tree-like regression

models, which are used to determine the correlations among the settlement characteristics. The new approach is non-parametric, which means that it imposes no assumptions about the underlying distribution of the predictor variables. The model identifies multiple combinations of key explanatory variables, providing more insights on observed differences in NRD settlement amounts.

BACKGROUND

The Oil Pollution Act of 1990 allows federal and state government agencies, acting as Trustees for natural resources on behalf of the public, to collect monetary damages from parties responsible for oil spills in the United States for the concomitant natural resource injuries. The monetary damages are usually based on the cost of projects to restore the lost or impaired services resulting from injuries to natural resources.¹ Oversight and monitoring costs for the restoration projects are often included in settlement amounts.

In this paper we develop a statistical model that explains a substantial portion of the variation in NRD settlement amounts for past oil spills. The next section of our paper describes the sources of the NRD settlement data, explains the challenges in determining the NRD amount for oil spill settlements, and presents the NRD settlement data. Then, the following section

¹ In some instances the responsible party implements one or more of the restoration projects rather than give the cost of the project(s) to the Trustees for them to implement. In many of these instances, the cost of these restoration projects is not included in the NRD settlement, which means that the settlement amount underestimates the total cost of the settlement. As explained later, we take this into account in our statistical model.

provides our statistical model. The penultimate section presents the results of our statistical model, while the final section describes future research plans.

NRD SETTLEMENT DATA

We obtained information for our analysis of NRD settlements from a variety of sources. The consent decrees accompanying settlements were our preferred source for settlement amounts. When consent decrees were not available, we used NRD documents such as Damage Assessment and Restoration Plans for settlement amounts. We also used *Federal Register* notices, press releases, and newspaper articles for settlement amounts, when necessary.

We encountered several difficulties in developing our NRD settlements database. The main difficulty was in isolating the NRD amount in “global” settlements, which often included reimbursement for response costs, penalties and fines, and other non-NRD elements. In several settlements, assessment costs were combined with a portion of response costs or other non-NRD costs. Even when assessment costs were not combined with non-NRD costs, the settlement documents often excluded earlier payments of assessment costs. (For example, sometimes the responsible party had paid for some of the Trustees’ assessment costs in an earlier stage of the assessment, so those costs were not part of the settlement amount.) Furthermore, the assessment costs in NRD settlements only reflect the Trustees’ assessment costs. The assessment costs incurred by the responsible party are not included in the settlement amount. Therefore, we excluded assessment costs from our measure of NRD amounts, focusing exclusively on primary restoration costs, compensatory restoration costs, compensable values, and oversight/monitoring costs for restoration projects. Furthermore, we excluded payments for unspecified support of governmental and non-governmental programs (e.g., oil spill prevention programs), whenever possible, under the assumption that these payments were made in lieu of fines or penalties.

We also found that some of the NRD settlements included projects that the responsible party was going to implement for which there was no cost estimate in the settlement. For example, the responsible parties on the 1994 Tampa Bay oil spill purchased some coastal property and developed a mangrove marsh on that property as part of the NRD settlement, but the settlement documents do not include a dollar value for that compensatory restoration project. Therefore, the Tampa Bay settlement amount in our database understates the full cost of that NRD settlement.

Finally, different sources of information on NRD settlements sometimes did not agree on the amount of the settlement or other characteristics of the spill (e.g., the amount of oil spilled). In such instances we assumed that the consent decree or the NRD document from the trustees (if we did not have the consent decree) was the most reliable source of the information in dispute. We also used the lower end of the range provided for the amount of oil spilled.

Our analysis includes every oil spill having an NRD settlement, except for five spills having no estimate of the spill amount and three “anomalous” NRD settlements. For example, we excluded the NRD settlements for both the *Exxon Valdez* (\$0.9 billion paid over 10 years) and the *Deepwater Horizon* (\$8.8 billion paid over 15 years) oil spills, because they are one or two orders of magnitude bigger than the next largest NRD settlement. We also excluded the \$0 NRD settlement for the *Mega Borg* oil spill, which involved almost five million gallons of oil but caused no measurable environmental damages. We excluded the NRD settlements for small oil spills in Washington based on that state's NRD formula and three spills outside Washington for which we had no information on the amount of oil spilled. Finally, we excluded NRD settlements for “chronic” oil spills occurring over decades (e.g., the diluent leaks in the

Guadalupe, California, oil field). In conclusion, our database includes NRD settlements for 91 oil spills ranging from 422 to 3,800,000 gallons.

Table 1 provides summary information on the oil spills and NRD settlements in our database. For the purposes of this paper we have listed the oil spills alphabetically by the name of the spill. The NRD settlements (excluding assessment costs) average \$4.35 million in 2016 dollars after adjusting for inflation, and range from \$0 to \$60.0 million.² The total for all of the NRD settlements in Table 1 is about \$396 million in 2016 dollars.

STATISTICAL MODEL

Variables Included in Our Statistical Model

Our statistical model uses a multiple-regression approach to explain variations in inflation-adjusted NRD settlement amounts (less assessment costs) in millions of dollars for oil spills in the United States (the “dependent” variable) as a function of the magnitude of several “explanatory” variables. In general, the explanatory variables are the characteristics of an oil spill that might affect the magnitude of natural resource damages, including:

- amount of oil spilled,
- type of oil spilled,
- year of the settlement,

² Thirteen oil spills in our database (one-seventh) have \$0 NRD settlements after excluding assessment costs. In all 13 of these oil spills the responsible party implemented the restoration projects required by the settlement, so the Trustees did not receive any payment for natural resource damages other than assessment costs. The impact of these settlements is quantified by the Unvalued Compensatory Restoration variable in our statistical model.

- number of years between the spill and the settlement,
- season in which the spill occurred,
- geographic region in which the spill occurred,
- whether the spill occurred in saltwater or freshwater,
- whether threatened or endangered species were injured by the spill,
- whether the spill closed a recreation area,
- whether unvalued compensatory restoration costs are part of the settlement,
- the number of trustees, and
- whether one or more of the trustees were tribes.

Other factors that might affect the magnitude of NRD settlements may include:

- the duration of the cleanup,
- the geographical extent of the spill,
- the number and type of injured threatened or endangered species,
- the duration of recreation area closures, and
- the availability of similar substitute recreation areas for the closed recreation areas.

Unfortunately, we could not find information on these factors for most of the oil spills in our database. Consequently, we could not include these factors in our statistical analysis. Table 2 describes the explanatory variables that were significant in our statistical model.

Regression Tree Prediction Model to Determine Significant Variable Splits

As noted above, a wide variety of spill characteristics may affect NRD settlement amounts, but the parameters that are significant drivers for the final settlement amount may not be readily apparent. Our multivariate statistical analysis utilizes a binary division of the sample

to create tree-like regression models (Kampichler et al. 2010) to determine the significant parameters that drive settlement amounts. Binary regression trees are a non-parametric technique that are not limited by the requirements of linear regressions regarding normalized distributions, and the initial distribution of the data can be used without having to transform any non-normal data (Lemon, 2003).

We used the binary regression-tree statistical method as an investigation tool to identify significant splits within the compiled dataset of NRD settlements. The binary regression-tree method revealed relationships between multiple variables within the data that might not be highlighted using other generalized multiple-regression models. We chose the binary regression-tree method because it highlights the underlying relationships of the prediction model and identifies splits within the compiled dataset. The continuing development and widespread use of this statistical method within the fields of health sciences, social sciences, statistics, computer science, environmental sciences and other fields speaks to the general acceptance of this method as an exploratory regression tool.

We used the statistical program R (R Core Team, 2016) with the package ‘tree’ (Ripley, 2016) for the regression-tree statistical analysis to create a prediction model of NRD settlement amounts. We used all of the variables listed in Table 2 as predictor variables with the NRD settlement amounts adjusted to 2016 dollars as the response variable. To build a regression tree, we split the sample population into two unique groups based on an identified significant division of the data by choosing one of the predictive factors. Then, we applied this process again, treating each new group as its own unique entity and finding the next set of variables that best divides the input population into two new groups. We repeated this process recursively until a minimum size was reached or a subgroup could no longer be subdivided (Therneau, et al. 2011).

Then, we pruned the initial tree to remove the least important nodes by recursively removing weak splits (Ripley, 2016) to prevent the tree from overfitting the data.

RESULTS

Regression Tree for Predicting NRD Settlement Amounts

The prediction regression tree created using the R library “tree” contained 4 significant splits using the variables Gallons, Region, and T&E Species, as described in Table 3. Figure 1 is a depiction of the regression tree, which created 5 unique groups using the significant variables. The bottom of Figure 1 shows the group size and boxplot for each terminal node group, with the first quartile, median and third quartile shown in the boxplots.

Trends Revealed by the Regression Tree

The significant splits within the dataset are easily seen when examining Figure 1. For example, when there is a threatened or endangered species present within an area impacted by a spill, the NRD settlement amount is 5 times larger than if a threatened or endangered species is not present within the impacted area. This can also be seen with respect to the region within the United States. Those spills that occur on either the Atlantic or Pacific coasts have NRD settlement amounts that are 8 times larger than a comparable spill in the Gulf of Mexico. Table 4 identifies each of the terminal nodes and shows the variable combinations that create each of the unique splits, the predicted NRD settlement amounts, and the range of values within each group.

Prediction Power of the Regression Tree

A previous multiple regression analysis of NRD settlement amounts using an older version of the dataset had an R squared of 0.78 (Dunford and Lynes, 2014), so it explained 78% of the variation in the settlement amounts. The relatively simple nature of the regression-tree

model results in a weaker fit to the NRD settlement amounts with an adjusted R squared of 0.39, as shown in Figure 2. However, the slope and intercept of the regression line were both very significant, and the regression-tree model clearly shows the thresholds at which differences in the circumstances of a spill event influence NRD settlement amounts. For example, the injury of a threatened or endangered species or the geographic location of the spill significantly impact the NRD settlement amount.

FUTURE RESEARCH

The use of the binary regression tree for creating a prediction model is one of many different machine learning algorithms that could provide insights on how different variables influence NRD settlement amounts. In the next phase of our research we are planning to explore other potential algorithms. Additionally, we are going to investigate alternatives for combining the insights gained from the regression-tree model and the ordinary-least-squares regression model from the previous statistical analysis of NRD settlement amounts by Dunford and Lynes (2014). In particular, the previous analysis revealed a highly non-linear relationship between the NRD settlement amounts and the explanatory variables. In contrast, the regression-tree approach identifies differences in the relationship between the NRD settlement amounts and the explanatory variables below versus above threshold levels of key explanatory variables. The threshold levels for the splits may be useful in creating new interaction variables that can be included in the non-linear regression approach. The goal will be to increase the predictive power of the non-linear regression model by incorporating the interaction variables created from the results of the regression-tree model.

There are two primary purposes for predicting NRD settlements from a statistical model. First, Trustees can use such a model to determine the appropriate level of effort for assessing

NRD. In particular, it is inefficient for assessment costs to exceed the NRD amount (e.g., it is inefficient to spend \$2 million in assessment costs to determine that the NRD amount is \$1 million). Trustees could use such a statistical model to estimate a range for NRD shortly after a spill, which would help them in determining the appropriate effort for assessing the NRD. Second, Potentially Responsible Parties (PRPs) for oil spills could use the statistical model shortly after a spill to determine an amount to reserve for an NRD settlement at a later time. As the assessment proceeds the PRPs can refine the initial reserve amount as warranted.

It is not possible for a statistical model to explain all the variation in NRD settlement amounts, which means that a final NRD settlement could fall outside a 90% confidence interval from a statistical model, for example. However, the likelihood of such an outcome is low (i.e., less than 10% in our example). Therefore, a statistical model should be a useful tool for both Trustees and PRPs in predicting NRD amounts shortly after an oil spill occurs.

REFERENCES

- Dunford, R.W. and Lynes, M.K. 2014. Predicting Natural Resource Damages from Oil Spills in the United States. Paper presented at the International Oil Spill Conference, Savannah, GA.
- Kampichler, Christian, et al. 2010. "Classification in conservation biology: a comparison of five machine-learning methods." *Ecological Informatics* 5.6 (2010): 441-450.
- Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. 2003. Classification and regression tree analysis in public health: Methodological review and comparison with logistic regression. *Annals Behavioral Medicine* 26.3:172-81.
- R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Ripley, B. 2016. Classification and Regression Trees. R package version 1.0-37.

<https://CRAN.R-project.org/package=tree>

Therneau T.M, Atkinson E.J. 2011. An introduction to recursive partitioning using the RPART routines. Mayo Foundation, January 11, 2011.

Table 1: Oil Spill Natural Resource Damage Settlements (Excluding Assessment Costs)³

Spill Name	Spill Year	Spill Location	Spill Amount (1,000 gallons)	Product Spilled	Inflation-Adjusted Settlement⁴ (\$1,000)	Settlement Year
Adak	2010	Adak, AK	142	No. 2 diesel fuel	\$0	2013
<i>Amazon Venture</i>	1986	Savannah River, SC	500	No. 6 fuel oil	\$2,515	1987
<i>American Trader</i>	1990	Huntington Beach, CA	400	Crude oil	\$24,167	1999, 1996, 1995
Anacortes	1991	Fidalgo Bay, WA	210	Crude oil	\$683	1998
<i>Anitra</i>	1996	Delaware Bay, NJ	40	Light crude oil	\$1,591	2004
<i>Apex Houston</i>	1986	San Francisco Bay to Long Beach Harbor, CA	25	Crude oil	\$8,703	1994
Arthur Kill	1990	Arthur Kill, NJ	567	Home heating oil	\$16,697	1991
<i>Athos I</i>	2004	Delaware River & Delaware Bay, NJ	263	Crude oil	\$30,027	2010
Avila Beach Tank	1992	Avila Beach, CA	25	Crude oil	\$2,125	1996
<i>B.T. Nautilus</i>	1990	Kill Van Kull, NJ	280	No. 6 fuel oil	\$5,436	1993
Beaver Creek	1999	Warm Springs Indian Reservation, OR	5	Unleaded gasoline	\$372	2006
Bill Williams River National Wildlife Refuge	2000	Bill Williams River National Wildlife Refuge, AZ	6	No. 2 diesel fuel	\$155	2003

³ Citations for the supporting documents for each oil spill settlement are available from the authors.

⁴ Settlement amounts are inflation-adjusted to 2016 dollars.

Spill Name	Spill Year	Spill Location	Spill Amount (1,000 gallons)	Product Spilled	Inflation-Adjusted Settlement ⁴ (\$1,000)	Settlement Year
Blackwater Creek	2002	Blackwater Creek watershed, MN	6	Crude oil	\$0	2009
<i>Bow Mariner</i>	2004	U.S. Exclusive Economic Zone, VA	190	Heavy fuel oil & diesel fuel	\$623	2008
<i>Buzzards Bay Barge 120</i>	2003	Buzzards Bay, MA	22	No. 6 oil	\$6,636	2010
<i>Cape Mohican</i>	1996	San Francisco Bay, CA	96	Fuel oil	\$5,880	1998
Cedar Creek	2001	Cedar Creek, OK	2	Oil	\$0	2005
Chalk Point/ Swanson Creek	2000	Aquasco, MD	140	Oil	\$3,588	2002
Chiltipin Creek	1992	Chiltipin Creek, TX	124	Light crude oil	\$209	1994
Christina River	2006	Wilmington, DE	2	Recycled waste oil	\$215	2009
<i>Cibro Savannah</i>	1990	Arthur Kill, NY & NJ	100	Home heating oil	\$481	1998
<i>Command</i>	1998	San Francisco Bay, CA	3	Crude oil	\$5,728	1999
<i>Cosco Busan</i>	2007	San Fransico Bay, CA	53	IFO 380, heavy fuel oil	\$34,195	2011
Danby Creek	2000	John Heinz National Wildlife Refuge, PA	192	Crude oil discharge	\$1,055	2005
Dixon Bay	1995	Dixon Bay, LA	11	Mixture of crude oil & natural gas	\$0	1996
<i>Dubai Star</i>	2009	San Francisco Bay, CA	0.4	Intermediate fuel oil	\$778	2012
East Walk River	2000	East Walker River, CA	4	#6 fuel oil	\$453	2003
El Segundo	1991	Santa Monica Bay, CA	21	Diesel fuel & naphthalene	\$247	1993
Elm Creek	2008	Elm Creek, IL	158	Crude oil	\$0	2012

Spill Name	Spill Year	Spill Location	Spill Amount (1,000 gallons)	Product Spilled	Inflation-Adjusted Settlement⁴ (\$1,000)	Settlement Year
Enbridge Energy	2010	Kalamazoo River, MI	840	Crude oil	\$59,984	2015
Equinox	1998	Lake Grande Ecille, LA	47	Oil	\$1,102	2005
<i>Ever Reach</i>	2002	Cooper River, SC	13	#6 fuel oil	\$125	2012
Fish Creek	1993	Fish Creek, IN	30	No. 2 diesel fuel	\$3,806	1996
<i>Foss Barge</i>	2003	Puget Sound, WA	5	#6 Bunker fuel	\$293	2008
Genesis Pipeline	1999	Leaf River, MI	336	Crude Oil	\$2,545	2004
<i>Genmar Progress</i>	2007	Guayanilla Bay, PR	45	Diesel fuel	\$2,667	2016
Greenhill Marsh	1992	Timbalier Bay, LA	96	Light crude oil	\$0	1994
High Island	1991	High Island, TX	10	Light crude oil	\$0	1992
<i>Irene</i>	1997	Santa Barbara County, CA	7	Petroleum product	\$3,173	2002
<i>Jahre Spray</i>	1995	Delaware River, PA	59	Crude oil	\$178	1996
<i>Jin Shiang Fa</i>	1993	Rose Atoll National Wildlife Refuge, AS	101	Diesel fuel & lube oil	\$1,880	2003
<i>Julie N.</i>	1996	Fore River & Casco Bay, ME	180	No. 2 fuel oil & No. 6 bunker oil	\$1,383	2000
<i>Kentucky</i>	1994	Delaware River, PA	11	Light crude oil	\$54	1995
Koch Pipeline	1994	San Patricio County, TX	90	Light crude oil	\$0	2008
Kure	1997	Humboldt Bay, CA	5	Intermediate fuel oil	\$3,113	2008
Kuroshima	1997	Summer Bay, AK	39	Bunker C	\$852	2002
Lake Barre	1997	Lake Barre, LA	275	Crude oil	\$0	1999
Lake Salvador	1991	Lake Salvador, LA	2	Light crude oil	\$0	1991

Spill Name	Spill Year	Spill Location	Spill Amount (1,000 gallons)	Product Spilled	Inflation-Adjusted Settlement ⁴ (\$1,000)	Settlement Year
Marcelinas Creek	1998	Marcelinas Creek, TX	40	Crude oil	\$619	2015
Martinez	1988	Suisuan & San Pablo Bay, CA	400	Crude oil	\$19,747	1990
McGrath Lake	1993	McGrath Lake, CA	87	Crude oil	\$1,951	1997
<i>Milos Reefer</i>	1989	St. Matthew Island, AK	250	Diesel & fuel oil	\$659	1993
Monongahela River	1988	Monongahela River, PA	3800	Diesel fuel	\$3,361	1989
Morris J. Berman	1994	San Juan, PR	1045	#6 fuel oil	\$13,108	2000
Neches River	1993	Neches River, TX	88	Light crude oil	\$297	1997
<i>Nestucca</i>	1988	Gray's Harbor, WA	231	No. 6 fuel oil	\$874	1991
<i>New Carissa</i>	1999	Coos Bay, OR	107	Bunker fuel oil & diesel fuel	\$3,948	2004
<i>North Cape</i>	1996	Block Island Sound, RI	828	Home heating oil	\$24,477	2000
North Pass	2002	Plaquemines Parish, LA	13	Crude oil	\$50	2006
Northridge	1994	Santa Clara River, CA	190	Crude oil	\$10,534	1997
Oil and Crooked Creeks	1990-1992	Oil & Crooked Creeks, IN	4	Jet fuel/kerosene	\$415	1996
Perth Amboy	2006	Arthur Kill, NJ	14	Dabo crude oil	\$53	2006
<i>Polar Tanker</i>	2004	Dalco Passage, WA	1	Crude oil	\$532	2010
Posavina	2000	East Boston, MA	60	IFO 380, heavy fuel oil	\$138	2000
<i>Presidente Rivera</i>	1989	Delaware River, PA	200-250	No. 6 fuel oil	\$3,528	1993
Quinnipiac River	1995	Quinnipiac River, CT	5	No. 4 fuel oil	\$61	1996
Reedy River	1996	Reedy River, SC	960	Diesel fuel	\$9,715	1998

Spill Name	Spill Year	Spill Location	Spill Amount (1,000 gallons)	Product Spilled	Inflation-Adjusted Settlement⁴ (\$1,000)	Settlement Year
<i>RTC-380</i>	1992	Long Island Sound, CT & NY	22	No. 2 fuel oil	\$321	1994
San Jacinto River	1994	Channelview, TX	406	Gasoline, crude oil, & fuel oil	\$1,318	2002
Sanborn Pond	2001	Sanborn Pond, ME	6	No. 6 fuel oil	\$144	2007
Santa Clara River	1991	Los Angeles & Ventura Counties, CA	56	Crude oil	\$3,508	2002
<i>Shinoussa</i>	1990	Galveston Bay, TX	694	Catalytic feedstock oil	\$2,110	1994
<i>Silver Creek</i>	2006	Mount Baker-Snoqualmie National Forest, WA	18	Diesel Fuel	\$567	2008
<i>Skaubay/Berge</i>	1995	Gulf of Mexico, TX	38	#6 fuel oil	\$2,241	1999
<i>Star Evviva</i>	1999	Atlantic Ocean, SC	24	#6 oil	\$2,428	2003
<i>Stuyvesant</i>	1999	Humboldt Bay, CA	2	Intermediate fuel oil	\$7,542	2006
Sugarland Run	1993	Potomac River, VA	408	No. 2 diesel fuel	\$4,022	1998
Suisan Marsh	2004	Solano County, CA	124	Diesel fuel	\$1,322	2007
<i>Tampa Bay</i>	1993	Boca Ciega Bay, FL	362	Jet fuel, diesel, gasoline & No. 6 fuel oil	\$4,210	1999
<i>Tenyo Maru</i>	1991	Cape Flattery, WA	475	Intermediate fuel oil, diesel fuel, & fish oil	\$8,291	1994
Tesoro Refinery	1998	Oahu, HI	5	Bunker fuel	\$686	2001
Texmo	2006	Bill Williams River National Wildlife Refuge, AZ	8	Diesel	\$1,398	2007
Turkey Creek	2007	Walker County, TX	276	Jet A fuel	\$0	2009
Umpqua River	2006	Umpqua River, OR	542	Oil	\$53	2008

Spill Name	Spill Year	Spill Location	Spill Amount (1,000 gallons)	Product Spilled	Inflation-Adjusted Settlement⁴ (\$1,000)	Settlement Year
Wabash River	2003	Griffin, IN	2		\$0	2009
Waiiau Marsh	1996	Pearl Harbor, HI	39	No. 6 fuel oil	\$3,359	1999
Westchester	2000	Plaquemines Parish, LA	550	Crude oil	\$0	2003
Whatcom Creek	1999	Bellingham, WA	236	Oil	\$4,468	2004
<i>World Prodigy</i>	1989	Narragansett Bay, RI	290	Home heating oil	\$992	1991
Yellowstone River	2011	Yellowstone River, MT	63	Crude oil	\$11,784	2016
Yoncalla Creek	1993	Yoncalla Creek, OR	6	Diesel fuel	\$336	1995

Table 2: Variables Used in Creating the Regression Tree

Variable Name	Description
Settlement	NRD settlement (2016 dollars)
Gallons	Amount of oil spilled in thousands of gallons
Winter	1 if spill occurred in the winter, and 0 otherwise
T&E Species	1 if a threatened or endangered species was injured in the spill, and 0 otherwise
Unvalued Compensatory Restoration	1 if the settlement included a compensatory restoration project having no stated dollar value, and 0 otherwise
Light Crude	1 if the oil spilled was light crude, and 0 otherwise
Combined Costs	1 if the Natural Resource Damage costs also included assessment costs
LT	1 if the spill occurred in Louisiana or Texas, and 0 otherwise
Season	1: Dec-Feb, 2: Mar-May, 3: Jun- Aug, 4: Sept-Nov
Saltwater	1 if the spill occurred in saltwater, and 0 otherwise
Recreation Closure	1 if the spill closed a recreation area, and 0 otherwise
Elapsed Time	The number of years between the spill and the natural resource damage settlement
Restoration	1 if compensatory restoration performed by the responsible party is part of the settlement, and 0 otherwise

Federal Trustees	The number of federal trustees
State Trustees	The number of state trustees
Tribal Trustees	The number of tribal trustees
Region	Region of the country: Atlantic Coast, Pacific Coast, Gulf of Mexico, and Inland
API Group	The American Petroleum Institute (API) base oil category.

Table 3: Variables for Significant Splits in the Regression Tree

Variable Name	Number of Splits	Significant Splits
T&E Species	1	Yes/No
Gallons	2	< or > 761,000 gallons, < or > 256,500 gallons
Region	1	Gulf of Mexico, or Atlantic/Pacific

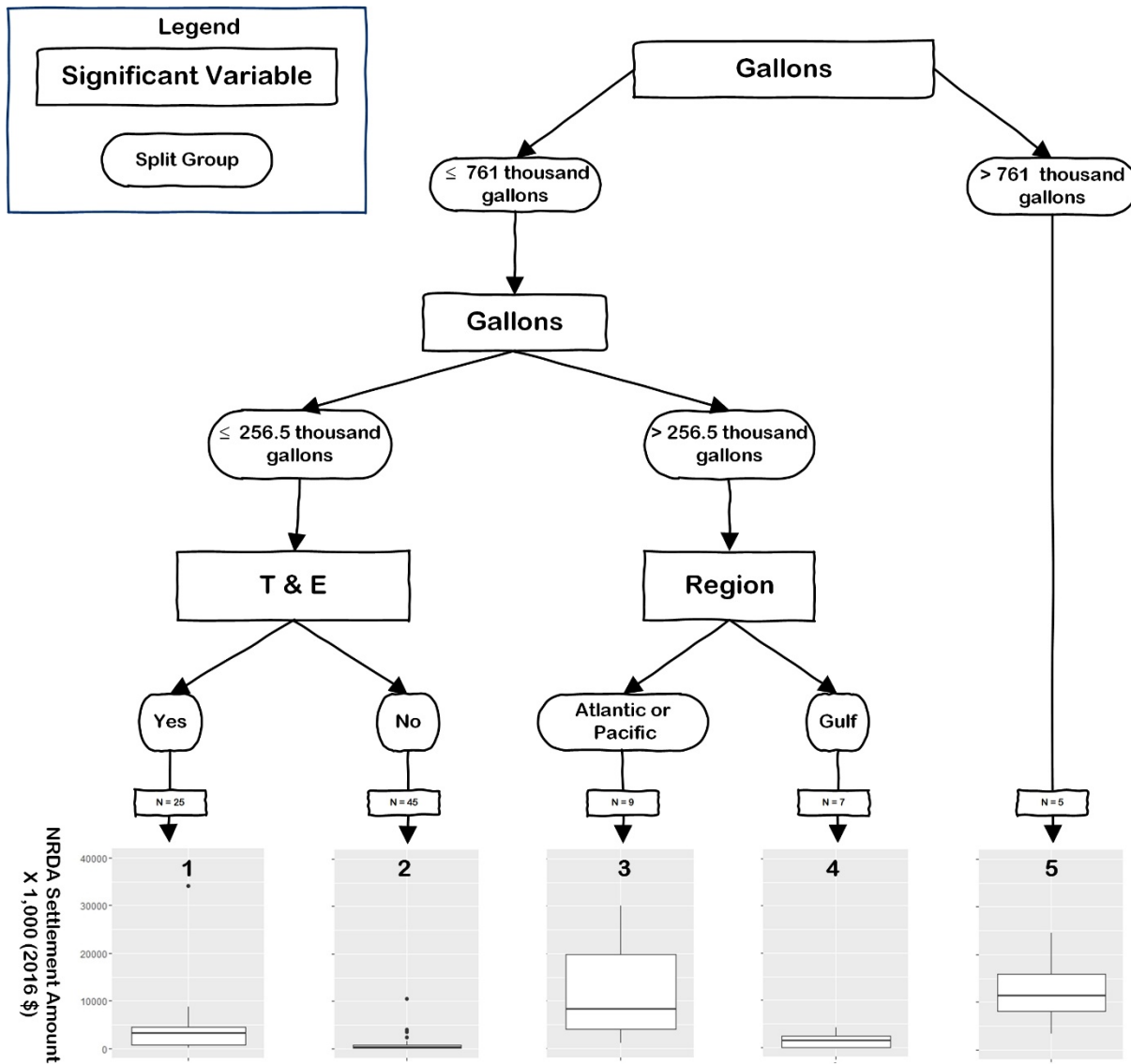


Figure 1: Regression Tree for Predicting NRD Settlement Amounts (in 2016 dollars)

Table 4: Regression-Tree Terminal Nodes and Attributes for the Splits Creating Each Node

Node #	Split Attributes	N	Predicted 2016 NRD Settlement Amount	Predicted 2016 NRD Settlement Range (Min – Max)
1	Spill Size ≤ 761,000 gallons, Spill Size ≤ 256,500 gallons, T&E Species = Yes	25	\$4,242,000	\$125,000 – \$34,200,000
2	Spill Size ≤ 761,000 gallons, Spill Size ≤ 256,500 gallons, T&E Species = No	45	\$847,700	\$0 – 10,530,000
3	Spill Size ≤ 761,000 gallons, Spill Size > 256,500 gallons, Region = Atlantic or Pacific	9	\$12,430,000	\$992,000 – \$30,030,000
4	Spill Size ≤ 761,000 gallons, Spill Size > 256,500 gallons, Region = Gulf	7	\$1,455,000	\$0 – \$4,210,000
5	Spill Size > 761,000 gallons	5	\$22,130,000	\$3,361,000 – \$59,980,000

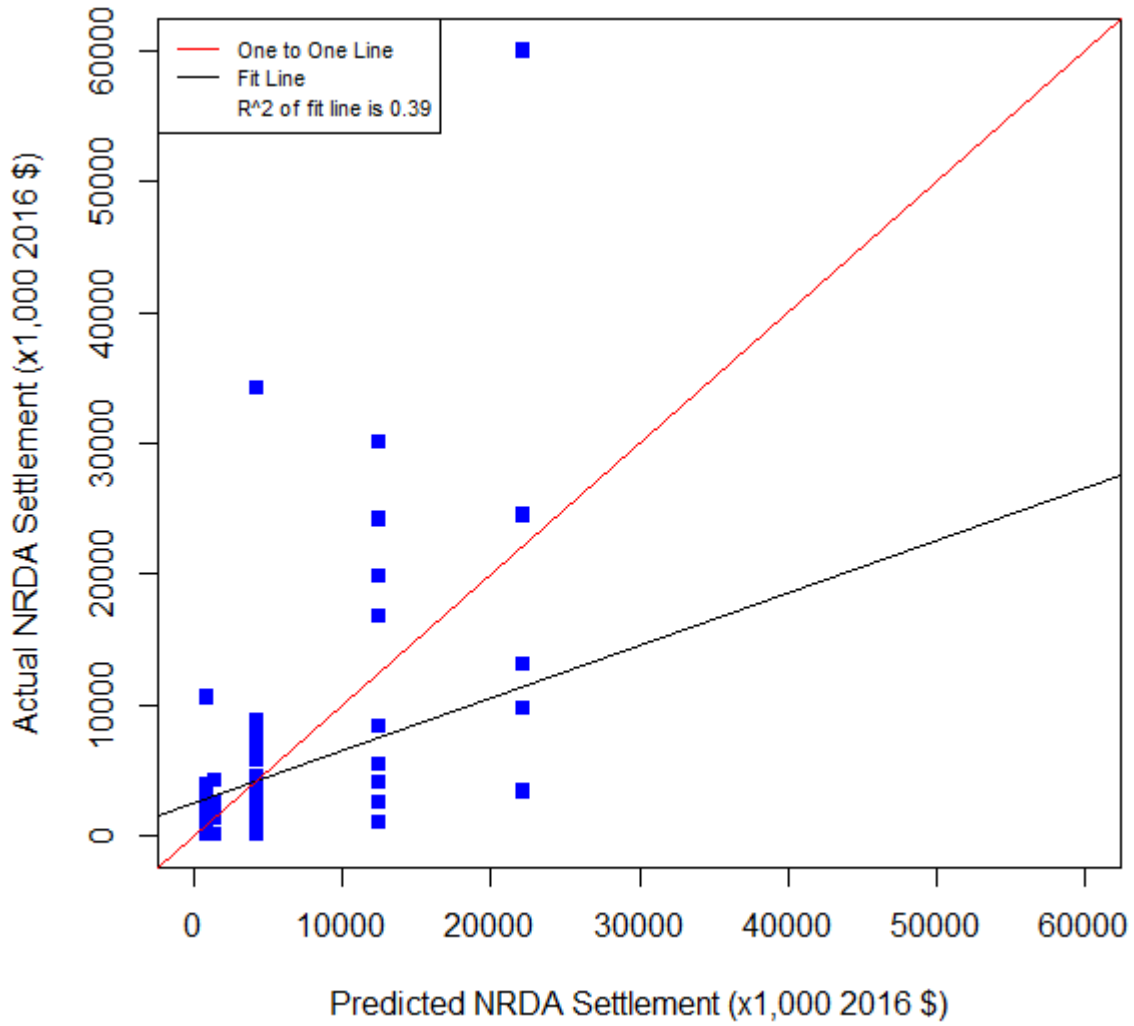


Figure 2: Actual Settlement Amounts Versus Predicted Settlement Amounts