Residential Links Under the Weather

Ramakrishna Padmanabhan[1,2,3]   Aaron Schulman[2]   Dave Levin[3]   Neil Spring[3]

[1]CAIDA   [2]UC San Diego   [3]University of Maryland

ABSTRACT

Weather is a leading threat to the stability of our vital infrastructure. Last-mile Internet is no exception. Yet, unlike other vital infrastructure, weather's effect on last-mile Internet outages is not well understood. This work is the first attempt to quantify the effect of weather on residential outages.

Investigating outages in residential networks due to weather is challenging because residential Internet is heterogeneous: there are different media types, different protocols, and different providers, in varying contexts of different local climate and geography. Sensitivity to these different factors leads to narrow categories when estimating how weather affects these different links. To address these issues we perform a large-scale study looking at eight years of active outage measurements that were collected across the bulk of the last mile Internet infrastructure in the United States.

CCS CONCEPTS

• **Networks** → **Network measurement**; **Network reliability**;

KEYWORDS

Internet reliability, Internet outages, weather, ping

1 INTRODUCTION

> "Severe weather is the number one cause of power outages in the United States and costs the economy billions of dollars a year in lost output and wages, spoiled inventory, delayed production, inconvenience and damage to grid infrastructure."

> *– President's Council of Economic Advisers,*
> "Economic Benefits of Increasing Electric
> Grid Reliance to Weather Outages", 2013

The quote above illustrates how weather-related damage to vital infrastructure can cause significant economic harm. Yet, little is known about the impact of weather-induced outages on the most pervasive infrastructure for Internet access: residential last-mile links. U.S. policy [59] requires telcos to report outages to the FCC, but the reporting threshold is large: the outage must last at least 30

minutes and have affected tens of thousands of customers [59]. Such widespread Internet failures have previously received attention from researchers, who have investigated undersea cable cuts [7, 11], natural disasters [15, 25], and backbone router failures [28].

In practice, most weather events are much more localized and not severe enough to generate such a large outage. For decades, this everyday weather has been known to lead to smaller scale outages of telecom infrastructure. For example, early telephone and cable television engineering documents describe how to avoid moisture in wires because it impedes signal propagation [31, 56]. Also, rain attenuates satellite signals above 10 GHz [27]. Further, point-to-point wireless links can experience multipath fading due to objects moving in the wind [10]. In short, residential links are vulnerable to everyday weather because residential equipment and wiring are often installed outdoors: wind can blow trees onto overhead wires, heat can cause equipment to fail, and rain can seep into underground equipment cabinets.

Surprisingly, for these everyday weather conditions, there are no public statistics on the frequency or magnitude of the outages they induce. Such information could help Internet-based companies determine how many customers they lose to nature and help regulators know which conditions and geographic areas deserve their attention. In this work, we provide the first comprehensive study that identifies the correlation between everyday weather and residential Internet last-mile outages. Specifically, we quantify the increase in the number and duration of outages observed during weather, when compared to non-weather periods.

The correlation between weather and outages cannot be determined with a short term study. Weather is *seasonal*: some weather conditions occur only at certain times of the year. Some weather events are also *rare*; and these rare events (e.g., snow in Louisiana) are particularly interesting in their effects. Targeted probing is needed because weather is *localized*: at any time only specific locations are exposed to interesting weather conditions. Broad observation of outages of several links will capture correlated outages of several hosts, such as the work by Heidemann et al. [23, 47], but it will not reveal failures of individual links as may be the case for weather. Although some systems can obtain detailed measurements at residential gateways [50, 57], the limited deployment of these measurement systems make them inadequate for studying at the scale needed to observe many different weather conditions, multiple times, in different geographic areas. Therefore, we performed an eight year longitudinal study with targeted measurements of residential hosts exposed to weather events.

In 2011, we introduced a measurement system for this task called ThunderPing [53], and have run this software for the past eight years. ThunderPing follows forecasts of weather in the U.S. and pings a sample of 100 hosts from each last-mile provider in the area for six hours before, during, and six hours after the forecasted weather event. The focus of the initial paper on ThunderPing was its probing methodology, but it also included a preliminary study

that looked at 66 days of data. Given how limited the dataset was, we could not draw statistically significant conclusions, particularly when looking at rare link types or rural locations. With only one summer, we saw only a small subset of possible conditions.

This paper analyzes 1,811 days of data across eight years, capturing a diversity of seasonal and regional weather conditions. This large dataset enabled us to address these significant limitations of our preliminary study.

There is a challenge with quantifying how weather correlates with outages: outages are relatively uncommon events, and thus every outage is a significant event. This is compounded by the fact that we wish to analyze subsets of our data to focus on, say, particular link types or locations. With so few outages observed compared to the time that hosts are responsive, it is difficult to determine if different weather conditions cause a statistically significant increase in outages. To address this issue, we apply hazard rate estimation from epidemiology to quantify, and establish statistical significance to, the *inflation* in outage probability, even though failures happen infrequently. We detail this approach in Section 3.1, as we believe it to be of general use to the community.

Another challenge is this metric could be artificially inflated by weather conditions coinciding with network state changes such as maintenance [49] or renumbering [44]. Such events would also result in previously responsive IP addresses becoming unresponsive, resulting in what we term a *dropout* event. Dropouts are therefore supersets of failures (including failures due to planned maintenance) and IP address renumbering. We verify that weather does not appear to be positively correlated with peak diurnal dropout periods, and to do recovery time analysis we select dropout durations unlikely to suffer from address renumbering, allowing us to ensure that we are probing the same residential link before and after the outage.

We present a dataset spanning eight years, all weather conditions, and 101 billion responsive pings to 10.8 million addresses throughout the U.S. We apply techniques from epidemiology to attribute statistically significant dropout rates to individual weather conditions. Our key findings span five broad areas of analysis:

Failure rates increase during weather (§4.1): Most weather conditions are accompanied by a statistically significant increase in the probability of failure; surprisingly, this even includes routine weather events such as light rain and snow. Hail, thunderstorms, gale-force winds, heavy rain, and tornadoes are particularly damaging, resulting in increases in failure rates of up to one additional failure per 30 hours.

Link type variations (various): Different link types experience weather in highly varying ways. For example, wireless link types (WISPs and satellite) experience orders of magnitude greater increases in failure rates than their wired counterparts (cable, DSL, fiber) during rainy conditions.

Geographic variations (§4.2): Different geographic regions are affected to varying degrees. For instance, Midwestern U.S. states are more prone to failures in thunderstorms and rain than coastal states. Also, southern states are the most prone to failures in snow.

Continuous variable analysis (§4.3): Most link types have highly nonlinear failure rates with respect to changes in wind speed and precipitation. Wired links experience significantly larger inflations in failure rates when there is more than an inch of precipitation.

On the other hand, the mere presence of precipitation results in statistically significant increases in failure rates for satellite links.

Recovery times (§5): Wired links can experience a two-fold increase in recovery time during some weather conditions. Wireless link types' recovery durations are less strongly correlated with weather and are often shorter during weather conditions.

Our findings have ramifications on how network outage detection and analysis should be performed; limiting measurements to any particular geographic region, link type, or time of year can introduce statistically significant bias. We believe our results also have implications for network administrators and policy-makers; an increased use of wireless links in rural areas (i.e., through the $2 billion Connect America Fund) has resulted in those states having many short failures during weather. In this work, we studied the frequency and severity of weather-induced failures on last-mile Internet; however, our dataset allows the investigation of other issues, including determining the root causes of weather-related outages. To this end, we have made our analyses of the ThunderPing dataset publicly available at: https://thunderping.cs.umd.edu

2 MEASUREMENT METHODOLOGY AND DATASETS

To perform our study, we make use of ThunderPing [53], a measurement infrastructure we developed to actively probe hosts within a weather event. In this section, we describe the features of ThunderPing that are salient to this paper: when, where, whom, and how ThunderPing probes, as well as the datasets it uses in making these decision. We also describe how we determine when there are "dropouts": transitions in responsiveness to pings indicative of a failure. Last, we summarize the dataset ThunderPing has constructed through eight years of active measurement.

2.1 Data collection methodology

The primary goals of ThunderPing are to measure the responsiveness of hosts before, during, and after weather events. Whereas other active probing techniques involve scanning virtually all networks at regular intervals [16, 47], ThunderPing takes a more focused approach, probing hosts in regions where weather events are likely. Here, we explain how ThunderPing performs its probing precisely where and when it is needed.

When and where ThunderPing probes: The U.S. National Weather Service (NWS) provides real-time "severe weather alert" forecasts. Each such forecast includes myriad metadata [53], but of critical importance to this study are predictions of the event's begin time (s), end time (t), and affected regions, specified by FIPS code—which roughly corresponds to a U.S. county. Of course, these are merely *predictions* of future weather events. ThunderPing uses them as a best-effort guide of where and when to probe, and then follows it up with ground-truth data of *observed* weather conditions.

In particular, ThunderPing probes during the interval [$s-6, t+6$]: up to six hours before the predicted start and six hours after the predicted end. This serves two purposes: it insulates against NWS's prediction errors, and it extends data collection to capture dropouts and their recovery times.

Whom ThunderPing probes: In this study, we seek to understand how weather affects different residential providers and link types.

	IPs	Airports	States	Dropout hours	Responsive hours
Cable					
Comcast	3,228,695	486	51	709,386	364,379,801
Charter	876,137	438	47	354,542	137,546,151
Suddenlink	224,182	165	28	202,616	46,251,813
Cox	188,980	286	47	70,284	32,260,861
DSL					
Qwest	874,623	728	51	1,080,377	170,272,132
Centurylink	373,341	245	33	486,161	98,108,287
Verizon DSL	400,023	206	30	226,147	48,083,070
Megapath	157,182	358	44	230,748	75,995,504
Fiber					
Verizon Fios	572,230	159	23	65,468	62,010,415
GVTC	20,779	7	1	10,273	2,518,882
Dickey	5,525	6	3	7,496	1,343,056
WISP					
RISE Bdbd.	66,557	99	24	75,609	22,105,668
Skyriver	4,529	30	6	5,680	2,548,989
Watch Comm.	5,206	11	2	16,540	1,866,487
Satellite					
ViaSat	211,684	775	51	1,272,226	47,371,118
SageNet	1,497	71	33	5,062	1,117,162
All	10,836,469	847	51	12,032,978	2,391,568,235

Table 1: Summary of dataset for large ISPs. "All" includes data from smaller ISPs. (Here, we count D.C. as a state.)

To this end, given a geographic region (FIPS code) where a weather event is predicted to occur, ThunderPing pings a sample of 100 residential hosts from each provider in the affected region.

IP addresses are found for a given location through the MaxMind location-to-address service, using data from the same year. Geolocation databases are poor at locating routers [18], but largely agree on locations of residential addresses (which are all that Thunder-Ping probes), suggesting reasonable correctness [55]. MaxMind self-reports 90% city-level accuracy within 100km in the U.S., further noting that "IP geolocation is more accurate for broadband IP addresses and less accurate for cellular networks" [40]. Even in the event in which residential addresses' geolocation data is incorrect, ThunderPing will overestimate weather's impact only when a host is in an even more severe weather alert than originally predicted. The only scenario in which we have observed this to be an issue is during hurricanes, wherein adverse weather conditions can be extremely widespread. Thus, to avoid this issue, we omit data during any interval of time when there is a hurricane *anywhere* in the U.S.

As before [53], we use reverse DNS (rDNS) records of IP addresses to determine their provider and link type, when clear. When an address's rDNS record contains a cable-only provider like Comcast, Charter, Suddenlink etc., we determine the link type of the address to be cable. Since ThunderPing uses rDNS for determining linktype, some providers like AT&T were not probed by ThunderPing since AT&T typically does not assign reverse DNS names to their residential customers. Providers that use a variety of media types to provide connectivity (typically rural providers) are included under "All" link types with the rest, but are not classified further.

How ThunderPing probes: ThunderPing sends pings to each of these hosts from up to 10 geographically dispersed PlanetLab vantage points every 11 minutes. This interval is inspired by Heidemann et al. [23] and was then also used by Trinocular [47]. When a PlanetLab node fails, ThunderPing is designed to replace it, but in our analyses, if the number of working vantage points drops

below three, we discard observations at that time as untrustworthy. ThunderPing retransmits failed ICMP requests: when a vantage point does not receive a response, it retries with an exponential backoff up to 10 times within the 11 minute probing interval.

ThunderPing relies on active probing, but providers may administratively filter ICMP requests and home routers may decline to respond. We assume that such providers are no more or less vulnerable to weather.

Determining actual weather conditions: Recall that weather alerts merely *predict* weather conditions; it is of course possible that other weather materializes. To analyze our data, we use publicly available NWS datasets of *observed* weather. The NWS operates a network of 900 automated "ASOS" weather stations, typically located at airports. The NWS weather stations record hourly observations of 24 weather variables in METAR format. The NWS makes these observations available online [41].

There are two types of weather information: categories that account for the common precipitation types (e.g., thunderstorm, hail, snow) and continuous variables (e.g., wind speed, precipitation quantity). In this paper, we analyze both.

2.2 Dropouts

A "dropout" occurs when the IP address representing a residential link that is responsive to pings from at least three vantage points transitions to being unresponsive from all of the vantage points for 11 minutes. This definition permits some vantage points to be faulty or congested; requiring three is meant to avoid classifying in-network faults as residential link faults, and is based on prior work [47]. A dropout can occur within a "responsive address hour," a continuous observation of an IP address in known weather conditions where the address is responsive to pings at the beginning of the hour. A responsive hour may or may not include a dropout, and the ratio of dropouts to responsive hours is our measure of outage likelihood. Two addresses observed in the same hour or a single address observed for two hours are equivalent. The "hour" duration aligns with the typical weather report interval. We consider at most one dropout per hour per address. An address that alternates between responsive and unresponsive will be counted as a single dropout per hour. This treats the potentially multiple dropouts as part of a single outage event, similar to how multiple lost packets in TCP are part of the same congestion event.

Observing a dropout is a sign that a residential link may (or may not) have experienced an outage, but *not all dropouts are necessarily outages*. Dropouts can also occur if the device re-attaches to the network with a new address after only momentary disconnection, typically through re-association of a PPP session for a DSL modem, but potentially through administrative renumbering of prefixes. In Sections 3.3 and 5, we describe how we account for these activities.

We annotate each responsive address hour for an address with the corresponding weather information associated with the geographically closest weather station to that address. Doing so allows us to find the number of responsive address hours and dropout address hours in specific weather conditions.

2.3 Dataset summary

ThunderPing has been running for eight years, and has collected 101 billion responsive pings to 10.8 million residential addresses. This dataset comprises observations from Apr 2011 to Dec 2018, though only 1,811 days included sufficiently many operating vantage points to classify a responsive address hour.

We show per-ISP highlights in Table 1. We observe major providers such as Comcast, Qwest, and ViaSat in all 50 states (and DC). Of the 2.39B responsive address hours from Table 1, 216M (9%) were hours where responsive addresses experienced rain, 91M (4%) snow, and 27M (1%) thunderstorm.

Hurricanes manifest as a combination of weather features and are so pronounced that their contribution to thunderstorm or rain outages would be disproportionate. We thus omit them from categorical weather classification (e.g., Figure 2). However, we consider data from Hurricane events when studying continuous variables (e.g., inches of rain and wind speed) where extremes are clearly distinguishable. Eleven hurricanes made U.S. landfall during our measurement. Collectively, these hurricane times account for less than 3% of responsive address hours and 4% of dropout hours.

3 QUANTIFYING WEATHER DROPOUTS

Our overarching goal is to quantify weather's impact on last-mile link failures. The challenge is to do so in a way that (1) permits a measure of statistical significance and (2) controls for other, non-weather-related causes of dropouts. In this section, we describe the challenges in doing this kind of analysis, and the methodology we apply to solve them.

3.1 Obtaining statistical significance

It is relatively straightforward to derive tight confidence intervals when there is a large amount of data. Although our dataset includes billions of pings, the events we consider in this paper—dropouts—occur exceedingly rarely. On average, we observed a link dropping out only once every 3–35 days that we were actively pinging and receiving responses from the link's address, depending on the link type. The inverse of the average dropout probability per link type—including the average across all link types—is as follows:

Link type:	Fiber	Cable	DSL	WISP	Sat	All
Days b/w dropouts:	35	22	14	11	3	8

Rare events complicate statistical significance: a single additional occurrence could, if not careful, drastically alter the results.

Fortunately, there is a well-established set of techniques from the field of epidemiology that permit statistical significance over rare events. Epidemiology—the study of the occurrence and determinants of disease—faces similar challenges when analyzing mortality: deaths ("failure") are rare, and subjects ("hosts") can be exposed to their disease ("weather condition") for different amounts of time until the time of death ("dropout"). Here, we describe the techniques we borrow from biostatistics [60] to address these concerns. Throughout our study, we will consider different "bins" of subjects: each "bin" is a group of addresses with a specific link type, geographic region, or combination thereof, experiencing a weather condition. For example, WISP addresses in the U.S. experiencing

gale-force winds is one possible bin, cable addresses in Florida experiencing thunderstorms is another, and so on.

Like in epidemiological studies, we focus on estimating the *hazard rate* (sometimes referred to as the instantaneous death rate). In essence, what a hazard rate gives us is the expected number of deaths per unit time. More concretely, for a given hazard rate λ, the probability of death over a short duration of time t is $\lambda \cdot t$.

When presenting our results, we use an hour as the short duration t. This allows us to align host availability with corresponding weather information (the NWS reports weather data every hour). Moreover, unlike with epidemiological studies, hosts can be revived after death—and can even die multiple times in an hour. To avoid biasing our results towards hosts that repeatedly fail and recover over a short period of time, recall from Section 2.2 that we treat multiple dropouts within a given hour as a single "dropout-hour". As a result, in our analysis, the hazard rate captures the hourly probability that a host drops out (one or more times).

The first challenge in estimating hazard rates is that different hosts in a bin may be observed over different periods of time: in our study, hosts that remain responsive can naturally be observed for longer periods of time than those that drop out. We track the number of hours O_i that we observe each address $i = 1, \ldots, n$ to be responsive to pings (and therefore had the opportunity to dropout), and we also count the number of hours where the addresses experienced at least one dropout, D. An *unbiased* estimate of the hazard rate $\hat{\lambda}$ for that bin can be obtained as follows [60, Chapter 15.4]:

$$\hat{\lambda} = \frac{D}{\sum_{i=1}^{n} O_i} \quad (1)$$

We exclude any bin if it lacks enough observed hours to permit computing confidence intervals. We adhere to the following rule of thumb [60, Chapter 6]: we accept a bin with n samples (responsive hours) and estimated hazard rate $\hat{\lambda}$ only if:

$$n \geq 40 \text{ and } n\hat{\lambda}(1 - \hat{\lambda}) \geq 10 \quad (2)$$

Thus, we require more samples to obtain statistical significance for very small (or very large) values of $\hat{\lambda}$; for example, to estimate $\hat{\lambda} = 0.00139$ (one dropout-hour per month), we would require 7205 samples. When the conditions in Eq. (2) do not hold, we do not report the results for the bin. We had enough samples to report results with statistical significance for almost all the bins analyzed in the paper; the exceptions were instances of particularly uncommon weather in an area, such as snow in Florida.

When the conditions in Eq. (2) hold, we can calculate 95% confidence intervals over the estimated hazard rate as follows [60, Chapter 6.3]:

$$\hat{\lambda} \pm 1.96 \cdot \sqrt{\frac{\hat{\lambda}(1 - \hat{\lambda})}{n}} \quad (3)$$

The above calculations yield the hazard rate along with its confidence intervals[1]; what remains is to *compare* two hazard rates, for instance, the overall hazard rate for a given link type and the hazard rate for that link type specifically in the presence of snow. Two estimated hazard rates $\widehat{\lambda_1}$ and $\widehat{\lambda_2}$ can be compared by simply subtracting them [60]. Fortunately, with sufficiently many samples,

---

[1]The constants in the equations above are recommendations from [60, Chapter 6]

the confidence intervals over the difference of two hazard rates is given by the addition of the confidence intervals over the original hazard rates.[2]

To summarize: throughout this paper, we bin addresses experiencing a weather event by link type, geography, or both, compute hazard rates for these bins using Eq. (1), discard any bins that do not satisfy Eq. (2), and compute confidence intervals using Eq. (3).

3.2 Controlling for other dropout causes

The mere fact that a dropout occurs during a weather event is not sufficient evidence to conclude that it was caused by (or even correlated with) the weather event itself. Dropouts can also be caused by completely weather-independent events—such as regularly scheduled network maintenance periods—or partially weather-dependent events—such as a car accident leading to an above-ground wire being knocked down. Our goal is to completely filter out the first and to filter out the non-weather related instances of the second.

To control for a "baseline" probability of dropouts, we measure the *additive difference* in dropout probabilities between times of inclement weather and times of no inclement weather. More concretely, we consider a set of 11 different kinds of weather events, listed in Figure 2, including tornado, thunderstorm, various severities of rain and snow, and so on. Any measurements taken at a time and location that is *not* experiencing these weather events is included towards our estimate of the baseline dropout probability. Note that some weather events may have after-effects: for instance, a "snow" event in the weather data from the NWS spans only the time when snow is actively falling, not the potentially problematic time when snow is on the ground. To account for such after-effects, we also exclude the 24 hours *after* any of these 11 weather types from our computation of the baseline.

We calculated a single baseline across the 8 years for each bin. This simple definition of "baseline" dropout probability represents the dropouts not due to weather for a bin. By considering 8 years of observations, we were able to (a) observe sample hours of weather conditions affecting various regions/link types on several occasions and (b) avoid overfitting the baseline even when we were analyzing per-state behavior and per-link-type behavior within each state. Nearly 60% of all of ThunderPing's sample hours, across all link types and geographical areas, are included in the baseline.

We experimented with other approaches to calculate the baseline probability of dropouts and found consistent results. For example, we excluded other durations (4, 12, 48 hours) in the aftermath of the 11 weather types in Figure 2 from our computation of the baseline; the results were similar. We also excluded hours with extreme temperatures from the baseline and found consistent results.

For a given bin, given the observed baseline dropout probability ($\lambda_{\text{baseline}}$) and the observed dropout probability during weather type w (λ_w), we attribute the weather-induced dropout probability to the difference between the two: $\lambda_w - \lambda_{\text{baseline}}$. We refer to this difference as the *inflation* of dropout probability (note that this value can be negative if a weather condition correlates with a decrease in dropout probability).

Controlling for baseline dropouts in this manner allows us to explore the *relative* impacts that weather has, but it does not allow us to attribute a specific cause. For instance, in Section 4.3, we show that precipitation and dropout inflation are correlated, but we cannot infer whether that is caused by water seeping into above-ground wires, increased car accidents, and so on. Nonetheless, we believe this to be a useful metric for many applications. Whatever the specific cause, understanding what link types are more resilient to different weather conditions in different regions of the country can inform consumers deciding which ISP to use, service providers anticipating user demand, or regulatory bodies seeking to ensure nationwide availability.

3.3 Validating a weather baseline

Our notion of a weather baseline applies if and only if there are no causes of dropouts that occur more often with weather but are not themselves caused by the weather event. For example, if an ISP happened to perform address reassignment at a particular time of day when rain was more likely to occur, then it would appear that rain caused more dropouts. To rule this out, we verify that weather events do not positively correlate with common dropout periods.
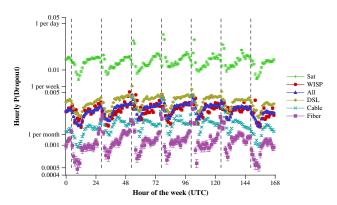
The first question we must answer is: Are there any hours of the week that have a significantly higher probability of dropout than other hours of the week? To this end, we evaluate the probability of dropouts in each hour of the week in the following manner: for each hour of the week, we counted the number of dropouts (recall that dropouts only occur at most once per hour per link) across all links observed during that hour, then we divided that by the number of hours in which the link was responsive. We did this for each link type separately, as some link types may be more likely to be renumbered [44]. Also, some link types may require maintenance more often than others.
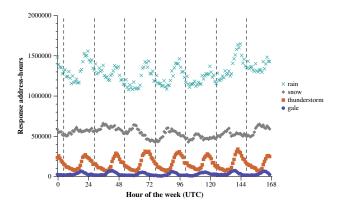
The results are shown in Figure 1(a). Hour 0 on the *x*-axis is Sunday midnight, UTC. As expected, the hourly probability of dropouts significantly varies in a diurnal pattern over the course of each week. Prior work suggests that the increased dropout probability during weeknights could be attributed to ISPs performing administrative maintenance [6, 12]. Interestingly, while each link type exhibits a markedly different dropout probability, all links experience their highest probability of dropouts in the evening, peak near midnight Eastern Standard Time (indicated with vertical dotted black lines), and decrease significantly until the early hours of the morning.

Given that we observe a diurnal variation in hourly likelihood of dropouts, and the fact that weather conditions also have a known diurnal pattern of occurrence [61], the next question we must answer is: Does hourly weather occurrence positively correlate with dropout probability? To answer this question, we count the total number of responsive hours that we observed in each hour of the week for each weather condition.

The results are shown in Figure 1(b). As expected, most weather conditions, possibly except for snow, have a diurnal pattern in their occurrence. Fortunately, none of the weather conditions have a positive correlation with the hourly probability of dropouts. This validates our use of a baseline to measure the relative impact of various weather conditions.

---

[2]This follows from the fact that var($\lambda_1 - \lambda_2$) is approximately var(λ_1) + var(λ_2) when Eq. (2) holds.

(a) Dropout probability has significant diurnal variation.

(b) Weather types vary in their most prominent hours.

Figure 1: Weather conditions are not correlated with what appear to be administrative dropouts.

Baseline probability of dropout depends on link type. The investigation into probability of dropout for each link type also provides additional justification for the selection of a metric that is based on the *increase* in failure probability due to weather. The dropout probability is significantly different for each link type, with Fiber being the lowest and Satellite being the highest (Figure 1(a)). With this metric, the baseline failure rate will be removed from all link types; including the diurnal variations in dropout probability.

4 WEATHER ANALYSIS

In this section, we analyze our collected data to understand how weather conditions affect dropouts.

4.1 Relative dropout rates

First, we analyze the relative rate of dropouts under various link types and weather conditions, after omitting all hurricane periods. We use categorical data from weather records (such as "thunderstorm present"), to assign weather conditions for each hour. When multiple weather condition occur in an hour, we count the hour under all the conditions. We experimented with assigning the most severe weather condition to each hour and obtained similar results.

The top of Figure 2 shows the number of responsive address-hours for which we measured the various link- and weather-types. Although there is a wide range in their absolute values (note the log-scale of the y-axis), the overall shape of the histograms remains mostly consistent across the different link types. This reflects the fact that, in their deployment throughout the US, different link types are exposed to very similar conditions, with one minor exception: we did not measure any fiber, WISP, or satellite links during tornadoes. The top figure also explains the larger range of confidence intervals of the bottom figure for severe weather types (such as tornadoes and hail) that observed fewer address-hours.

The bottom of Figure 2 shows the difference in the dropout rate during the presence of a weather condition and the baseline failure rate. For example, the "Thunderstorm" bar for fiber shows that fiber addresses experienced an inflated hourly dropout probability of 0.003496 (+/- 0.000158). We obtained this statistic by observing 702,639 sample hours from fiber addresses in thunderstorm, of which 3,195 were dropout hours (there were 44,266,075 hours in baseline of which 46,538 were dropout hours). In the bottom of the

figure, we also include confidence intervals on all bars; they are tight on almost all values, except tornadoes and hail.

Interpreting inflation in hourly P(Dropout). A value of zero signifies no observed difference with or without a particular weather condition; positive values indicate increased probability of dropout during that weather condition; and negative values indicate *fewer* failures during that weather condition. Further, the inverse of the inflation in hourly P(Dropout) yields the number of address-hours in a certain weather condition that will likely result in an additional failure. Thus, when fiber addresses receive 286 address-hours (or 12 days[3]) of thunderstorm, they will experience one dropout-hour more than they would have in baseline. Figure 2 shows on the y-axis the values that correspond to an additional failure once every month (0.001389) and once every week (0.005952).

We make three key observations from Figure 2. First, most weather conditions exhibit higher dropout probabilities across *all* of the link types we measured, although the extent of the inflation varies. Hail, thunderstorms, gale-force winds, heavy rain, and (for the link types that experienced it) tornadoes are particularly damaging; all yield a statistically significant increase in dropout probabilities of more than 0.001389 (one additional failure per month).

Second, for each given link type, heavier rates of precipitation (both rain and snow) yield higher probabilities of dropout. We analyze precipitation further in Section 4.3. Interestingly, the probability of dropouts is greater during thunderstorms than during heavy rain for all link types. This indicates that the causes of failures during thunderstorms extend beyond the rainfall, perhaps to increased wind or power outages.

Finally, the inflation in dropout probabilities of wired link types (cable, DSL, and fiber) are similar to one another, as are those of wireless link types (WISP and satellite), but wired and wireless link types are different from one another. For example, light rain has almost no discernible difference in dropout probabilities for wired links, but all forms of rain, including light rain, exhibit higher dropout probability for wireless links. Conversely, gale-force winds have a profound increase in dropout probabilities for wired links; the lowest inflation in hourly P(Dropout) for all wired links is

---

[3]Addresses can receive 286 address-hours of thunderstorm in multiple ways: if one address receives 12 days of thunderstorm, 286 addresses receive one hour, 143 addresses receive two hours etc.
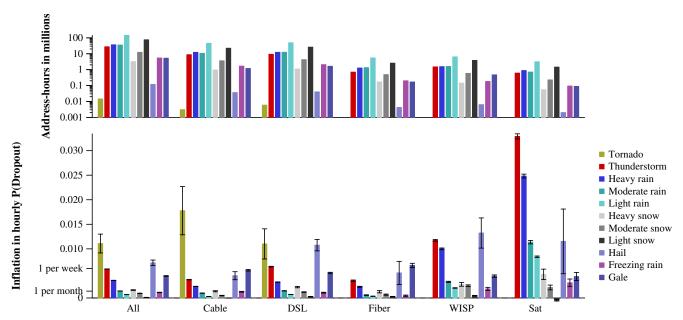
Figure 2: The number of responsive address-hours (in millions) for which we have measured various link types in various weather conditions (top, log-scale y-axis), and the additional ("inflated") probability of dropout experienced in those link- and weather-types (bottom).

for DSL (0.005082, corresponding to an additional failure every 8 days). However, wireless links are proportionally less likely to drop out during a gale. It is not surprising that strong winds can cause wired links to fail, for instance by knocking down above-ground cables. Although wireless links are not affected in the same way, it is surprising that higher failure rates would not be observed, given that such strong winds could destroy or blow away satellite dishes.

We also investigated the inflation in hourly dropout probability in the *aftermath* of weather conditions to check if failures are more likely to occur in the hours after a weather event. Only satellite links experienced statistically significant inflation in hourly dropout probability exceeding one per month in the aftermath of weather conditions, primarily in the hours after precipitation.

Summary and ramifications. The results from Figure 2 collectively show that different link types can experience weather in different ways. It is not surprising that different link types would differ in the *magnitude* with which they experience dropouts; but the differing patterns of high inflation in dropout probability for wired and wireless link types across weather patterns are suggestive. These results have ramifications on network measurement methodology: when performing outage analysis, it is important to account for both link type *and* weather condition.

4.2 Geographic variation

Next, we investigate the extent to which different geographic regions experience weather in different ways. Of course, different states experience different *amounts* of weather (for instance, we did not observe a statistically significant amount of snow in Florida). As in Section 4.1, we present the inflated probability of hourly dropouts, comparing address-hours with a particular weather condition for a state (e.g., snow) against baseline address-hours for that state.

This gives us an apples-to-apples comparison across states, even if they experience weather conditions in varying amounts.

In Figure 3, we present the hourly dropout probability inflation across all 50 U.S. states (and DC) for thunderstorms, rain (excluding hurricanes), snow, and gale-force winds (we omit other severe weather conditions from Figure 2 since many states did not have statistically significant results for them). We make two key observations. First, there is a high variation of increased dropout probability across states. For example, during thunderstorms, North Dakota experiences an average increased hourly dropout probability of 0.0124 (2.1 additional failures per week), while Washington D.C. increases by only 0.0015 (0.9 additional failures per *month*). Moreover, as shown by the 95% confidence intervals in the figure, these differences are statistically significant. We believe this to be an important result because it shows the role that geography plays in network outages.

Second, while the raw dropout inflation varies among states, the relative impact of weather types is common across *most* states: the increase in dropouts during thunderstorms tends to be greater than in rain, which in turn tends to be greater than in snow. There are a few notable exceptions. Louisiana and Mississippi have more inflated dropouts in snow than in thunderstorms, and Hawaii tends to experience similar amounts of failures in rain as it does in thunderstorms. By controlling for geography and the total amount of time spent in weather, this result shows that some weather conditions have more pronounced impact on dropouts.

Below Figure 3, we present a breakdown of the classified link types in each state, weighted by responsive address-hours in probing. We consult this graph to determine whether the outliers in the top graph are a direct function of the link types that are prevalent
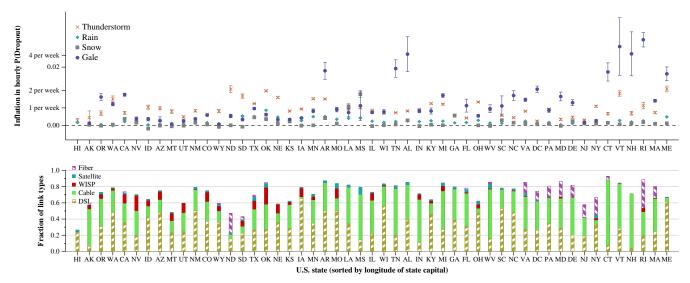
Figure 3: *Top:* **Inflation in hourly dropout probability by U.S. state for thunderstorm, rain, snow, and gale-force winds (95% confidence intervals).** *Bottom:* **Fraction of link types by U.S. state (the remaining fraction are of unknown type).**

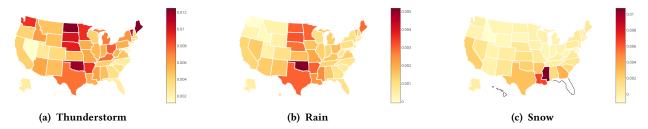**(a) Thunderstorm**　　　　**(b) Rain**　　　　**(c) Snow**

Figure 4: The mean hourly inflation in dropout probability by U.S. state for thunderstorm, rain, and snow. This is the same data as in Figure 3 (which includes the confidence intervals), re-visualized here to show regional trends. Large geographic regions can exhibit common behavior; northern states are more prone to failures in thunderstorms, Midwestern states in rain, and southern states in snow. (Note the different scales for each.)

in a state. North Dakota has a substantial and exceptional deployment of fiber: 50% of the link-type-classified responsive hours are from fiber addresses. Although our sampling approach is based on finding 100 addresses in each provider in a region, and thus is not meant to sample the distribution of link types used by customers, we note that this is consistent with published reports that "60 percent of the households, including those on farms in far-flung areas, have fiber" [42]. Although there are instances where top and bottom graphs appear related—Vermont (VT) and Maine (ME) show both a high vulnerability to thunderstorms and a relatively large proportion of DSL compared to immediate neighbors CT, NH, MA—it appears that geography is more important than link type at determining the inflation in probability of dropout in precipitation.

Next, we look beyond individual states to see if there are *regional* correlations of dropouts. In Figure 4, we show maps with the hourly inflation in dropout probabilities during thunderstorms, rain (excluding hurricanes), and snow. [4] For each U.S. state, this figure presents the mean hourly inflation in dropout probability values that we previously reported in Figure 3; we plot these values

on a U.S. map to enable regional comparisons visually. Figure 3 contains the confidence interval ranges for the mean values reported in Figure 4.

During thunderstorms (Figure 4(a)) and rain (Figure 4(b)) Midwestern states tend to experience greater inflation of dropouts than other regions. (Maine is an outlier; its dropout inflation during thunderstorm and rain is due to an abnormally powerful series of storms in October 2017.) Recall from Figure 2 that WISP and satellite links fail more often in thunderstorms and rain than other link types. One possible explanation for higher dropout rates in the Midwest would be that these states have more wireless links. This hypothesis is confirmed in Figure 3, which shows that Midwestern states have more satellite and WISP links than other states.

During snow (Figure 4(c)), we see more pronounced dropout inflation in southern states. [5] Louisiana and Mississippi experienced drastically higher probability of dropouts in snow. Unlike rain and thunderstorm, this disparity cannot be explained by link type alone, as no link types experience drastically higher dropout rates than others. Our insight is that snow seems to affect states where snow is less common.

---

[4]These plots and their data (confidence intervals, address hours, and dropout hours) are available at https://thunderping.cs.umd.edu.

[5]We do not include data for Florida or Hawaii, as we did not observe enough responsive hours of snow to achieve statistical significance (they did not satisfy Eq. (2)).
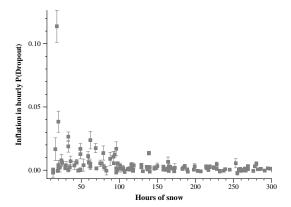
Figure 5: Inflation in hourly dropout probability of hosts (all link types) as a function of the number of hours the hosts' nearest U.S. airport received snow (truncated to only those with fewer than 300 hours in snow). The less common snow is in a region, the more impact it tends to have.

One possible explanation for the regional effects is therefore that regions that are less "familiar" with a particular weather condition may be more heavily affected by it. To evaluate this hypothesis, we plot in Figure 5 the inflation in hourly dropout probability of each U.S. airport as a function of the number of hours each airport has spent in snow. The results in this figure confirm our hypothesis for snow: the less familiar a location is to snow, the more often it tends to experience dropouts. Areas with very small amounts of snow do not experience large inflation (ostensibly because there is not enough time for it to cause damage). Conversely, areas with snow beyond a threshold are more resilient to snow. A likely reason for this is that regions that are more used to snow tend to invest more in infrastructure to prepare for and mitigate it [26]. We also performed this analysis under thunderstorms and rain (figures not shown), but did not observe as strong an effect. We hypothesize that this is because all of the airports we measured experienced enough thunderstorm and rain to grow accustomed to them.

Summary and ramifications. We conclude from these results that different geographic regions can be affected by weather to varying degrees. We attribute this geographic variation to two leading factors: (1) the predominance of some link types over others (e.g., wireless links are more common in the Midwest), and (2) how familiar a region is with a particular weather condition (and thus how prepared for it the region is). Our results have several interesting ramifications on outage analysis. First, when performing outage analysis, it is important to consider a representative set of locations and link types; measuring only, say, cable links would risk overestimating the Midwest's resilience to dropouts. Second, it is important to note the time and weather conditions when outage measurements are taken; collecting measurements only during Spring months, when thunderstorms are more common, would risk overestimating dropouts year-round.

4.3 Continuous weather variables

Thus far in our analysis, we have considered various *binary* classifications of weather—rain (or not), snow (or not), tornado (or not),
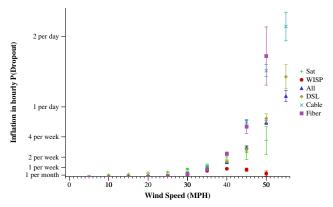

Figure 6: Inflation in hourly dropout probability as a function of wind speed across multiple link types. All link types experience greater dropout probabilities at higher wind speeds, but satellite and WISP links increase the least.
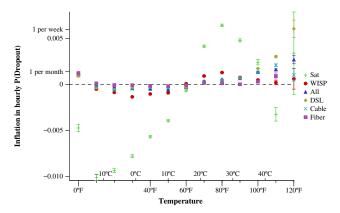

Figure 7: Inflation in hourly dropout probability as a function of temperature. All link types exhibit non-monotonic effects, typically increasing at higher and lower temperatures (satellite being an exception).

etc. Although these classifications are standard (they are included in the weather reports we collect), they risk masking the precise effect that weather conditions can have. Here, we evaluate dropouts as a function of several *continuous* weather variables: wind speed, precipitation, and temperature.

Figure 6 shows the inflation in the hourly dropout probability of various link types as a function of wind speed. For all link types, we see almost no inflation in dropout probability when wind speed is less than 30 mph. Beyond 30 mph, there are significant increases in dropout probability for wired links (cable, DSL, and fiber), but comparatively smaller effects on wireless links (WISP and satellite). This is reflected in Figure 2, which showed that wireless links were not as affected by gale-force winds. Figure 6 expands on this by showing that, as wind speed increases, dropout inflation increases at a super-linear rate—beyond 40 mph winds, Cable links' dropout inflation increases by an order of magnitude.

In Figure 7, we show dropout inflation as a function of temperature. There are several surprising observations in this figure. First,
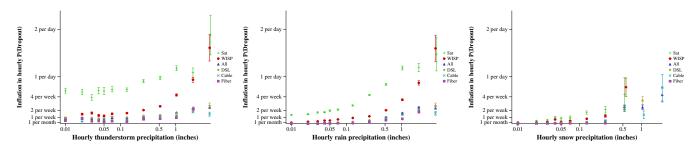
Figure 8: Inflation in hourly dropout probability as a function of precipitation during thunderstorm (left), rain (center), and snow (right), across multiple link types. All link types experience higher dropout probabilities with more precipitation, but to widely varying magnitudes. (Note the different ranges of the x-axes.)

satellite links are highly sensitive to temperature; at low temperatures, satellite links are far less likely to experience dropouts, but this increases steadily, until at approximately 70° F when satellite links become more likely to fail. Surprisingly, at approximately 80° F, there is an inflection point at which satellite links again become significantly more reliable. We hypothesize that there is a confounding factor: satellite links are less reliable when there is no line-of-sight visibility (e.g., due to fog), and we suspect that higher temperatures result in less fog.

All of the other link types we measured exhibit similar behavior to one another. They have highly variable dropout probabilities at low temperatures; they remain mostly steady until 60° F, then they increase slightly with higher temperatures. Unlike with our other results, WISPs more closely resemble wired links than satellite links; we hypothesize that this, too, is because satellite links are affected by line-of-sight while WISPs and wired links are not.

Finally, in Figure 8 we measure various link types' dropout inflation as a function of precipitation in thunderstorms, rain, and snow. All link types exhibit increased dropout inflation with increased precipitation, regardless of the overarching weather condition. However, surprisingly, the magnitude of increase varies significantly across link types. Again, satellite tends to be the most sensitive to change. Other link types are not as consistent across different types of precipitation; WISP links exhibit nearly the same increase in dropouts at high thunderstorm precipitation as satellite, but far less during non-thunderstorm rain.

There appears to be an inflection point with snow and rain: prior to 0.1 inches of precipitation in rain or snow, non-satellite links experience little change in their dropout probabilities. After these points, they increase significantly and quickly.

Summary and ramifications. Weather conditions are often described with binary categories: rain (or not), snow (or not), and so on. These continuous variable results show that such categories can be overly coarse; the mere presence of rain or snow does not necessarily affect most link types, unless there is more than 0.1 inches of precipitation. Like with our prior results, different link types can exhibit widely varying behaviors, lending further motivation to incorporate link types into future outage analyses.

5 RECOVERY TIME

In this section, we analyze ThunderPing probes to determine if there is a relationship between the weather condition during which

a dropout occurred, and the time to recover from that dropout. We discovered that during gale-force wind conditions, wired links take twice as long to recover than they do in baseline weather. Also, during precipitation (including snow), wired links take 1.5× longer to recover than they do in baseline weather. Surprisingly, although wireless links tend to have a higher likelihood of failures during weather than wired links, the recoveries are not significantly longer than those during baseline weather. Combining the observations of recovery time in this section with our observations of the likelihood of failures in the prior section (Section 4), we provide insight into the severity of weather-correlated failures.

Recent work has shown that ThunderPing probes (i.e., pings) alone are insufficient to accurately measure recovery time: residential links often recover with a different address after a failure due to dynamic addressing [44].[6] In this work, we overcome this limitation of the ThunderPing dataset by verifying recoveries with a dataset of host-side measurements of address continuity [49] obtained from a large CDN. With the CDN dataset, we were able to check the validity of ~1% of the recoveries observed by ThunderPing.

5.1 Our metric: Host-verified recovery time

We begin by describing how we determine if a recovery time observed with ThunderPing is valid. First, we describe the CDN dataset that provides host-side confirmation that an address did not change when a host recovered from a failure. Then, we describe how we combined this address continuity dataset with our ThunderPing dataset to identify valid recovery times.

Host-side address continuity dataset overview. The large CDN offers a service to content owners whereby end users can elect to install software that will improve the performance the client experiences when accessing the content through the CDN. The CDN records logs collected from its software installations on users' desktops and laptop machines. The CDN software logs contain a timestamp, unique identifier of the software installation (ID) on the client machine and the public source IP address visible to the CDN. Each logline contains (among other fields) the timestamp at which the logline was created, the unique identifier of the software installation on the machine (the ID), and the public IP address seen by the CDN's infrastructure at this time. Loglines in the CDN software dataset are dependent on user activity, and therefore, their

---

[6]Dynamic address reassignment after failure [44] invalidates the preliminary recovery time analysis from ThunderPing in 2011 [53].

	Verifiable	Valid
Cable	14,588	12,991 (89.1%)
DSL	25,999	5,315 (20.4%)
Fiber	1,445	609 (42.1%)
WISP	4,216	2,985 (70.8%)
Sat	3,926	3,025 (77.1%)
Total	60,264	31,418 (52.1%)

Table 2: The host-side CDN data verified that many recovery times were valid. It also made it possible for us to study recovery times for link types that often change addresses after a dropout (i.e., DSL and Fiber).

frequency varies. These logs contained data from 04/2013 to 12/2018. We described this dataset previously in Richter et al. [49].

Verifying recovery times observed by ThunderPing. We developed the following procedure to verify if a residential link has the same address after it recovers from a failure. First, we determine if a recovery time is verifiable by checking if the CDN logs contain a host ID for the address that failed; we allow for a one week window before the dropout (as we did in previous work [43, 49]), because the CDN logs are sporadic. Even though there was no coordination between the selection of addresses probed by ThunderPing and those observed by the CDN, *one percent* of ThunderPing's dropout's were verifiable (from 04/2013 to 12/2018).

To determine if a recovery time is valid, we then looked for the first host ID log in the CDN data after the recovery occurred. If the address associated with the ID is the same as the address before the dropout, then we consider that recovery time valid. [7] If the address is different, we do not know if the recovery time is valid—a laptop may have just moved to a different network.

Table 2 shows the recoveries that could be verified, by link type. In total, 52% of the recovery times were valid (the address did not change). The table also shows that nearly all dropouts for addresses with cable connections are not accompanied by address changes; this corroborates prior work [44]. Satellite and WISP addresses also tend to remain persistent across dropouts. Although DSL and fiber links are more likely to have their address reassigned when they recover from a dropout, there were still a small number of recoveries that were valid. This demonstrates the importance of incorporating the CDN logs into this study: these logs make it possible to study the recovery times of DSL and fiber.

5.2 Weather analysis

We now describe how recovery time correlates with the weather conditions when the link failed. Our hypothesis was that failures during weather conditions that presented an increase in likelihood of failure (e.g., snow, rain, thunderstorm, and gale), will take more time to recover than failures during baseline weather conditions. We observed that recovery times for wired links are indeed longer in these conditions, whereas, surprisingly, recovery times of wireless links are not significantly longer.

---

[7] It is possible that the address associated with the ID was reassigned to a different customer and then reassigned back; however, prior work indicated that the new address assigned to a customer link upon reassignment is often not even in the same /16 address block (50% of the time [44]).
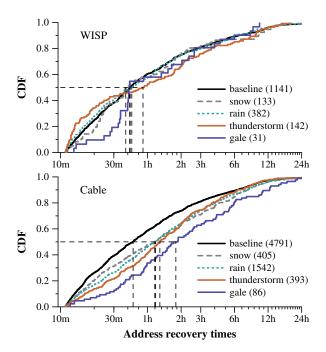

Figure 9: Recovery times for cable hosts tend to be longer when dropouts occur during snow, rain, thunderstorm, and gale-force winds. WISP host recovery times do not appear to be strongly correlated with weather.

For example, Figure 9 shows the distribution of recovery times for dropouts that occurred on cable (bottom) and WISP (top) links in several weather conditions (note the log-scale on the x-axis). For comparison, we also show the distributions for dropouts that occurred in baseline weather conditions. The numbers in parentheses show the number of recovery time samples in each curve, gridlines show median values. We clip the x-axis at 24h since more than 98% of recovery times (for all weather conditions) are below this value. Due to space constraints, we only show the recovery time for cable (wired) and WISP (wireless); however, these are representative of the distributions of the other wired and wireless link types. We observe that the baseline recovery time is similar for both wired and wireless link types: they both have a median recovery time of 45 minutes. This allows us to directly compare the effect of severe weather conditions on recovery times of cable and WISP links.

Cable host recovery times tend to be longer when dropouts occur during snow, rain, thunderstorm and gale compared to baseline. The median recovery time for dropouts beginning in snow is 71 minutes, in rain is 70 minutes, in thunderstorm is 77 minutes, and in gale-force winds is 108 minutes (more than 2× the baseline). We suspect that recovery times in gale-force winds tend to be the longest because they can cause physical damage to wires (especially ones in close proximity to trees), and recovering from such damage requires a truck roll. Recovery times during thunderstorms, rain, and snow, have surprisingly similar durations, in spite of potential lightning damage during thunderstorms—which could also lead to truck rolls. These results highlight that weather does not just increase the likelihood of outages for cable hosts, but that the outages tend to be longer than routine failures.

Conversely, WISP host recovery times do not appear strongly correlated with weather conditions. This means that although wireless links fail more often during certain weather conditions (Figures 2 and 8), they can recover quickly from these failures. We were surprised by this result because we expected that failure of wireless links would be catastrophic (e.g., dish goes out of alignment). The rapid recovery of wireless links may be because wireless infrastructure has a small physical footprint (no wires).

Even though there are relatively few recovery times that we could verify compared to the number observed with ThunderPing, we verified that these failures come from diverse airports. For example, for cable, even the 86 samples during gale-force winds were obtained from hosts geolocating to 54 different hurricane airports in diverse locations around the U.S. The samples for other weather conditions for the remaining link types are similarly diverse.

6 RELATED WORK

Internet outages are widely studied [1, 3–5, 8, 13, 14, 19, 21, 22, 24, 29, 30, 33, 34, 37, 39, 45, 47–49, 54, 58], but this work is the first to systematically investigate the effect of weather on outages. Of course, the analysis framework we introduced in Section 3 is generic and can be applied to outages detected by prior systems.

We focus this section on efforts that study residential links and that use active probes. We omit discussion of prior approaches that infer outages from BGP activity [19, 32, 35, 36, 62].

Residential measurements Behind-the-NAT residential network measurements can provide detail about events as well as precise location information. Such efforts include RIPE Atlas [50, 54], BIS-Mark [20, 57], and SamKnows [9, 51], as well as host-based software such as DIMES [2], Dasu [52], and PeerMetric [38]. In contrast, our goal is breadth and discovering rare weather related outages, particularly of atypical link types used for rural deployments.

Internet outage detection with active probes Chan et al. [11] study undersea cable cuts using TCP measurements. Cable cuts are widely reported outages that visibly affect BGP routes.

PlanetSeer [63] uses passive monitoring of wide-area traffic, not weather forecasts, to trigger active outside probing of hosts from PlanetLab nodes. Their results motivate our work: they found that 63% of outages occur within three hops of the edge of the network.

Paxson [46] surveys outages with large scale traceroutes to hosts. He classifies route outages based on the number of consecutive failed traceroutes. Consecutive failures can be common for residential link types (e.g., Satellite), so we use a conservative definition of a dropout that requires on 11 minutes of no responses, missing a set of pings from different vantage points across that interval.

Zmap was used to investigate outages in the wake of Hurricane Sandy [16]. However, their proof of concept weather measurement is limited to three days of a well-forecast hurricane.

Quan et al. [47] introduce "Trinocular" to analyze ping responses from a sample of hosts in every /24 (3.4 M) in the IPv4 address space and have also analyzed the effect of hurricanes [25]. Their approach considers the aggregated dropouts of a sample of a /24 prefix to represent a meaningful failure. We adopt parameters of their approach, for example the 11 minute safe probing interval and use of at least three vantage points. Quan et al. note that studying individual addresses as we do means that customer-initiated events,

such as turning off the access point, may be interpreted as failures. However, our approach expects that such address-specific events are typically independent of weather (and thus controlled for by §3.2), and if they occur in correlation with weather events, are no less meaningful individually.

7 CONCLUSIONS

Using an eight-year dataset collected by probing residential IP addresses in the U.S., we showed that a variety of weather conditions can inflate the likelihood of Internet dropouts. We quantified this inflation and show that it varies depending upon the type of weather, link type, and geographic location. We also showed that the time to recover from a dropout increases during weather events.

Our datasets and analyses enable a wide range of future studies. We focused our study on the weather conditions that appeared in a prior study to have significant impact on residential links [53], but our datasets span *all* weather reports. Our data therefore support myriad other inquiries, such as the combined effects of wind speed and rain, the effects of cloud cover, and so on. It is also possible to extend our analysis of recovery times, especially for link types like cable and satellite, where dynamic address reassignment is uncommon. For such link types, most recovery times measured by ThunderPing will be valid (as shown in Table 2), permitting additional inquiries such as the geographic variation of recovery times during weather conditions, the relationship between the duration of a weather condition (such as gale-force winds or a thunderstorm) and recovery times etc. Our dataset, collected over 8 years (and counting), allows the analysis of longitudinal trends in dropouts across link types, geographic areas, and weather conditions. Analyses of these trends can answer questions about whether the deployment of new technologies in an area results in improved reliability. Further, in this study, we only investigated "dropout" events, which by definition, only constitute complete packet-loss from three or more vantage points for at least 11 minutes; however, a promising avenue of future research is the study of how shorter events or partial failures are affected by weather. Another important area of future study is attributing *root causes* to dropouts under different weather conditions. Eto et al. [17] observed that residential U.S. homes experience ~1.5 power outages annually on average; we find considerably higher outage rates. While we cannot precisely compare two such disparate datasets, the large difference suggests that power outages are not the sole cause of dropouts.

To support researchers, network operators, and policy-makers, we make our data and code publicly available at: https://thunderping.cs.umd.edu. This work does not raise any ethical issues.

REFERENCES

[1] Abdulla Alwabel, John Healy, John Heidemann, Brian Luu, Yuri Pradkin, and Rasoul Safavian. 2015. *Evaluating Externally Visible Outages.* Technical Report ISI-TR-701. USC/Information Sciences Institute.

[2] Oded Argon, Anat Bremler-Barr, Osnat Mokryn, Dvir Schirman, Yuval Shavitt, and Udi Weinsberg. On the dynamics of IP address allocation and availability of end-hosts. arXiv preprint arXiv:1011.2324. (2010).

[3] Guillermo Baltra and John Heidemann. 2019. *Improving the Optics of Active Outage Detection (extended).* Technical Report ISI-TR-733. USC/Information Sciences Institute.

[4] Ritwik Banerjee, Abbas Razaghpanah, Luis Chiang, Akass Mishra, Vyas Sekar, Yejin Choi, and Phillipa Gill. 2015. Internet outages, the eyewitness accounts: analysis of the outages mailing list. In *Passive and Active Network Measurement Conference (PAM).*

[5] Robert Beverly and Matthew Luckie. 2017. The Impact of Router Outages on the AS-level Internet. In *ACM SIGCOMM.*

[6] Robert Beverly, Matthew Luckie, Lorenza Mosley, and kc claffy. 2015. Measuring and characterizing IPv6 router availability. In *Passive and Active Network Measurement Conference (PAM).*

[7] Tomasz Bilski. 2009. Disaster's impact on Internet performance – case study. In *International Conference on Computer Networks (CN).*

[8] Zachary S. Bischof, Fabián Bustamante, and Nick Feamster. 2018. The growing importance of being always on – a first look at the reliability of broadband internet access. In *Research Conference on Communications, Information and Internet Policy (TPRC).*

[9] Zachary S. Bischof, Fabián E. Bustamante, and Rade Stanojevic. 2014. Need, want, can afford: Broadband markets and the behavior of users. In *ACM Internet Measurement Conference (IMC).*

[10] Helmut Bölcskei, Arogyaswami J. Paulraj, K. V. S. Hari, Rohit U. Nabar, and Willie W. Lu. 2001. Fixed broadband wireless access: State of the art, challenges, and future directions. *IEEE Communications Magazine* (2001).

[11] Edmond W. W. Chan, Xiapu Luo, Waiting W. T. Fok, Weichao Li, and Rocky K. C. Chang. 2011. Non-cooperative diagnosis of submarine cable faults. In *Passive and Active Network Measurement Conference (PAM).*

[12] Giovanni Comarela, Gonca Gürsun, and Mark Crovella. 2013. Studying interdomain routing over long timescales. In *ACM Internet Measurement Conference (IMC).*

[13] Alberto Dainotti, Claudio Squarcella, Emile Aben, Kimberly C. Claffy, Marco Chiesa, Michele Russo, and Antonio Pescapè. 2011. Analysis of country-wide Internet outages caused by censorship. In *ACM Internet Measurement Conference (IMC).*

[14] Amogh Dhamdhere, Renata Teixeira, Constantine Dovrolis, and Christophe Diot. 2007. NetDiagnoser: Troubleshooting network unreachabilities using end-to-end probes and routing data. In *ACM Conference on emerging Networking EXperiments and Technologies (CoNEXT).*

[15] Ramakrishnan Durairajan, Carol Barford, and Paul Barford. 2018. Lights out: Climate change risk to internet infrastructure. In *Applied Networking Research Workshop (ANRW).*

[16] Zakir Durumeric, Eric Wustrow, and J Alex Halderman. 2013. ZMap: Fast internet-wide scanning and its security applications. In *USENIX Security Symposium.*

[17] Joseph H. Eto and Kristina Hamachi LaCommare. 2008. *Tracking the Reliability of the U.S. Electric Power System: An Assessment of Publicly Available Information Reported to State Public Utility Commissions.* Technical Report LBNL-1092E. Ernest Orlando Lawrence Berkeley National Laboratory.

[18] Manaf Gharaibeh, Anant Shah, Bradley Huffaker, Han Zhang, Roya Ensafi, and Christos Papadopoulos. 2017. A look at router geolocation in public and commercial databases. In *ACM Internet Measurement Conference (IMC).*

[19] V. Giotsas, C. Dietzel, G. Smaragdakis, A. Feldmann, A. Berger, and E. Aben. 2017. Detecting Peering Infrastructure Outages in the Wild. In *ACM SIGCOMM.*

[20] Sarthak Grover, Mi Seon Park, Srikanth Sundaresan, Sam Burnett, Hyojoon Kim, Bharath Ravi, and Nick Feamster. 2013. Peeking behind the NAT: an empirical study of home networks. In *ACM Internet Measurement Conference (IMC).*

[21] John Heidemann. Outage clustering: From leaves to trees. Talk at CAIDA Workshop on Active Internet Measurements (AIMS). (2018).

[22] John Heidemann, Yuri Pradkin, and Guillermo Baltra. 2018. The policy potential of measuring Internet outages. In *Research Conference on Communications, Information and Internet Policy (TPRC).*

[23] John Heidemann, Yuri Pradkin, Ramesh Govindan, Christos Papadopoulos, Genevieve Bartlett, and Joseph Bannister. 2008. Census and survey of the visible Internet. In *ACM Internet Measurement Conference (IMC).*

[24] John Heidemann, Yuri Pradkin, and Aqib Nisar. 2018. *Back Out: End-to-end Inference of Common Points-of-Failure in the Internet (extended).* Technical Report ISI-TR-724. USC/Information Sciences Institute.

[25] John Heidemann, Lin Quan, and Yuri Pradkin. 2012. *A Preliminary Analysis of Network Outages During Hurricane Sandy.* Technical Report. USC/ISI.

[26] Chris Hill. 23 state DOTs spent more than $1 billion on snow, ice maintenance this winter. https://www.equipmentworld.com/

23-state-dots-spent-more-than-1-billion-on-snow-ice-maintenance-this-winter/. (May 2015).

[27] D.C. Hogg and Ta-Shing Chu. 1975. The role of rain in satellite communications. *Proc. of the IEEE* (1975).

[28] Gianluca Iannaccone, Chen nee Chuah, Richard Mortier, Supratik Bhattacharyya, and Christophe Diot. 2002. Analysis of link failures in an IP backbone. In *ACM Internet Measurement Workshop (IMW).*

[29] Internet Outage Detection and Analysis (IODA). Internet outage detection and analysis (IODA). https://www.caida.org/projects/ioda/. ([n. d.]).

[30] Umar Javed, Italo Cunha, David R. Choffnes, Ethan Katz-Bassett, Thomas Anderson, and Arvind Krishnamurthy. 2013. PoiRoot: Investigating the root cause of interdomain path changes. In *ACM SIGCOMM.*

[31] Frank B. Jewett. 1909. The modern telephone cable. In *Proceedings of 26th annual convention of the American Institute of Electrical Engineers.*

[32] John P. John, Ethan Katz-Bassett, Arvind Krishnamurthy, Thomas Anderson, and Arun Venkataramani. 2008. Consensus routing: The Internet as a distributed system. In *Symposium on Networked Systems Design and Implementation (NSDI).*

[33] Ethan Katz-Bassett, Harsha V. Madhyastha, John P. John, Arvind Krishnamurthy, David Wetherall, and Thomas Anderson. 2008. Studying black holes in the internet with hubble. In *Symposium on Networked Systems Design and Implementation (NSDI).*

[34] Ethan Katz-Bassett, Colin Scott, David R. Choffnes, Ítalo Cunha, Vytautas Valancius, Nick Feamster, Harsha V. Madhyastha, Thomas Anderson, and Arvind Krishnamurthy. 2012. LIFEGUARD: Practical repair of persistent route failures. In *ACM SIGCOMM.*

[35] Nate Kushman, Srikanth Kandula, and Dina Katabi. 2007. Can you hear me now?!: It must be BGP. *ACM SIGCOMM Computer Communication Review (CCR)* 37, 2 (March 2007).

[36] Craig Labovitz, Abha Ahuja, Abhijit Bose, and Farnam Jahanian. 2000. Delayed Internet routing convergence. In *ACM SIGCOMM.*

[37] Craig Labovitz, Abha Ahuja, and Farnam Jahanian. 1999. Experimental study of internet stability and wide-area backbone failures. In *International Symposium on Fault-Tolerant Computing (FTCS).*

[38] Karthik Lakshminarayanan and Venkata N. Padmanabhan. 2003. Some findings on the network performance of broadband hosts. In *ACM Internet Measurement Conference (IMC).*

[39] Ratul Mahajan, David Wetherall, and Thomas Anderson. 2002. Understanding BGP misconfiguration. In *ACM SIGCOMM.*

[40] MaxMind. Geoip2 city accuracy. https://www.maxmind.com/en/geoip2-city-accuracy-comparison?country=&resolution=100. (Jan. 2019).

[41] NOAA. Automated surface observing system (ASOS). https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/automated-surface-observing-system-asos. ([n. d.]). ftp://ftp.ncdc.noaa.gov/pub/data/noaa/.

[42] Carolyn Orr. A look at how and why North Dakota became a leader in deployment of fiber optic Internet. http://www.csgmidwest.org/policyresearch/0616-fiber-optic-North-Dakota.aspx. (Jun 2016).

[43] Ramakrishna Padmanabhan. 2018. *Analyzing internet reliability remotely with probing-based techniques.* Ph.D. Dissertation. University of Maryland.

[44] Ramakrishna Padmanabhan, Amogh Dhamdhere, Emile Aben, kc claffy, and Neil Spring. 2016. Reasons dynamic addresses change. In *ACM Internet Measurement Conference (IMC).*

[45] Ramakrishna Padmanabhan, Aaron Schulman, Alberto Dainotti, Dave Levin, and Neil Spring. 2019. How to find correlated internet failures. In *Passive and Active Network Measurement Conference (PAM).*

[46] Vern Paxson. 1997. End-to-end routing behavior in the Internet. *IEEE/ACM Transactions on Networking* (Oct. 1997).

[47] Lin Quan, John Heidemann, and Yuri Pradkin. 2013. Trinocular: Understanding Internet Reliability Through Adaptive Probing. In *ACM SIGCOMM.*

[48] Lin Quan, John Heidemann, and Yuri Pradkin. 2014. When the Internet sleeps: Correlating diurnal networks with external factors. In *ACM Internet Measurement Conference (IMC).*

[49] Philipp Richter, Ramakrishna Padmanabhan, Neil Spring, Arthur Berger, and David Clark. 2018. Advancing the art of internet edge outage detection. In *ACM Internet Measurement Conference (IMC).*

[50] RIPE NCC. RIPE Atlas. http://atlas.ripe.net. ([n. d.]).

[51] SamKnows. Test methodology white paper. https://availability.samknows.com/broadband/uploads/Methodology_White_Paper_20110701.pdf. (2011).

[52] Mario A. Sánchez, John S. Otto, Zachary S. Bischof, David R. Choffnes, Fabián E. Bustamante, Balachander Krishnamurthy, and Walter Willinger. 2013. Dasu: Pushing experiments to the internet's edge. In *Symposium on Networked Systems Design and Implementation (NSDI).*

[53] Aaron Schulman and Neil Spring. 2011. Pingin' in the rain. In *ACM Internet Measurement Conference (IMC).*

[54] Anant Shah, Romain Fontugne, Emile Aben, Cristel Pelsser, and Randy Bush. 2017. Disco: Fast, good, and cheap outage detection. In *Network Traffic Measurement and Analysis Conference (TMA).*

[55] Yuval Shavitt and Noa Zilberman. 2011. A geolocation databases study. *IEEE Journal on Selected Areas in Communications* 29, 10 (2011), 2044–2056.

[56] W. T. Smith and W. L. Roberts. 1966. Design and characteristics of coaxial cables for Community Antenna Television. *IEEE Transactions on Communication Technology* (1966).

[57] Srikanth Sundaresan, Walter de Donato, Nick Feamster, Renata Teixeira, Sam Crawford, and Antonio Pescapè. 2011. Broadband Internet performance: A view from the gateway. In *ACM SIGCOMM*.

[58] Daniel Turner, Kirill Levchenko, Alex C. Snoeren, and Stefan Savage. 2010. California fault lines: Understanding the causes and impact of network failures. In *ACM SIGCOMM*.

[59] U.S. Government. 47 CFR part 4 section 4.9: Outage reporting requirements threshold criteria. ([n. d.]).

[60] Gerald van Belle, Patrick J. Heagerty, Lloyd D. Fischer, and Thomas S. Lumley. 2004. *Biostatistics: A Methodology for the Health Sciences (Second Edition)*. John Wiley & Sons.

[61] J. M. Wallace. 1975. Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States. *Monthly Weather Review* (1975).

[62] Feng Wang, Zhuoqing Morley Mao, Jia Wang, Lixin Gao, and Randy Bush. 2006. A measurement study on the impact of routing events on end-to-end internet path performance. In *ACM SIGCOMM*.

[63] Ming Zhang, Chi Zhang, Vivek Pai, Larry Peterson, and Randy Wang. 2004. PlanetSeer: Internet path failure monitoring and characterization in wide-area services. In *Symposium on Operating Systems Design and Implementation (OSDI)*.