G-dWOLS Analysis of Different Fire Suppression Methods' Effectiveness on Alberta Wildfires

Jiewen Liu
Junior Student at McGill University
jiewen.liu@mail.mcgill.ca

August 21, 2020

Abstract

Based on the idea of precision medicine, the best treatment strategy can be determined on individual level if we can mathematical model it precisely. Under the dynamic treatment regime (DTR) framework, G-dWOLS is a novel and powerful regression based method that can be used in developing such decision rules to give the best treatment, or intervention. Beyond medical cases, we hope to implement G-dWOLS analysis on other scenarios and figure out the optimal choice of intervention. In this work, our goal aims at measuring different forest fire suppression methods' effectiveness on Alberta's wildfires, so as to develop a strategy of fire suppression that is tailored to an individual fire characteristics. There are many environmental factors that can affect the growth of forest fires and knowing the best intervention strategy to minimize the loss is of great importance. G-dWOLS analysis reveals something detailed about how each fire suppression method impacts on an individual fire which inspires us to come up with possible explanations on previous (counter-intuitive) results all while suggesting possible targeted intervention strategies.

1 Introduction

Forest fires refer to uncontrolled fires in rural areas that are rich in burnable vegetation. They represent one of the most common natural disasters in forest ecosystems. Severe fires can vaporize millions of acres of forestland and bring tremendous economic losses. To better combat against it, some institutions have been established and many factors which might be related to forest fires are monitored and recorded. For instance, many studies have been done to show the close relationship between the moisture, temperature, specific forest cover, etc. and the distribution of size of forest fires.[6,10,13,17] These actions could help us better understand the trend of fires and adopt a corresponding fire attack strategy to render the fires under control as soon as possible. Therefore, naturally it would be of particular interest to learn how effective the fire attack method selected is and which one is the best for individual fires.[8,14]

Quantifying the effect of fire attack methods would provide us more insight of it, however, it is difficult to do so since the progress of forest fires involves many factors and randomized trials are unrealistic in this setting.[5] Under the circumstances, observational study is the only choice available to untangle this complex process and incomplete data have posed great challenges on revealing the true power of fire attack method. To causally attribute the difference in outcome to the method, or say treatment, many methodologies have been investigated and proposed to address the issue of confounding, which is introduced by the difference in each treatment group.[2] Ideally, if we have grasped enough background knowledge and all confounders are properly taken into considerations, imbalance within data can be removed and hence quantifying the effect of the treatment through observational data will be possible.

Generalized dynamic weighted ordinary least squares framework, or G-dWOLS, is one of these candidates in eliminating the bias brought by confounding and extracting the causal effect from an observational dataset. It is a regression based method which is built upon on the theory of dynamic treatment regimes (DTRs).[20] In brief, a DTR is probabilistic framework that takes the subject's information as the input and output the most recommended treatment.[4] It can be realized by value-search or regression, and is allowed to be extended to the multi-stage treatment scenarios.[16] For example, value-search approaches like inverse probability weighting (IPW), overlap weighting (OW), outcome weighted learning, targeted maximum likelihood estimating (TMLE) etc. and their augmented forms have been studied in seeking the best treatment.[11,12,18,27,30] For the regression based approach, the treatment effect is quantified and the best treatment is derived through correctly specifying the outcome model.[28] Starting from DTRs, Q-Learning[15,25] and G-estimation[21,22] were first established and an improved, hybrid method – dynamic weighted ordinary least squares (dWOLS) – came later, which combines the double-robustness of G-estimation and simplicity of Q-Learning. Nonetheless, dWOLS didn't, in its original development, go beyond the case of binary treatment. A more generalized version G-dWOLS[23] was later developed, making it suitable for dealing with continuous and multi-level treatment settings.

Back to the context of forest fire, the treatment we are interested in is the fire attack method and the outcome, or response in other words, is related to the forest fire size. Take the response "duration till being-held" as an example. We want to measure the influence of different fire attack method on the time it takes for fire to reach the being-held stage, which means there is no further size increment and the fire is fully under control. In this paper, we investigate a small number of meaningful outcomes and quantify the effect of different treatment under G-dWOLS framework, using data from wildfire record provided by Alberta's Agriculture and Forestry ministry.

2 Methodology and Dataset

2.1 Method Introduction and General Concept

Given the information, a list of variables we observe, of a patient, a business case or an ecosystem, we are always intrigued to know what the optimal strategy is that maximize or minimize the outcome, for example blood pressure, profits or air quality. The optimal strategy may come with sequential decision making at the different stages and the dynamic treatment regimes (Robin, 2004)[20] is proposed to offer such a blueprint of this decision making system. Hence, our goal turns to finding the optimal DTR which orient us toward the best decisions.

To realize the DTR through a regression based method, the outcome model, which is composed of a treatment-free model plus a blip function, needs to be specified first. In simple terms, outcome model takes the outcome of a decision, treatment, as response and the subject's information encoded in a chain of numerical variables, treatment and interactions as predictors. There are many research contributions in this field about constructing regression model properly by observational studies.[7,9,28]

Q-Learning, a classical reinforcement learning algorithm (Watkins, 1989; Sutton and Barto, 1998),[25] has been well adapted to estimating the best decision in a DTR framework. It is a pretty straightforward approach with great simplicity in computation by using ordinary least squares (OLS) to estimate coefficients of each terms directly at each stage. G-estimation (Robins, 2004),[21] a more elaborate framework in surveying the treatment effect, has been well applied on structural nest mean models and enjoy the nice property of double-robustness. However, G-estimation involves solving intimidating estimating equations which has greatly hampered its uptake in practical applications. On the basis of Q-Learning, dynamic weighted ordinary least squares (dWOLS) (Wallace and Moodie, 2015)[29] is devised to circumvent the difficulty in G-estimation while maintaining the double-robustness property. It is achieved by applying dynamic weights at each regression equation. dWOLS targets the case where the treatment is binary, and later a more versatile method – generalized dynamic weighted ordinary least squares (G-dWOLS) (Schulz and Moodie, 2020)[23] – was developed to address the multi-level or continuous treatment case.

2.2 Notation and Setup

Before going to details of G-dWOLS method, let us be clear about the notation and the setup first. To be consistent with the mainstream DTR literatures, we use following notations:

1) Y: The continuous outcome we are interested in. Usually larger Y is preferred but it depends.

2) A: The treatment assigned to each subject. It can be continuous or multi-level.

Variable	Definition
Initial Spread Index (ISI)	The value which reflects the expected rate of fire spread.
Fire Weather Index (FWI)	The value which gives the fire intensity.
Year	The year in which the fire occurred
Month	The month in which the fire occurred.
Fuel Type	The fuel elements that are associated with the behaviour of forest fires.
Period	Time of the day when the initial action against fire is taken (PM or AM).
Response Time	Time interval between the report of fires to action taken.
Count Fire Overlap	The number of fires active at the time of initial assessment.
Treatment	The method used to suppress forest fires.
Detection Agent	The general type of detection agent responsible for discovering the wildfire.
Log-transformed IA size	The log transformed fire sizes at the initial attack.
Eco-Region Name	The type of the ecological region where the fire occurs.

Table 1: Definition of Variables Included in Models.

3) X: The subject's covariates. It contains subject's information and is known before the assignment of treatment.

To implement G-dWOLS method, we have some extra notations:

1) $\pi(A = a|X = x)$: The probability of treatment a is picked conditional on the subject's covariate x. It is also known as "propensity score" (or generalized propensity score, in the case of treatments that are not binary, and represents a distribution or density function, as appropriate to the treatment type).

2) $w(A = a; X = x)$: The weight in each estimating equation for subject with covariate x and treatment a assigned.

The outcome which is associated to X and A through the treatment free model $f(x^\beta; \beta)$ and the blip function $\gamma(x^\psi, a; \psi)$[23] is modelled by the analyst according to:

$$E[Y|X = x, A = a] = f(x^\beta; \beta) + \gamma(x^\psi, a; \psi) = \beta x^\beta + \psi a x^\psi.$$

Both x^β and x^ψ are subsets of the X, the subject's covariates observed. The treatment free model is parameterized with β and the blip function is parameterized with ψ. The treatment free model, just as the name suggests, contains the subject's covariate terms which are believed to be associated with outcome linearly when no treatment or baseline treatment is applied. The blip function includes the treatment's main effect and the interactions between treatment and a possible subset of a subject's covariates in linear form. Usually, x^β and x^ψ are the same sets or x^ψ is a subset of x^β. Here is an illustration. When there are two covariates x_1 and x_2, and a three-level treatment that only interacts with x_2 is studied, the model is:

$$E[Y|X = x, A = a] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \psi_1 a_1 + \psi_2 a_2 + \psi_3 a_1 x_2 + \psi_4 a_2 x_2.$$

The key problem in casual inference is precisely measuring the effect of treatment when only one outcome is available for each subject following the assigned treatment. Since for the same subject, the outcome under another treatment, or the counterfactual outcome, is never observable.[20] To quantify the effect of treatment causally, we need to build up and make some assumptions for our models as most DTR literatures do:

1) Consistency. The subject's outcome observed is the same as its potential outcome under the specific treatment applied. For example, if the patient $id = i$ take the binary treatment $A_i = 1$, then his outcome we observed is $Y_i = Y_i(1)$ instead of $Y_i(0)$ under our factual model.

2) No unmeasured confounders (NUC). We have measured all possible confounders and included them in subject's covariates. When conditional on these, no other factors could impact on the selection of treatment and outcome at the same time, $A \perp Y(A) \mid X$.

3) Positivity. The probability of all possible treatments conditional on the subject's covariates is greater than 0 and less than 1. $0 < P(A = a|X = x) < 1$.

	Air Tanker	HAC1R	HAC1H	HAC1F	Ground-based action	SMD before weighting	SMD after weighting
n	920	495	3017	540	467		
FWI_Dm1 (mean (SD))	13.18 (8.77)	13.33 (9.46)	12.31 (9.24)	13.68 (9.28)	12.57 (9.54)	0.076	0.050
FWI_Dm2 (mean (SD))	11.52 (8.87)	12.67 (11.87)	11.50 (8.99)	12.30 (8.89)	12.15 (10.88)	0.063	0.050
FWI_D0 (mean (SD))	14.15 (8.89)	13.09 (9.06)	12.10 (9.11)	13.01 (8.90)	12.43 (10.40)	0.103	0.037
FWI_D1 (mean (SD))	9.76 (9.23)	9.62 (9.05)	8.70 (9.06)	9.34 (9.57)	8.88 (9.52)	0.062	0.014
FWI_D2 (mean (SD))	9.00 (8.99)	8.98 (9.24)	8.17 (8.99)	9.36 (9.78)	8.21 (9.25)	0.068	0.011
Year (mean (SD))	5.64 (3.34)	5.36 (3.09)	6.26 (3.32)	5.51 (3.23)	5.84 (3.05)	0.133	0.036
Detection_Agent (%)						0.252	0.050
AIR	146 (15.9)	103 (20.8)	738 (24.5)	136 (25.2)	77 (16.5)		
LKT	553 (60.1)	228 (46.1)	1482 (49.1)	248 (45.9)	188 (40.3)		
UNP	221 (24.0)	164 (33.1)	797 (26.4)	156 (28.9)	202 (43.3)		
Eco_Region (%)						0.356	0.052
Clear Hills Upland	169 (18.4)	174 (35.2)	831 (27.5)	107 (19.8)	171 (36.6)		
Mid-Boreal Uplands	680 (73.9)	251 (50.7)	1764 (58.5)	348 (64.4)	194 (41.5)		
Other	71 (7.7)	70 (14.1)	422 (14.0)	85 (15.7)	102 (21.8)		
Fuel_Type (%)						0.234	0.031
C2	729 (79.2)	334 (67.5)	2077 (68.8)	378 (70.0)	263 (56.3)		
M2	84 (9.1)	77 (15.6)	543 (18.0)	82 (15.2)	106 (22.7)		
Other	107 (11.6)	84 (17.0)	397 (13.2)	80 (14.8)	98 (21.0)		
Period = PM (%)	834 (90.7)	439 (88.7)	2646 (87.7)	486 (90.0)	398 (85.2)	0.082	0.023
Month (%)						0.180	0.052
August to October	93 (10.1)	64 (12.9)	583 (19.3)	77 (14.3)	105 (22.5)		
July	427 (46.4)	243 (49.1)	1373 (45.5)	254 (47.0)	201 (43.0)		
June or May	400 (43.5)	188 (38.0)	1061 (35.2)	209 (38.7)	161 (34.5)		
log_Duration_BH (mean (SD))	2.09 (1.49)	1.77 (1.55)	1.89 (1.68)	1.94 (1.63)	2.45 (1.70)	0.192	0.057
log_Response_Time (mean (SD))	6.53 (2.37)	6.58 (2.46)	6.51 (2.56)	6.35 (2.67)	7.14 (2.77)	0.127	0.058
log_Count_Fire_Overlap (mean (SD))	3.93 (1.03)	3.72 (1.00)	3.67 (0.98)	3.89 (0.96)	3.86 (1.09)	0.139	0.007
log_IA_size (mean (SD))	-1.72 (2.32)	-3.22 (1.90)	-2.84 (1.96)	-2.75 (1.97)	-2.36 (2.65)	0.320	0.030

Table 2: Summary of fire's characteristics conditional on the observed treatment.

2.3 Generalized dWOLS

Now, our attention moves to acquiring the precise estimation of coefficients in the outcome model, especially the ones related to treatment terms. The first and the most important step toward double-robust estimation relies on identifying the weight applied OLS estimating equation. Schulz and Moodie (2020)[23] established G-dWOLS method and proved that if weights satisfy the balancing equations, the result from weighted OLS would be double-robust and unbiased.[23] The balancing equations are (1) in continuous treatment cases and (2) in categorical treatment cases:

1) Given the predefined continuous treatment $A \subset R$, for all $a_j, a_k \in A$ *and* $a_j \neq a_k$, $\pi(a_j|x)w(a_j;x) = \pi(a_k|x)w(a_k;x)$ is satisfied.

2) In the discrete setting $A = \{a_1, a_2, ..., a_l\}$ for all $j, k \in A$, $\pi(a_j|x)w(a_j;x) = \pi(a_k|x)w(a_k;x)$ is satisfied.

There are several feasible weights that will suffice. Schulz and Moodie (2020)[23] also discussed a category of applicable weights. Nonetheless, since the best choice for weights is still unknown and the property of consistency as well as double-robustness remains unchanged, overlap weights are used in the regression framework in later analysis and discussion. The forms of overlap weight are (1) in continuous treatment cases and (2) in categorical treatment cases:

$$1)\ w(a_k;x) = \frac{1/\pi(a_k|x)}{\int_A 1/\pi(u|x)du}, \qquad 2)\ w(a_k;x) = \frac{1/\pi(a_k|x)}{\sum_{j=1}^{l} 1/\pi(a_j|x)}.$$

2.4 Data

Forest fire can bring enormous damages in various aspects. Its occurrence and spread are affected by weather, humidity, ground type and so on. The ultimate goal for forest management agency is to deploy resources appropriately and adopt the optimal fire suppression strategy to minimize the loss brought by fires.[19,24] In our regression model, there are a few candidates to be set as the outcome, e.g. how long it takes to reach the stage being-held, the fire size at the stage being-held or extinction divided by its size at the initial attack, whether the fire size increased or not at the stage being-held and so on. Talbot et al. (2019)[2] explored the model with whether the fire size increased or not at the stage being-held as response in detail and compared the averaged effect of different treatment on the population level. In our study, light is shed on the first two outcomes. The

```
Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                 -0.4116051  0.1477125  -2.787 0.005346 **
FWI_Dm1                     -0.0047668  0.0029494  -1.616 0.106108
FWI_Dm2                      0.0060355  0.0026381   2.288 0.022190 *
FWI_D0                       0.0088771  0.0026820   3.310 0.000939 ***
FWI_D1                       0.0081141  0.0026355   3.079 0.002089 **
FWI_D2                       0.0009721  0.0024812   0.392 0.695238
Year                        -0.0094366  0.0059616  -1.583 0.113502
Detection_AgentLKT           0.1242358  0.0524571   2.368 0.017904 *
Detection_AgentUNP          -0.0560937  0.0519862  -1.079 0.280631
Eco_RegionMid-Boreal Uplands  0.2485979  0.0422868   5.879 4.38e-09 ***
Eco_RegionOther              0.1521400  0.0585132   2.600 0.009345 **
Fuel_TypeM2                 -0.1842337  0.0503814  -3.657 0.000258 ***
Fuel_TypeOther               0.0978238  0.0497554   1.966 0.049339 *
TreatmentHAC1R              -0.2976996  0.0562195  -5.295 1.24e-07 ***
TreatmentHAC1H              -0.3904240  0.0560635  -6.964 3.70e-12 ***
TreatmentHAC1F              -0.4454897  0.0561190  -7.938 2.47e-15 ***
TreatmentGround-based action -0.3558919  0.0559226  -6.364 2.13e-10 ***
PeriodPM                    -0.3675217  0.0601983  -6.105 1.10e-09 ***
MonthJuly                   -0.0883081  0.0573209  -1.541 0.123475
MonthJune or May             0.1130137  0.0566168   1.996 0.045971 *
log_Response_Time            0.0080232  0.0076984   1.042 0.297367
log_Count_Fire_Overlap       0.3067648  0.0204457  15.004  < 2e-16 ***
log_IA_size                  0.0198130  0.0093404   2.121 0.033949 *
```

Figure 1: The coefficients of model with main terms only. The outcome is the logarithm of fire size at the stage being-held divided by its size at the initial attack.

dataset used for analysis is the same as the one investigated by Talbot, which is an publicly available data source issued by Alberta's Agriculture and Forestry ministry.

Table 1 provides definition of variables considered in models.

There are four kinds of action, or say treatments, to suppress the fires and we are interested in their effects and whether these should be tailored to fire characteristics. From top to bottom, they are believed to be decreasing in their potency:

1) Air Tanker,
2) HAC1R (Heli-attack crew with helicopter and rappel capability),
3) HAC1H (Heli-attack crew with helicopter but no rappel capability),
4) HAC1F (Fire-attack crew with or without a helicopter and no rappel capability),
5) Other ground-based action.

3 G-dWOLS Analysis of Methods' Effectiveness

3.1 Analysis

To begin with, the original dataset is filtered and preprocessed in the same way performed by Talbot et al. Samples are selected from years between 2003 and 2014, and fires on private land are discarded. In this step, the

Table 3: Characteristics under each predicted best treatment following the model with logarithm of fire size at the stage being-held divided by its size at the initial attack as response.

	Air Tanker	HAC1R	HAC1H	HAC1F	Ground-based action	SMD
n	624	1985	1128	1503	199	
FWI_Dm1 (mean (SD))	11.18 (8.41)	12.94 (8.67)	13.52 (9.04)	10.20 (7.79)	29.42 (9.74)	0.905
FWI_Dm2 (mean (SD))	10.29 (8.41)	11.89 (9.21)	11.66 (9.62)	10.49 (8.60)	24.79 (9.83)	0.651
FWI_D0 (mean (SD))	9.28 (8.21)	14.54 (9.20)	10.08 (7.22)	11.43 (8.23)	28.40 (8.97)	1.001
FWI_D1 (mean (SD))	7.57 (8.03)	9.18 (9.42)	8.46 (8.13)	8.43 (8.43)	20.27 (13.14)	0.486
FWI_D2 (mean (SD))	7.75 (8.18)	8.83 (9.62)	8.37 (8.22)	7.53 (8.10)	15.89 (14.15)	0.318
Year (mean (SD))	8.29 (3.80)	5.61 (3.11)	5.82 (3.13)	5.59 (3.01)	5.76 (3.38)	0.324
Detection_Agent (%)						0.317
AIR	94 (15.1)	347 (17.5)	183 (16.2)	536 (35.7)	40 (20.1)	
LKT	288 (46.2)	1067 (53.8)	698 (61.9)	543 (36.1)	103 (51.8)	
UNP	242 (38.8)	571 (28.8)	247 (21.9)	424 (28.2)	56 (28.1)	
Eco_Region (%)						1.474
Clear Hills Upland	135 (21.6)	448 (22.6)	1 (0.1)	817 (54.4)	51 (25.6)	
Mid-Boreal Uplands	387 (62.0)	1486 (74.9)	1002 (88.8)	353 (23.5)	9 (4.5)	
Other	102 (16.3)	51 (2.6)	125 (11.1)	333 (22.2)	139 (69.8)	
Fuel_Type (%)						0.264
C2	383 (61.4)	1370 (69.0)	896 (79.4)	1013 (67.4)	119 (59.8)	
M2	145 (23.2)	368 (18.5)	97 (8.6)	249 (16.6)	33 (16.6)	
Other	96 (15.4)	247 (12.4)	135 (12.0)	241 (16.0)	47 (23.6)	
Period = PM (%)	49 (7.9)	1974 (99.4)	1122 (99.5)	1471 (97.9)	187 (94.0)	1.795
Month (%)						0.238
August to October	183 (29.3)	310 (15.6)	126 (11.2)	254 (16.9)	49 (24.6)	
July	223 (35.7)	927 (46.7)	557 (49.4)	709 (47.2)	82 (41.2)	
June or May	218 (34.9)	748 (37.7)	445 (39.5)	540 (35.9)	68 (34.2)	
log_Duration_BH (mean (SD))	2.50 (1.61)	1.72 (1.58)	2.12 (1.50)	1.86 (1.72)	2.70 (1.81)	0.315
log_Response_Time (mean (SD))	7.49 (1.71)	7.64 (1.51)	6.89 (1.64)	4.44 (3.22)	6.97 (2.06)	0.596
log_Count_Fire_Overlap (mean (SD))	3.38 (1.00)	3.68 (0.97)	3.92 (0.99)	3.84 (0.98)	4.11 (1.23)	0.327
log_IA_size (mean (SD))	-2.74 (2.00)	-3.72 (1.39)	-0.58 (1.87)	-2.86 (1.96)	-1.45 (2.64)	0.795

sample size reduces from 8591 to 5439. Variables that seems not to be so relevant or are obviously correlated with others are excluded. Some factor variables have categories merged to avoid small cell counts, like for the factor month "June" and "May" are combined to be "June or May". Additionally, for stability's sake, log transformation is done on the variables response time and count fire overlap because they are heavily right skewed. Year is treated as a continuous variable which is an integer ranging from 0 to 11, indicating the calendar years 2003 - 2014.

The expert's choice among these fire suppression methods depends on many environmental factors assessed and known at that time, and not all fire suppression methods would be practical in some fires. Thus, we create table and observe the characteristics of fires classified by their fire suppression methods applied. Then, a multinomial generalized linear model is used to estimate the probability of each fire suppression method. According to the formula given, overlap weight for each fire sample is calculated through the estimated probability. Next, standardized mean differences (SMDs) are used to verify if the imbalance within different treatment group is well reduced.[3] Finally, weighted regression is exercised on the outcome model following the linear specification shown above. Two outcomes investigated here are logarithm of time till being-held stage and logarithm of fire size at the stage being-held divided by its size at the initial attack. A little more attention is paid on variable selection for the model with the latter one as outcome. To better view the regression result, the estimated optimal fire suppression method for each fire sample is computed using the fitted model. A new table is created to summarize the characteristics of fires categorized by their best treatment estimated, to give a sense of which fire characteristics are associated with each particular recommended suppression method. In addition, models with main terms only are studied as well since in this situation, the coefficients of the treatment terms represent the effect at the population level and hence we would be capable of comparing the result with the previous works.

3.2 Result

As testified by the summary table in Table 2, fire samples grouped by their fire suppression method applied are intrinsically very different. Significant imbalance in many fire characteristics is observed. SMD checking indicates the great reduction on all variables brought by overlap weights which is reported in last column in Table 2.

However, no matter which one is set as the model response, air tanker, which is believed to be the most aggressive action, is neither the most effective one in reducing the time till being-held stage nor the increment in fire size ratio. Figure 1 gives the example where the logarithm of fire size ratio is set as the response. All negative coefficients for other treatment terms means air tanker, as the baseline, is the least powerful in controlling fires from spreading. On a contrary, helicopter related method and ground-based action are estimated to be the most ideal choice. Having a look of the model coefficients in Figure 2 and the summary of fire characteristics categorized by the predicted best treatment in Table 3, we observe prominent differences across characteristics of fires stratified by their predicted best treatment, which implies the optimal choice of fire suppression method are closely related to particular traits of an individual fire. Variable selection is done on the model with logarithm of fire size at the stage being-held divided by its size at the initial attack as response. After removing some insignificant interaction terms, the overall trend remains unchanged, but the pattern becomes more clear. The result shows that though air tanker is not the most powerful action, it is an extreme action compared to other fire suppression method since air tanker is set to be the baseline and all interaction terms give either positive or negative effect as shown in Figure 2, and it is to be recommended for certain fire types but certainly not all.

4 Discussion

In this work, we first have provided a review of the regression based DTR framework and G-dWOLS method, and then applied the method to data concerning wildfire suppression in order to estimate an optimal wildfire strategy. The mechanism of occurrence and spread of wildfires is complicated and there is no doubt proper intervention strategy will play an important role in diminishing ecological and financial loss. There are many difficulties in analyzing the power of fire suppression methods, like possible unmeasured confounders, unknown model specification, etc. With the help of G-dWOLS's double-robustness property, we still hope we could capture the effects of fire suppression methods causally and quantify them. Some works have been done in this Alberta's wildfire dataset. Talbot and Pier-Olivier[2,26] both suggest their results were counter-intuitive and there may exist confounding by indication. For example, air tanker, which is considered to be the most aggressive fire suppression method, was not found to be the most effective method in their analyses. Instead, although perceived to be the least powerful approach, ground-based action appears to be the most effective action in some models which take a specific combination of variables or use some fire size related variables as the outcome.

G-dWOLS provides a similar result as exhibited by the previous work. Nevertheless, we don't stop at overall effects, but dig further to look for tailored suppression strategies. Thanks to the resistance to incorrect model specification brought by double-robustness property and the regression based DTR framework, we obtain more insights. In previous work in this dataset, only a single value, the population level averaged effect of each fire suppression method, is given. G-dWOLS analysis provides us the quantified effect of each interaction terms, which indeed tell us more about this story. These terms show that the power of a specific fire suppression method is largely affected by the fire's characteristics at the individual level, such as fuel type, ecological region, logarithm of fire size at initial attack and so on. In best treatment table provided in Figure 3, we also see that when the logarithm of fire size ratio is set as the outcome, each treatment group gives an unique combination of characteristics for which that fire suppression method is predicted to yield the best outcome.

What it is reflected by the G-dWOLS analysis gives us new inspirations on the interpretation on the effectiveness of each method. If there is little or no unmeasured confounding, air tanker might truly be the least effective method on population level. Diving deep into the description of each method and the official Alberta's 2015 wildfire review,[1] we have the information:

1) **Helicopter**: Forest firefighters take the helicopter to fire sites rapidly, which make it possible for them to quickly measure the fire's situation and take action quickly. If well geared up and rappel is available, they

```
Coefficients:
                                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                                    -2.235860   0.260880  -8.570  < 2e-16 ***
FWI_D0                                          0.020474   0.004539   4.511 6.59e-06 ***
FWI_D1                                          0.009133   0.002170   4.209 2.61e-05 ***
Year                                           -0.010078   0.005951  -1.694 0.090392 .
Detection_AgentLKT                              0.129601   0.051946   2.495 0.012628 *
Detection_AgentUNP                             -0.043431   0.051651  -0.841 0.400471
Eco_RegionMid-Boreal Uplands                    0.228232   0.041791   5.461 4.94e-08 ***
Eco_RegionOther                                 0.152478   0.057780   2.639 0.008341 **
Fuel_TypeM2                                    -0.181967   0.049865  -3.649 0.000266 ***
Fuel_TypeOther                                  0.082705   0.049053   1.686 0.091847 .
TreatmentHAC1R                                  2.328445   0.342577   6.797 1.18e-11 ***
TreatmentHAC1H                                  1.693989   0.344039   4.924 8.74e-07 ***
TreatmentHAC1F                                  2.461398   0.343615   7.163 8.94e-13 ***
TreatmentGround-based action                    1.268188   0.348199   3.642 0.000273 ***
PeriodPM                                       -0.779910   0.130037  -5.998 2.13e-09 ***
MonthJuly                                      -0.097282   0.056563  -1.720 0.085513 .
MonthJune or May                                0.111092   0.055970   1.985 0.047212 *
log_Response_Time                               0.082103   0.016972   4.838 1.35e-06 ***
log_Count_Fire_Overlap                          0.667781   0.042104  15.860  < 2e-16 ***
log_IA_size                                    -0.052792   0.019404  -2.721 0.006537 **
FWI_D0:TreatmentHAC1R                          -0.026001   0.006285  -4.137 3.57e-05 ***
FWI_D0:TreatmentHAC1H                          -0.013792   0.006091  -2.264 0.023597 *
FWI_D0:TreatmentHAC1F                          -0.019448   0.006298  -3.088 0.002024 **
FWI_D0:TreatmentGround-based action            -0.005049   0.005997  -0.842 0.399948
TreatmentHAC1R:PeriodPM                         0.365259   0.182572   2.001 0.045482 *
TreatmentHAC1H:PeriodPM                         0.720043   0.178477   4.034 5.55e-05 ***
TreatmentHAC1F:PeriodPM                         0.449124   0.181470   2.475 0.013357 *
TreatmentGround-based action:PeriodPM           0.668383   0.187894   3.557 0.000378 ***
TreatmentHAC1R:log_Response_Time               -0.104035   0.023277  -4.469 8.01e-06 ***
TreatmentHAC1H:log_Response_Time               -0.089251   0.023158  -3.854 0.000118 ***
TreatmentHAC1F:log_Response_Time               -0.085090   0.023349  -3.644 0.000271 ***
TreatmentGround-based action:log_Response_Time -0.083382   0.021109  -3.950 7.91e-05 ***
TreatmentHAC1R:log_Count_Fire_Overlap          -0.414000   0.058959  -7.022 2.46e-12 ***
TreatmentHAC1H:log_Count_Fire_Overlap          -0.441786   0.058572  -7.543 5.38e-14 ***
TreatmentHAC1F:log_Count_Fire_Overlap         -0.592688   0.059072 -10.033  < 2e-16 ***
TreatmentGround-based action:log_Count_Fire_Overlap -0.387069   0.057543  -6.727 1.92e-11 ***
TreatmentHAC1R:log_IA_size                      0.127965   0.027115   4.719 2.43e-06 ***
TreatmentHAC1H:log_IA_size                      0.098549   0.027423   3.594 0.000329 ***
TreatmentHAC1F:log_IA_size                      0.082761   0.027776   2.980 0.002899 **
TreatmentGround-based action:log_IA_size        0.048814   0.026901   1.815 0.069651 .
```

Figure 2: The coefficients of model with interaction terms. The outcome is the logarithm of fire size at the stage being held divided by its size at the initial attack.

can land nearby the fire site and utilize many firefighting tools to deliver a precise attack. Besides that, the helicopters can also carry water and drop retardant if possible.

2) **Air tanker**: Fixed-wing aircrafts carry water tanks which are filled at the ground base or by skimming over the river, lake or reservoirs. Like bombers, they fly over the fire sites and drop water and retardant also sometimes.

3) **Ground-based Action**: It is inclined to be viewed as defensive action. It suits well when limited flying and limited opportunities for aerial ignition (a technique that use planes to remove the fuel) due to high build up index (BUI), a measure of the amount of combustible forest materials, and drought codes can be a problem for aerial attack method and exposure to fire is dangerous. Heavy equipment can be applied in ground-based action and it becomes an excellent option in dry area.

Furthermore, according to the fire review, it worth noticing the resources for helicopter based method (48,747 hours flown, 190 hired crew members, 321 helicopter deployed) are almost 10 times as rich as for air tanker (5,235 hours flown, 35 hired crew members, 35 aircrafts deployed) is. There is also sufficient equipment like pumps, chainsaws, hose, etc. for helicopter based method and ground based action.[1] Unquestionably, the sufficiency in resources contributes greatly to the power of helicopter based method and ground-based action. Restrictions on resources and some other factors like terrain, BUI code, a measure of the amount of combustible forest materials, may be the reason why it comes to big fires, air tanker pales in comparison to other methods less desirable as a "blanket" (un-tailored) strategy. Interestingly, whichever outcome and predictor combination are selected, merely around 15~19% of the best fire suppression method estimated correspond to the ones which were actually executed. The randomness might be explained by the availability. Alberta's fire management agency has a set of monitoring systems and the resources are deployed according to the risk estimated by system. The review also said "Detection is the crucial link between pre-suppression preparedness, where resources are placed on standby and strategically positioned, and suppression activities where the resources are deployed on new wildfires. Information collected during the detection process supports decision making regarding initial attack strategies and tactics aimed at containing wildfires while they are small." It is reasonable to think the availability of resources nearby is carefully considered and the decision on fire suppression method is made upon this in addition to the characteristics of the fire itself.

G-dWOLS analysis reveals some details of this complex mechanism, how the intervention influences the progress of wildfires. The result of regression perhaps cement the argument above. For instance, disadvantageous conditions for air tanker could be represented by the variable the number of additional fires burning and the response time so that the coefficient for interaction between treatment and these two are negative. High response time may indicate there are few resources that can be deployed near to the fire site and too many fires at the same time may magnify the restriction on resources if an air tanker is chosen. Notwithstanding a few plausible explanations, the effectiveness of different fire suppression methods is an open question to which we have provided some additional pieces of evidence. Confounding by indication remains to be a likely cause. To sum up, although it is challenging to evaluate the effect of fire suppression methods using observational data, G-dWOLS, as a regression based method, help demystify what really happens as well as enlighten us to search for clues from other perspectives.

5 Acknowledgement

The research is funded by the McGill University's Arts Research Internship Awards (ARIA) and the Erica Moodie's Discovery Grant from Natural Sciences and Engineering Research Council of Canada (NSERC). It is finished under the professor Erica Moodie's supervision. Thank professor Erica Moodie and Denis Talbot for supports and discussions of background and new findings.

References

[1] Review of Alberta Agriculture and Forestry's Wildfire Management Program and the 2015 Fire Season. MNP LLP Suite 1600, MNP Tower 10235 – 101 Street NW Edmonton, AB T5J 3G1, 2016.

[2] D. S. Arona, D. Thierry, C. Steven, D. Awa, and D. Talbot. Confounding adjustment methods for multi-level treatment comparisons under lack of positivity and unknown model specification, 2019.

[3] P. C. Austin and E. A. Stuart. Moving towards best practice when using inverse probability of treatment weighting (iptw) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in medicine*, 34(28):3661–3679, 2015.

[4] B. Chakraborty and E. E. Moodie. *Statistical methods for dynamic treatment regimes*. Springer, 2013.

[5] C. J. Conway and C. Kirkpatrick. Effect of forest fire suppression on buff-breasted flycatchers. *The Journal of Wildlife Management*, 71(2):445–457, 2007.

[6] W. J. De Groot et al. Interpreting the Canadian forest fire weather index (FWI) system. In *Proc. of the Fourth Central Region Fire Weather Committee Scientific and Technical Seminar*, 1998.

[7] P. Gardoni, A. Der Kiureghian, and K. M. Mosalam. Probabilistic capacity models and fragility estimates for reinforced concrete columns based on experimental observations. *Journal of Engineering Mechanics*, 128(10):1024–1038, 2002.

[8] K. M. Islam and D. L. Martell. Performance of initial attack airtanker systems with interacting bases and variable initial attack ranges. *Canadian Journal of Forest Research*, 28(10):1448–1455, 1998.

[9] J. D. Kang, J. L. Schafer, et al. Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical science*, 22(4):523–539, 2007.

[10] J. J. Keetch and G. M. Byram. *A drought index for forest fire control*, volume 38. US Department of Agriculture, Forest Service, Southeastern Forest Experiment, 1968.

[11] F. Li et al. Propensity score weighting for causal inference with multiple treatments. *The Annals of Applied Statistics*, 13(4):2389–2415, 2019.

[12] F. Li, K. L. Morgan, and A. M. Zaslavsky. Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, 113(521):390–400, 2018.

[13] A. S. López, J. San-Miguel-Ayanz, and R. E. Burgan. Integration of satellite sensor data, fuel type maps and meteorological observations for evaluation of forest fire risk at the pan-European scale. *International Journal of Remote Sensing*, 23(13):2713–2719, 2002.

[14] D. L. Martell, R. J. Drysdale, G. E. Doan, and D. Boychuk. An evaluation of forest fire initial attack resources. *Interfaces*, 14(5):20–32, 1984.

[15] E. E. Moodie, B. Chakraborty, and M. S. Kramer. Q-learning for estimating optimal dynamic treatment rules from observational data. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 40(4):629–645, 2012.

[16] E. E. Moodie, T. S. Richardson, and D. A. Stephens. Demystifying optimal dynamic treatment regimes. *Biometrics*, 63(2):447–455, 2007.

[17] P. J. Murphy. *Methods for evaluating the effects of forest fire management in Alberta*. PhD thesis, University of British Columbia, 1985.

[18] A. I. Naimi, E. E. Moodie, N. Auger, and J. S. Kaufman. Constructing inverse probability weights for continuous exposures: a comparison of methods. *Epidemiology*, pages 292–299, 2014.

[19] J. Podur, D. L. Martell, and K. Knight. Statistical quality control analysis of forest fire activity in canada. *Canadian Journal of Forest Research*, 32(2):195–205, 2002.

[20] J. M. Robins. Optimal structural nested models for optimal sequential decisions. In *Proceedings of the second seattle Symposium in Biostatistics*, pages 189–326. Springer, 2004.

[21] J. M. Robins. *Optimal Structural Nested Models for Optimal Sequential Decisions*, pages 189–326. Springer New York, New York, NY, 2004.

[22] J. M. Robins, D. Blevins, G. Ritter, and M. Wulfsohn. G-estimation of the effect of prophylaxis therapy for Pneumocystis *carinii pneumonia* on the survival of AIDS patients. *Epidemiology*, pages 319–336, 1992.

[23] J. Schulz and E. E. M. Moodie. Doubly robust estimation of optimal dosing strategies. *Journal of the American Statistical Association*, 2020.

[24] S. L. Stephens and L. W. Ruth. Federal forest-fire policy in the United States. *Ecological Applications*, 15(2):532–542, 2005.

[25] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[26] P.-O. Tremblay, T. Duchesne, and S. G. Cumming. Survival analysis and classification methods for forest fire size. *PLoS one*, 13(1):e0189860, 2018.

[27] M. J. Van Der Laan and D. Rubin. Targeted maximum likelihood learning. *The International Journal of Biostatistics*, 2(1), 2006.

[28] M. P. Wallace, E. E. Moodie, and D. A. Stephens. Model assessment in dynamic treatment regimen estimation via double robustness. *Biometrics*, 72(3):855–864, 2016.

[29] M. P. Wallace and E. E. M. Moodie. Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics*, 71(3):636–644, 2015.

[30] Y. Zhao, D. Zeng, A. J. Rush, and M. R. Kosorok. Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107(499):1106–1118, 2012.