



SAVVY MAINTENANCE / OPINION

Borescope meets AI

A status report on a work in progress

BY MIKE BUSCH

IN MY COLUMN in the June 2025 issue of *AOPA Pilot* (“Borescope Initiative”), I reported on the initiative we launched on January 1, 2024, to promote and improve the use of borescopes as the gold standard for assessing cylinder condition for piston aircraft engines. Among other things, we defined a standard 11-image protocol for inspecting each cylinder, and we created a cloud-based borescope image repository for uploading, identifying, analyzing, archiving, and reporting on images. In just over a year from launch, this repository has

grown to more than 100,000 borescope images. Roughly 70 percent of those were uploaded by SavvyAnalysis clients who asked our team of experts to analyze and report on the cylinder condition, and the remaining 30 percent were uploaded by users who did their own analysis.

The image analysis—regardless of who’s doing it—is accomplished using a screen that displays each image and offers a series of buttons for labeling the image. For instance, if the image being viewed is the head of an exhaust valve

(i.e., the part you see when the valve is closed), the buttons are labeled:

- Normal
- Uneven heat pattern—early
- Uneven heat pattern—mid-stage
- Uneven heat pattern—advanced
- Crack or chip
- Eroded
- Undetermined
- Bad image

We use the term “uneven heat pattern” or “heat distressed” because we think it’s more accurate than “burned valve.” We subdivide heat distress into three gradations: early, mid-stage, and advanced. Exhaust valves exhibiting early or mid-stage heat distress are excellent candidates for lapping in place, whereas valves exhibiting advanced heat distress may require pulling the cylinder and having the valve and seat replaced.

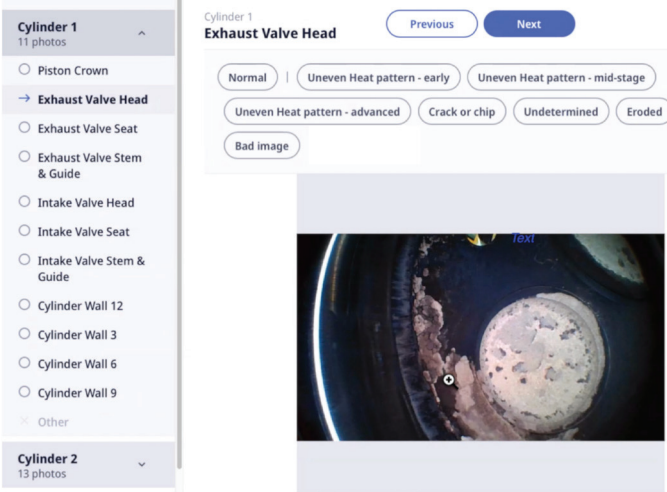
Can image analysis be automated?

Presently, analysis of these borescope images requires human expertise and judgment. But we wondered whether this was something we could train an AI model to do. We’ve been experimenting with this for the past few months. While this project is still a work in process and not yet ready for prime time, the early results are encouraging.

We started out with a simple goal: to create an AI model that could look at images of an exhaust valve head and figure out whether the valve looks normal or heat distressed. This would involve training the AI model on a bunch of exhaust valve images that our human analysts have already analyzed and labeled as normal or heat distressed. Hopefully the model could be trained to discriminate between normal valves and heat distressed ones.

At this point, we have only a limited number of images we could use to train the model. Of the more than 100,000 images uploaded to the repo, only about 10,000 of them are exhaust valve heads. (The rest are seats, stems, intake valves, cylinder walls and piston crowns.) Of those 10,000 exhaust valve head images, fewer than 800 are heat-distressed, and only a small fraction of those are in an advanced

STEPHANIE DALTON COWAN



state of distress. That’s not much training data. We probably won’t have enough images to train the model well until several hundred thousand more images have been uploaded to the repo. But we were still interested in finding out how well we could do with what we’ve got.

The AI model

Our initial model actually consists of two submodels. The first one is simply an object detection model that locates the exhaust valve within the image and crops the image to include only the exhaust valve head.

The second submodel takes the cropped image as input and generates a prediction of the valve condition. To keep things simple, it’s just a binary classifier that identifies the image as “normal” or “distressed”—it doesn’t (yet) attempt to predict whether a distressed valve is early, mid-stage, or advanced. In addition to its binary verdict, the classifier reports a confidence score. So, it might predict “I am 99 percent confident that this valve is normal” or “I am 67 percent confident that this valve is burned.”

This is a supervised machine learning model, which means that it was trained on images that were previously reviewed by our human analysts and labeled as normal or heat distressed. We used 1,827 labeled images of exhaust valve heads—70 percent were used for training, 20 percent for validation, and 10 percent for testing.

That’s not much data. If we had more, the model could be better trained and

achieve better performance. But despite the limited amount of training data, this first-try AI model did surprisingly well.

How did it perform?

We evaluated the model using a test set of 182 labeled images to see how the AI’s performance compared to our human analysts. Model performance was evaluated on two measures:

Recall (“catch rate”): Given an image labeled as “normal” or “distressed” by a human analyst, how likely is it that the AI

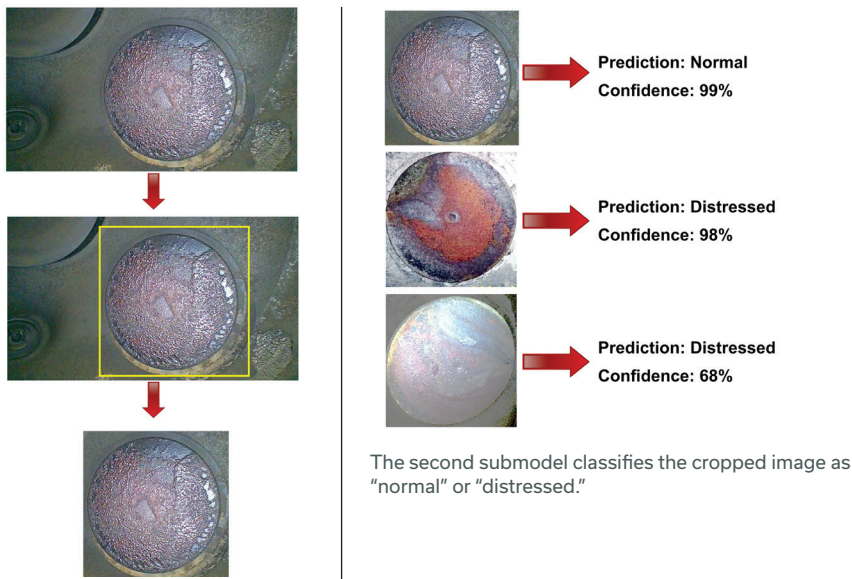
Borescope repository analysis screen for an exhaust valve head image.

model’s prediction agrees with the human analyst? A high recall of distressed valves means that the AI model catches most distressed valves, although it might also incorrectly catch some normal valves in the process (false positives).

Precision (“correctness”): Given an image predicted as “normal” or “distressed” by the AI model, how likely is it that the human analyst agreed with the AI model’s prediction? A high precision of distressed valves means that the AI model’s predictions that a valve is distressed is usually correct, although it might incorrectly classify some distressed valves as being normal (false negatives).

The test results are shown in a table known as a “confusion matrix.” It shows that the model caught 93 percent of the normal valves but only 62 percent of the distressed valves (about two out of three). That’s not surprising, because the distressed valves were mostly early and mid-stage valves where the cues can be subtle, even for our human analysts. The model did a lot better on valves that were in advanced distress. Of the 22 such valves in the test set, the model caught 20 of them for a recall rate of 91 percent.

Precision was pretty good: 81 percent for valves predicted as “normal” (19



The first submodel locates the exhaust valve and crops the image.

The second submodel classifies the cropped image as “normal” or “distressed.”

		AI Model		
		Normal	Distressed	Recall
Human	Normal	108	8	93%
	Distressed	25	41	62%
Precision		81%	84%	

AI model test results “confusion matrix.”

percent false positives) and 84 percent for those predicted as “distressed” (16 percent false positives). Precision can be increased by ignoring model predictions whose confidence score is less than, say, 98 percent, and classifying the valve as “undetermined” for predictions with lower confidence scores. The downside of doing this is that recall (catch rate) is degraded. Tuning the model is always a tradeoff between recall and precision.

Good enough to be useful?

At this early stage with limited training data, the model shows promise but it’s certainly not good enough to take the place of our human analysts. It presently catches only about two out of three distressed valves, although it’s quite adept at catching valves in advanced distress. It does quite well catching normal valves (about 13 out of 14). In both cases, the false positive rate is less than one in five.

Based on our prior experience with predictive analytics—notably our Failing Exhaust Valve Analytics (FEVA) model that predicts exhaust valve failure risk by analyzing digital engine monitor data—it can be risky to expose AI predictions of maintenance issues directly to our clients unless the false positive rate is very low. We find that aircraft owners tend to overreact to AI predictions that their engine might have a problem and then get angry if the prediction turns out to be wrong. So, we might think twice before exposing an AI

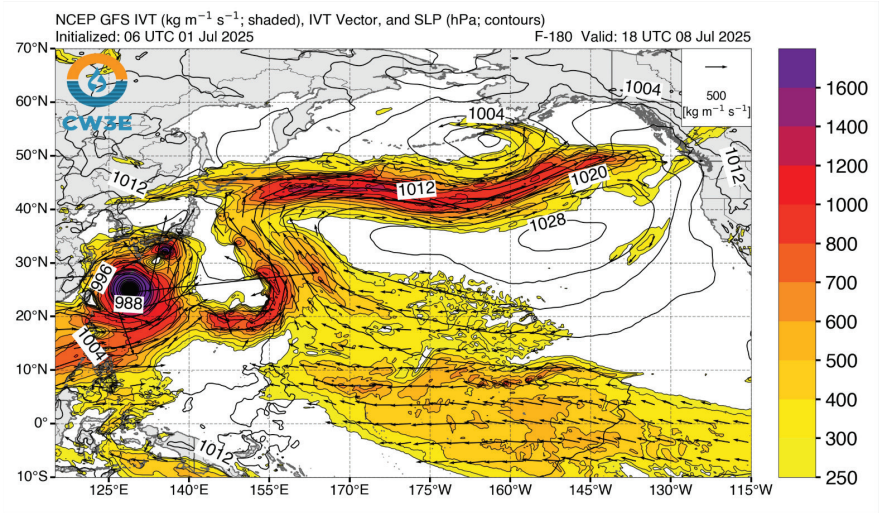
prediction of a distressed exhaust valve to a client without first passing it by one of our human analysts.

On the other hand, the model’s current performance makes it quite useful for flagging uploaded images that warrant human scrutiny. We have started doing that. As more borescope images are uploaded to our repo and can be used to train the AI model and improve its performance (particularly its precision), we feel confident that it will reach the point where we can expose its predictions directly to our clients, and perhaps free up our human analysts to do other things.

We live in interesting times. For the record, I wrote this article without help from AI. I prompted both ChatGPT4 and Grok3 to “write a 1,500-word article in the style of Mike Busch’s Savvy Maintenance columns in *AOPA Pilot* magazine” but was dissatisfied with either result. I’ll keep trying, and my best guess is that it won’t be long. ■
 mike.busch@savvyaviation.com
 ▶ savvyaviation.com

Ask the A&Ps /

Mike Busch, Paul New, and Colleen Sterling answer your toughest aviation maintenance questions on our Ask the A&Ps podcast. Submit questions to podcasts@aopa.org ▶ aopa.org/ask-the-a-and-ps



Computer models (above) help predict locations of atmospheric rivers a week into the future. Here we see a Category 5 event on its way to the Pacific Northwest. When it draws nearer to shore, Doppler, datalink, and airborne weather radars will provide real-time views of precipitation contours.

WX WATCH

Rivers of rain

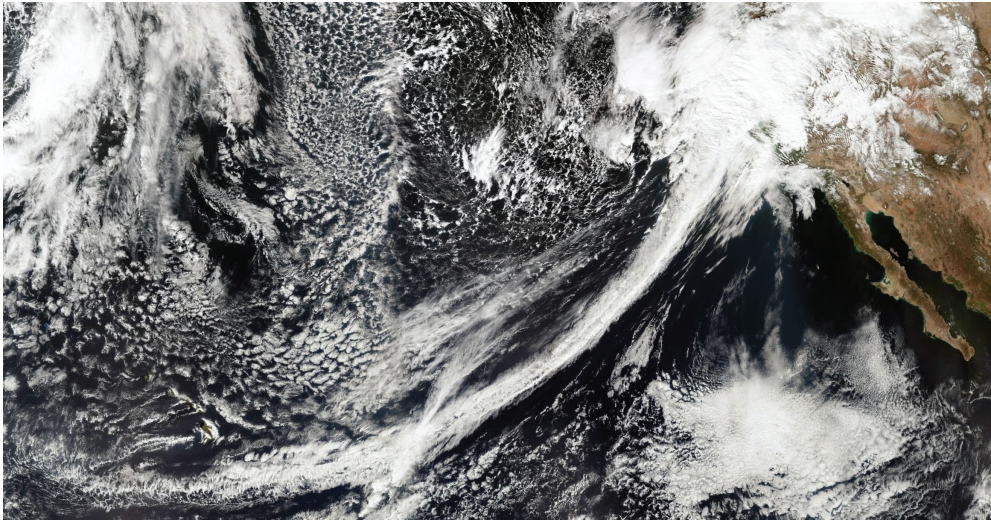
The main driver of West Coast torrential rain

BY THOMAS A. HORNE

NOT LONG AGO, the custom was to use the term “Pineapple Express” to describe a huge rainfall event that strikes the West Coast. That’s a pretty whimsical name for such a destructive, long-lasting climatic feature. Now they’re called atmospheric rivers, a term that gives more respect to these thousands-of-miles long, hundreds-of-miles wide cascades of flooding rains that routinely make landfall in the fall and winter months. Typically, dozens of atmospheric rivers make landfall each year, usually in California, but Oregon and Washington aren’t exempt. In the 2023 season, there were 46 atmospheric river events in California. And when they arrive, they last for as long as three to four days, soaking the terrain for hundreds of miles inland. Two to four feet of total rainfall is routine.

As you might expect, the result is widespread flooding and mudslides. On the other hand, atmospheric rivers do have benefits. They bring half of

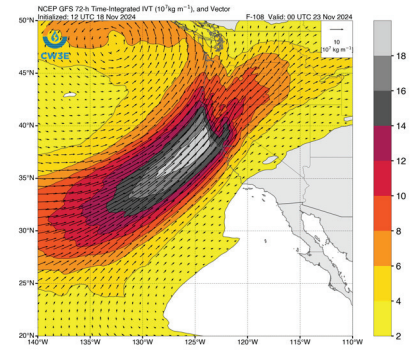
California’s total annual rainfall, fill reservoirs, and put an end to drought and wildfires. As for aviation, it takes little imagination to think of the impacts: widespread instrument meteorological conditions in cloud and rain, plus icing



conditions as the atmospheric rivers’ flows move up and over high terrain.

Atmospheric rivers tend to form when low pressure occurs in the central areas of the Pacific Ocean. (It’s those that originate near the Hawaiian Islands that prompted the Pineapple Express handle.) Convergent winds then lift the plentiful, warm, moisture-soaked air up into the westerly trade winds and then this soggy plume is on its way east, picking up more and more water vapor as it goes.

Atmospheric rivers are quantified by measurements of integrated vapor transport (IVT)—in other words, the weight of the water vapor flowing through the air, in a column from the surface to the flight levels. Depending on the amount



The Global Forecast System (GFS) model (above) is predicting a chart-busting amount of water vapor/rainfall (gray shades) for most of California. A geostationary satellite image (below) shows that atmospheric river plumes can span the Pacific Ocean.