**Plots, planes, and pixels: integrating data sources to quantify mass forest mortality extent, severity, and vulnerability**

**Project Summary**

In 2022, the USDA Forest Service Forest Health Protection Aerial Detection Survey (ADS) documented extensive mortality of true firs (*Abies* spp.) in Oregon. Referred to as “Firmageddon”, concerns about its potential economic and ecological impacts highlight the need to develop unbiased and verifiable information on forest mortality events, and improve our understanding of forest vulnerability in the future. However, available datasets about forest disturbances differ in various important aspects, including their spatial and temporal resolution, extent, agent attribution, and severity assessment. This makes gaining a cohesive picture of large-scale forest mortality events difficult. This project compares and integrates three forest disturbance data sources – plot-based data from the USDA Forest Service Forest Inventory and Analysis Program (FIA), aerial classification of disturbed patches from the Aerial Detection Survey (ADS) program, and remotely-sensed satellite time series imagery – in order to improve estimates of forest disturbance extent, severity, and vulnerability.

**Objectives**

1. *Data source comparisons*

Review and quantify the strengths, weaknesses, and complementary aspects of FIA, ADS, and remotely-sensed data. Compare three core metrics (area impacted, attribution, and severity) across data sources in the context of two case-studies: 2022 Firmageddon event in OR and 20XX pine beetle outbreak in XXXXX.

1. *Data source integration*

Develop a data processing and modeling workflow for integrating ADS, multispectral satellite imagery, and FIA plot measurements to provide spatially-explicit estimates of proportional tree mortality.

1. *Landscape contextualization*

Combine modeled mortality estimates with imputed forest attribute data from the LEMMA GNN project, interpreting model outputs as on-the-ground impacts.

1. *Retrospective and prospective vulnerability*

Use mortality estimates together with GNN forest attribute data and ancillary climatic and biophysical data to model the drivers of large-scale forest mortality. Assess our ability to retrospectively predict vulnerability (i.e., using Firmageddon and pine beetle case studies), and map potential future vulnerability to similar events.

**Approach**

*Data source comparisons*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **FIA** | **ADS** | **Landsat** |
| **Extent** | *temporal* | 2000-2019 (for annual design) | 1970s? - present | 1984 - present |
|  | *spatial* | national | Depends on annual flown area | global |
| **Scale** | *temporal* | difference between inventories | ? | annual? |
|  | *spatial* | estimation strata; 90 m plot | variable -- watershed? | pixel |
| **Resolution** | *temporal* | decadal or annual sample | annual? | annual |
|  | *spatial* | 90m plot; individual tree | variable | 30 m |
| **Effects** | *attribution* | direct observation, trained crews | indirect observation, expert evaluation | remote observation, spectral analysis |
|  | *specificity* | variable; low specificity for dead, higher specificity for live damage | qualitative determination; variable by agent and region | quantitative determination; generic agent |
|  | *severity* | tree-level effects | variable; stand to watershed-level; binned | quantitative; severity requires context |

We will review literature on field, ADS, and spaceborne approaches to forest mortality estimation and mapping, and catalogue the relevant aspects of each (*e.g.,* Table 1). We will quantitatively compare three aspects of forest mortality estimation across all three data sources – area impacted, attribution, and severity. For FIA-based area estimates, we will follow FIA estimation procedures (Bechtold and Patterson, 2005). For ADS estimates, area impacted can be directly gleaned from survey polygons. Remotely-sensed area estimates will be made using the Landscape Change Monitoring System (LCMS) and/or LandtrendR. Attribution … etc build out.

**Table 1.** Comparison of data attributes.

*Data source integration and landscape contextualization*

Building on prior research in the Pacific Northwest (Meigs et al., 2011, 2015), we will build a workflow integrating FIA, ADS, and remotely sensed data for mortality mapping and estimation (*e.g.,* Figure 1). This workflow has three major components: (1) Merging ADS polygons with Landsat-based disturbance mapping to generate multispectral imagery maps of attributed mortality with pixel-scale spatial resolution, (2) develop statistical models relating Landsat-spectral change to observed mortality on FIA plots, yielding pixel-scale spatial estimates of tree mortality, and (3) combine the results of 1-2 with gradient nearest neighbor (GNN, <https://lemma.forestry.oregonstate.edu/>) maps of forest attributes to contextualize and interpret mortality estimates.

*Retrospective and prospective vulnerability*

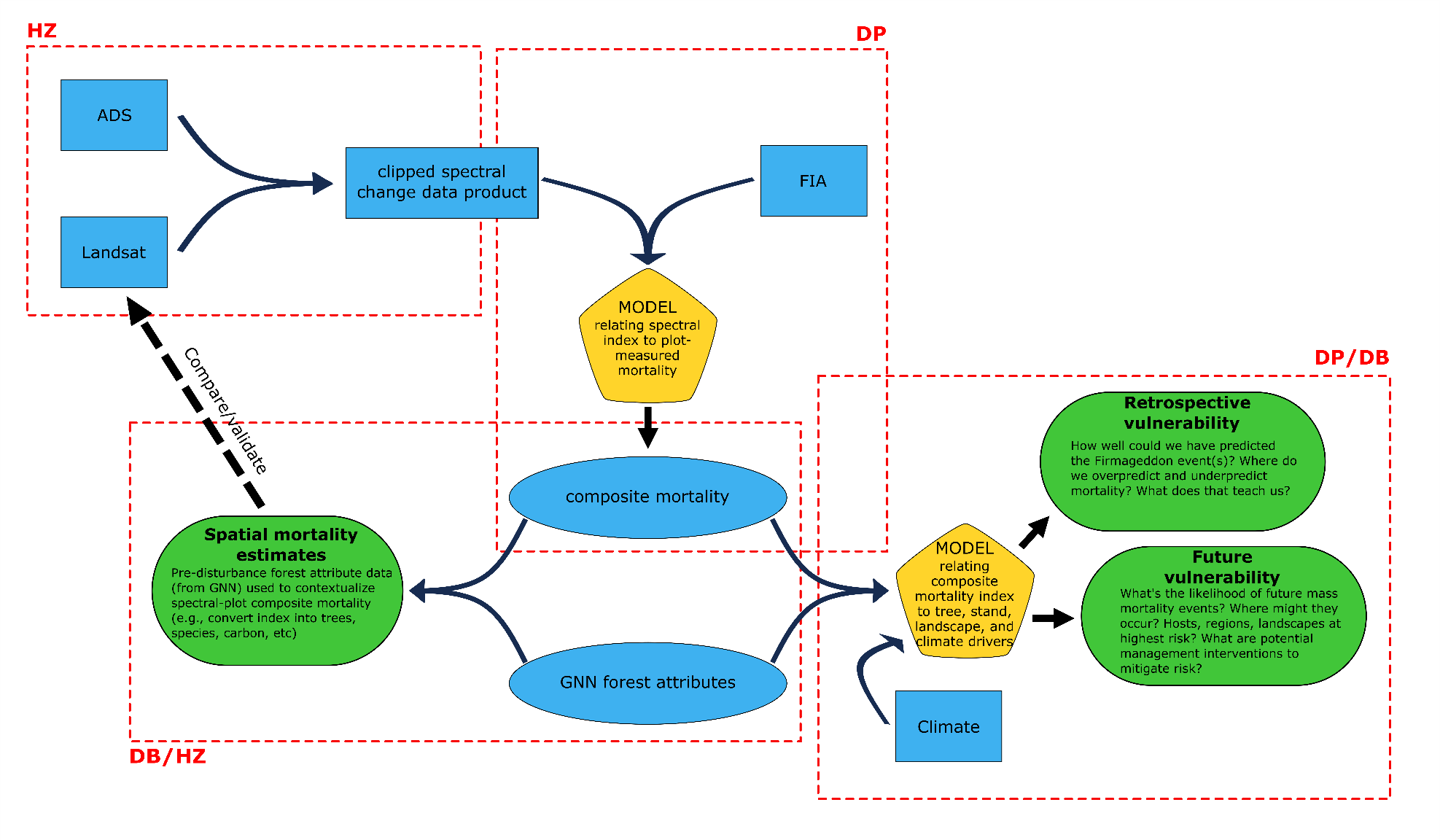
The data source integration we propose will generate fine-scale spatial mortality estimates that can be combined with tree-, stand-, and landscape-scale data to examine the drivers of mass mortality events. In particular, we will build models exploring how mortality varies with tree size, stand density, climate, biophysical setting, and topographic position (e.g., Bell et a. 2015, Shriver et al. 2021). Understanding how these variables modulate the probability of mass mortality events could guide management prescriptions for resilient forests in the future (Bradford & Bell 2017). We will challenge our mortality models to retrospectively predict spatial patterns of mortality for both our Firmageddon and pine beetle case studies, to both assess our ability to make these predictions and gain insight into when and where mortality may be more/less predictable. We will also use future climate projections to identify and map areas vulnerable to future large-scale mortality events.

**Questions and points of engagement for Danny**

* How should we think about the spatial precision and accuracy of ADS polygons?
* We should think about ADS data as having a higher rate of accuracy then precision. Depending on the forest damage. If mortality is discrete and obvious, like a mountain pine beetle outbreak in a pure stand of lodgepole then the data should be high on both counts. In contrast, if the damage is hard to see like western spruce budworm and widespread, then lower levels of confidence until it gets really bad.
* Are there nuances to the temporal aspects of ADS data (e.g., surveyed area, data processing lags, seasonality of surveys, etc) Yes, always caveats associated but all manageable. Some nuances to be aware of, the ADS maps mortality when we see the crowns change (same as space based change detection) so for many disturbance agents the infestation actually happened the year before. The ADS data is just a snap shot in time, so we have mortality that bleeds into the following year, depending on when those trees finally succumbed to the pressure. We don’t have the funds to fly areas multiple times, so if a heat dome type event comes along after the survey and pushes forests to the breaking point after we have surveyed that area we may not capture that until the following year.
* We process data and turn it into the WO most years by November 15. The survey window goes from Late June to October 1st. We do our best to start in low elevation and south and work our way up from there. We do our best to get wall to wall survey for most forest types, that said it doesn’t always happen. We then normalize the data by looking at the ratio of area surveyed / damage recorded when looking at trends over time. We also try to only analyze damage trends over very large areas and don’t suggest using this data for very localized projects with out ground level verification.
* What are the most common ways that people misuse ADS data (so we can avoid doing so) Some of the ways ADS data is misused is taking the data as straight fact when the story is more complex. For example, I was just contacted by a Ph. D student from Idaho. They found the ADS data for Region 6 and see 75 years of data on what appears to be western pine beetle data. Reality is that was just the one bark beetle they went with to describe mortality in larger ponderosa, which is the more common player, but there are several bark beetles in those areas and most likely in the same tree. Ips on top, Western in the middle and turpentine in the bottom of the bole. They wanted to take the data to make a model for just western pine beetle and I had to tell them that is not accurate and an inappropriate use of the data. Other damages show just the entomology side in the data, when diseases play a significant role in tandem depending on the host.

People also take the data and say acres of mortality when we are recording acres with mortality. Then in the acres with part we assign a percent severity. We can get to acres of mortality, but more math is needed to make that cross walk. The TPA data from historical data is not very dependable, kind of like the TPA from the recent LiDAR data. It is close, but not sure I would use it as a major component of my analysis.

* Is our assumption that fir mortality is harder to detect than e.g., pine beetle mortality correct? Not sure I understand this question. Fir mortality is harder to detect than pine mortality because it is easier to confuse with Douglas-fir mortality and lodgepole mortality when the lodgepole is in a mixed conifer stand. They all kind of turn the same color and we then try to decipher the host by crown structure. True firs are have a more pointy top and the crown is more likely to go lower on the bowl then a Douglas-fir. The other problem with true fir is we can’t tell the difference between most of them. The exception is subalpine fir, and Pacific fir (if the lighting is just right). The rest of the firs we lump together. Do those case studies really represent the “end members” of that spectrum? Not following this question, but we can discuss when we meet. There are more bits and pieces with above questions as well, but maybe this is enough for now and we can flush out the rest when we meet?



**Figure 1** (*previous page*). **Conceptual workflow diagram for data integration**. Diagram shows the workflow combining primary data sources (blue rectangles; ADS, Landsat, FIA) to generate modeled data (blue ellipses; composite mortality, GNN forest attributes), which are either directly combined or used to train a model (yellow polygon) to generate end products (green ovals; spatial mortality estimates, retrospective vulnerability, future vulnerability). Red dashed lines roughly delineate processing/analysis steps, with initials assigning primary responsibility.