

The VespAI logo is centered within a large red speech bubble. The word "VespAI" is written in a white, sans-serif font. The background of the slide features a light gray pattern of concentric circles and dashed lines, creating a sense of depth and movement.

# VespAI

Testing results

# Contents



Models trained

Dataset index



Testing data

Dataset index



Performance observations

Headlines for the paper: 2021 models on 2022 data  
Results on increasing training data size  
Insect pollutants



Gotchas

Wooden Hornets  
Good news

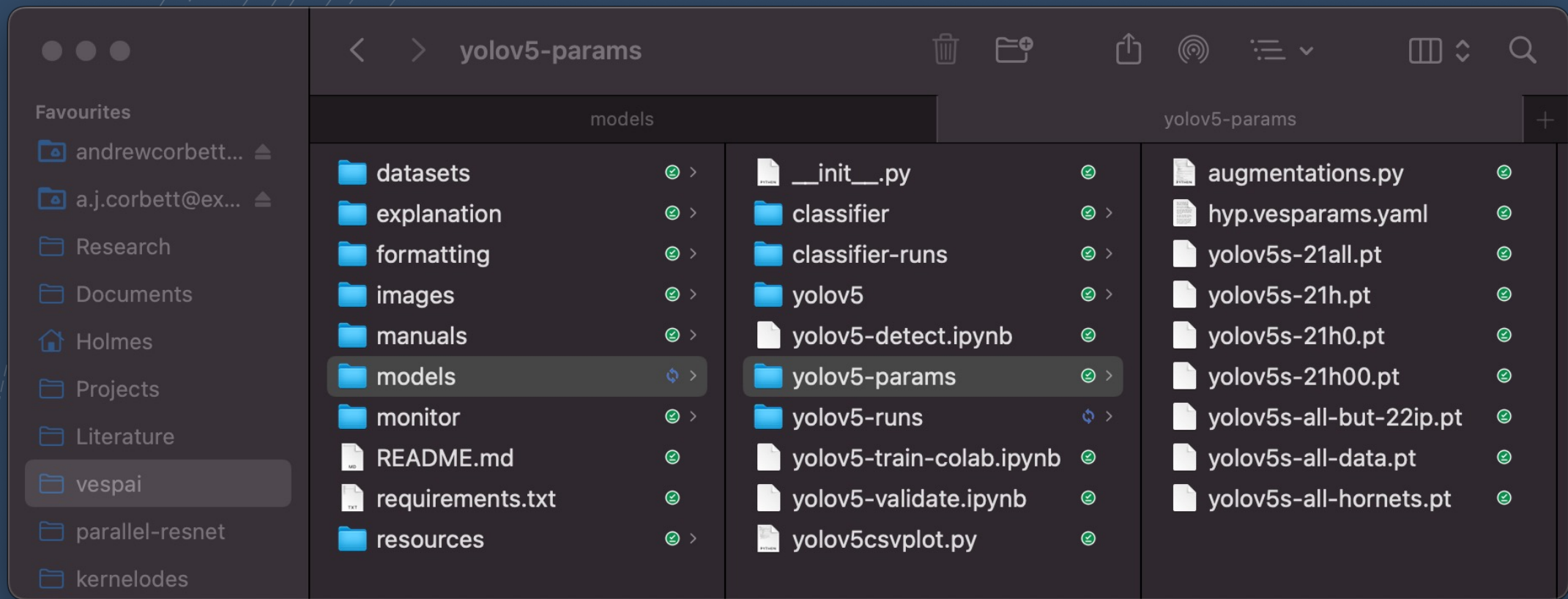
# Models Tested

## 2021 Data

- 21all: Sets 01-06; all 2021 data [2745 images]
- 21h: Sets 01-05; 2021 hornets, inc. the KQ [2147 images]
- 21h0: Sets 01-02; original 2021; TOW & PK [1698 images]
- 21h00: Set 04; updated 2021 TOW & E [285 images]
- 21ip: Set 06; 2021 insect pollutants [598 images]

## 2022 Data

- all-data: Sets 01-07; all labelled data [3302 images]
- all-but-22ip: Sets 01-06 + 08 (excl. 2022 IPs) [3178 images]
- all-hornets: Sets 01-05 + 08 [2580 images]



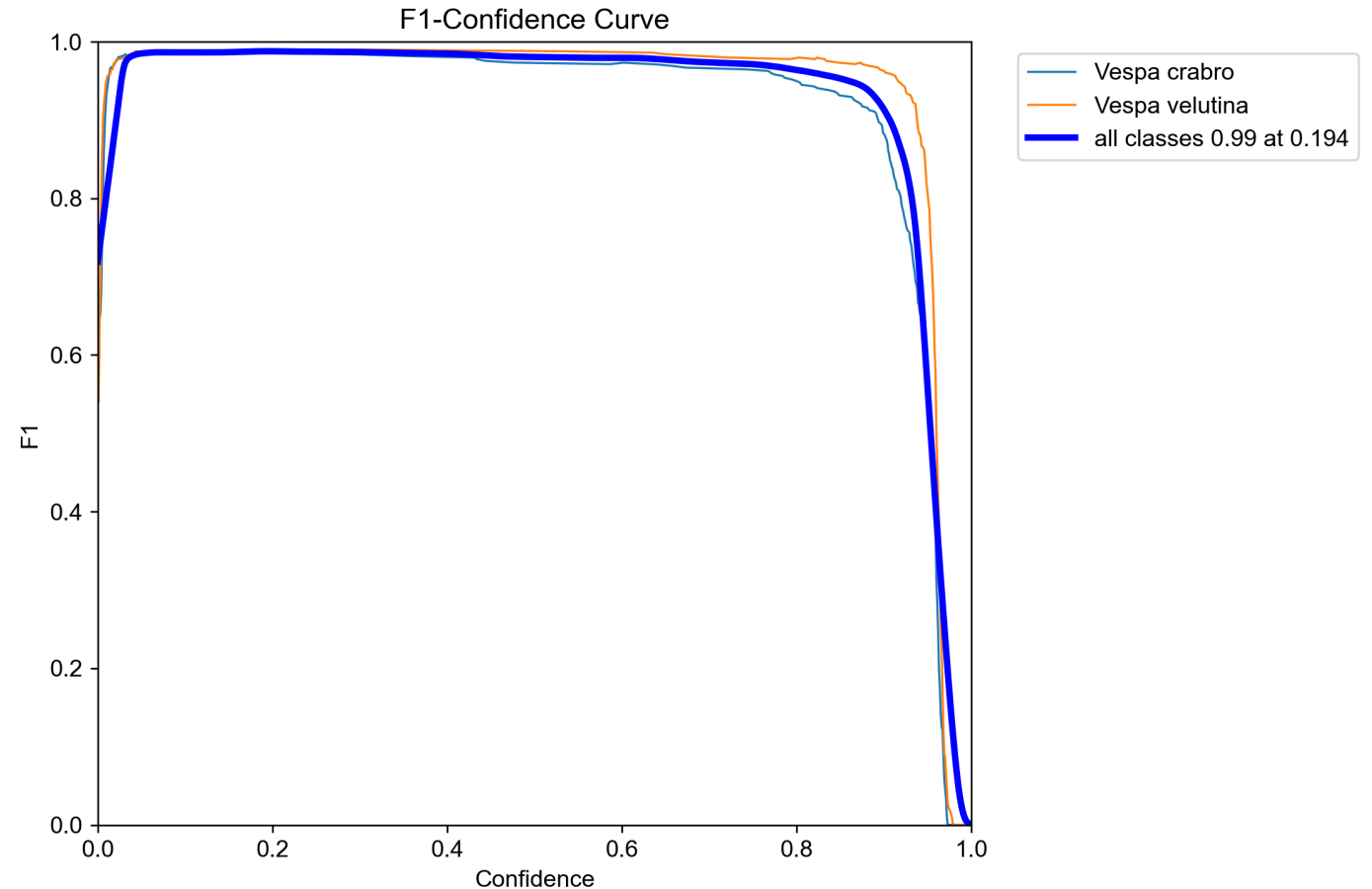
# Datasets

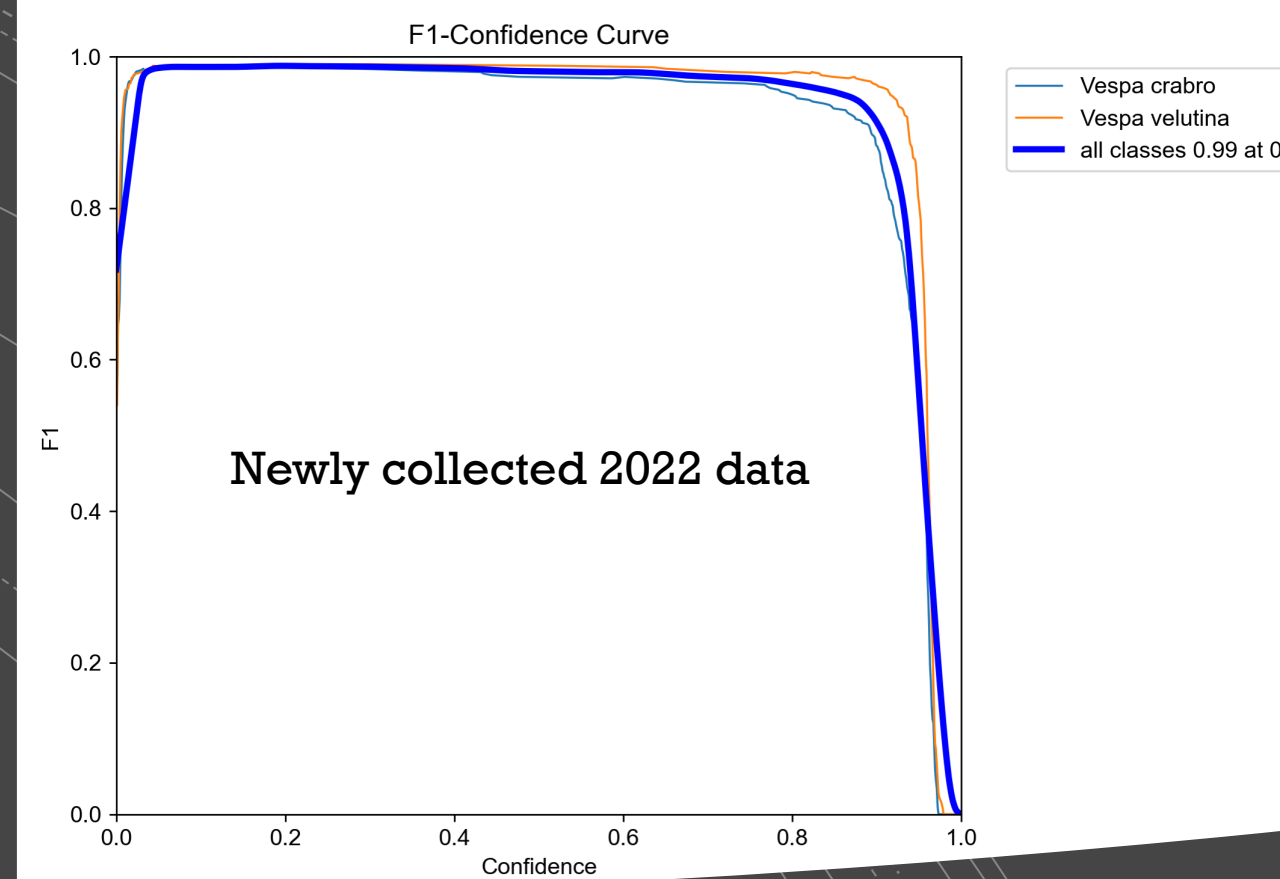
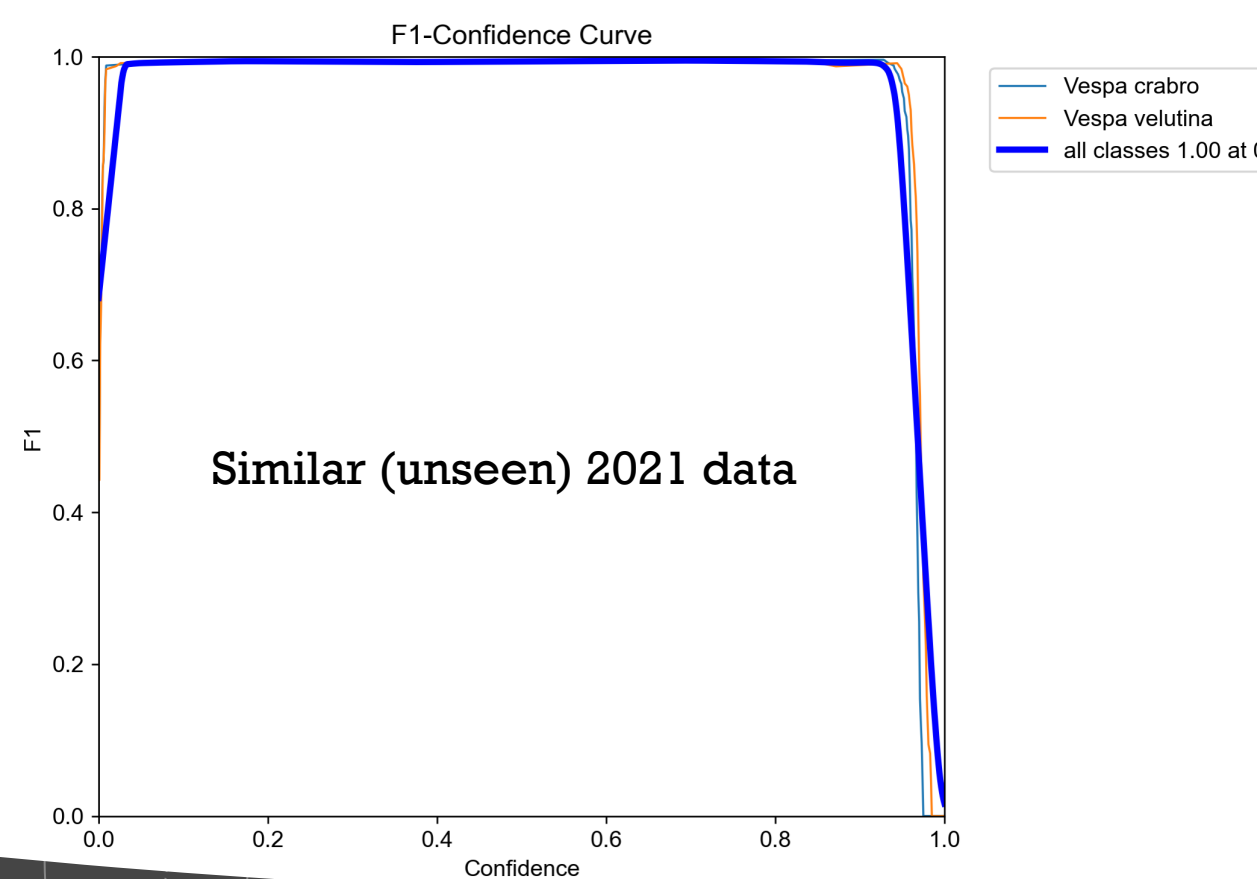
## 2022 Data

- 22all: Set 07 (= 08 + 09); all 2022 data [557 images]
- 22h: Set 08; all 2022 hornets [433 images]
- (22ip: Set 09; all 2022 insect pollutants [124 images])
- all-ip: Sets 06 + 09; 2021 & 2022 Ips [722 images]

“Models trained on 2021 data were strongly validated on 2022 testing data.”

Headlines





# Headlines

“Compare 2021 model performance on 2022 data vs. similar but unseen 2021 data.”

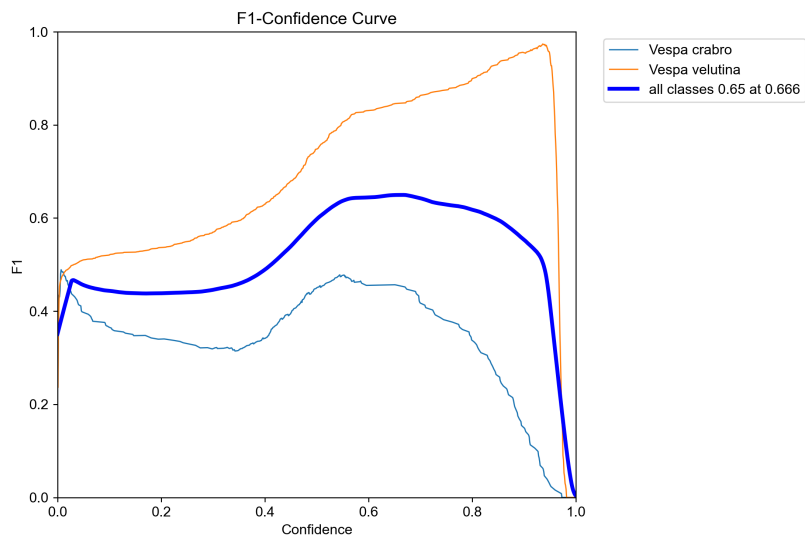
Improvement  
with more data

## 2021 Hornet-only trained models

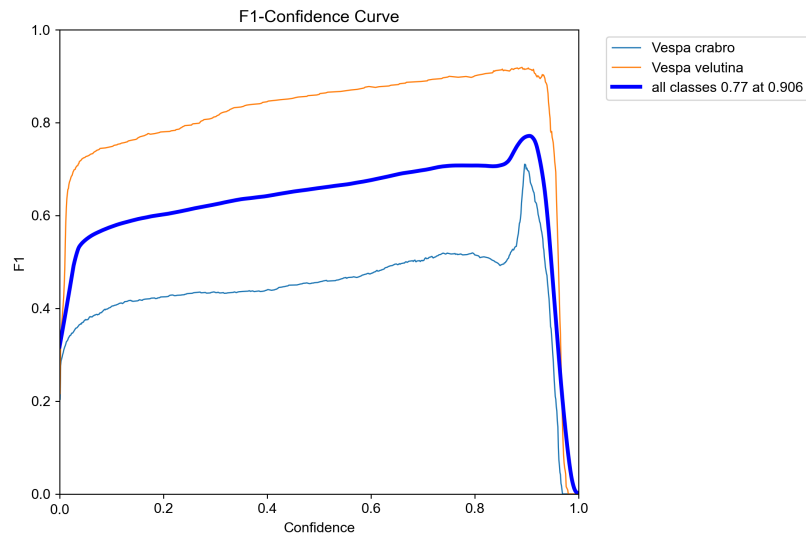
- 21h: Sets 01-05; 2021 hornets, inc. the KQ [2147 images]
- 21h0: Sets 01-02; original 2021; TOW & PK [1698 images]
- 21h00: Set 04; updated 2021 TOW & E [285 images]

“Success on 2022 data increases from 21h00->21h0->21h.”

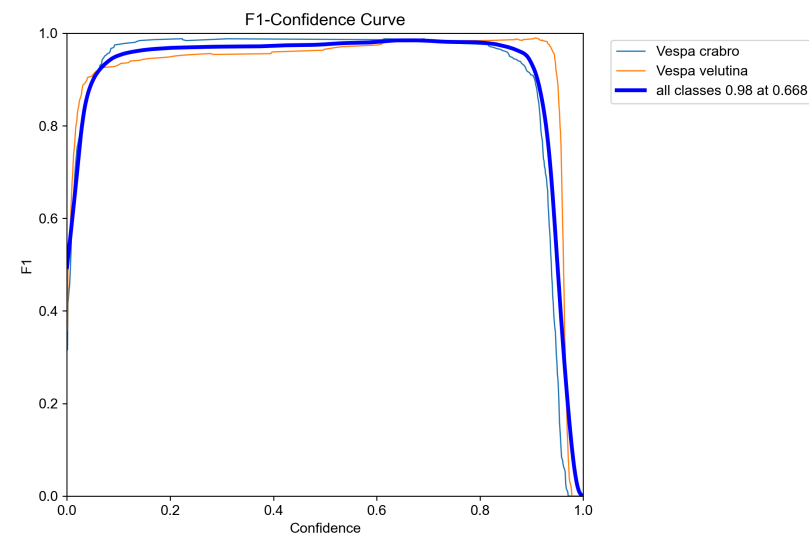




21h00 [285 images]

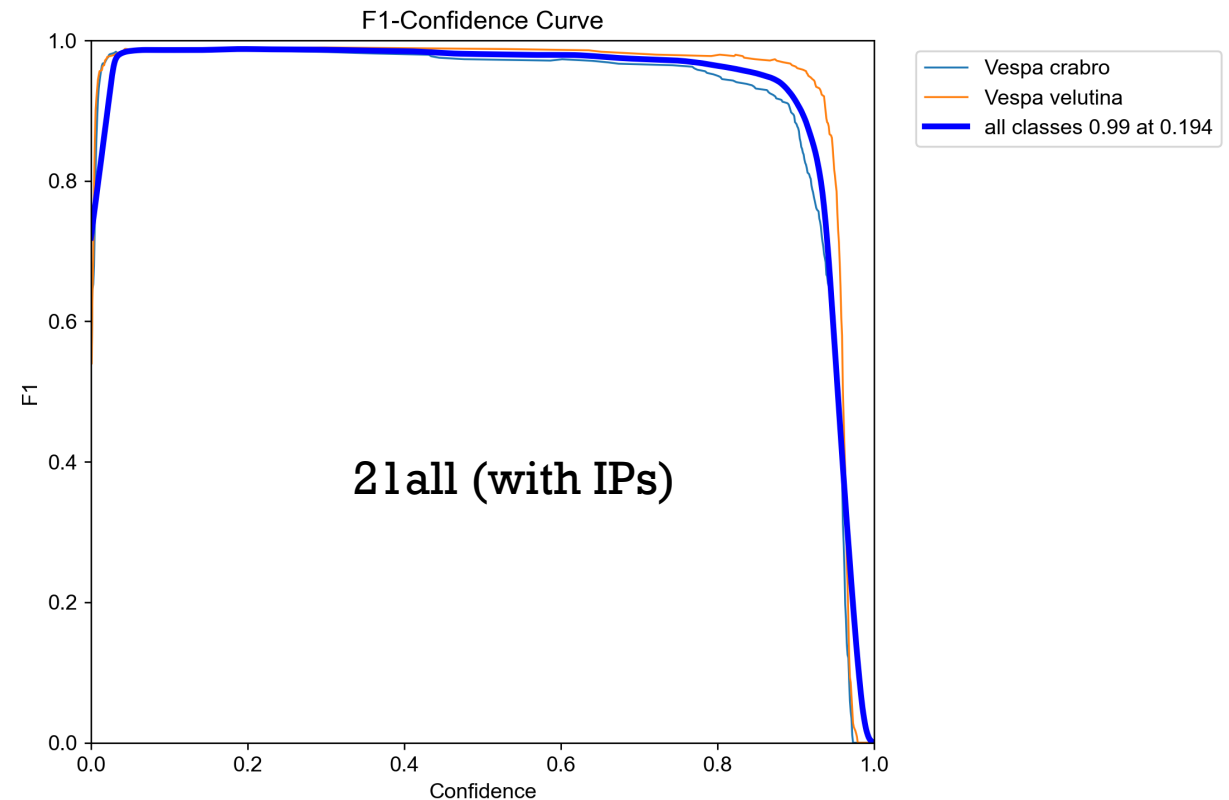
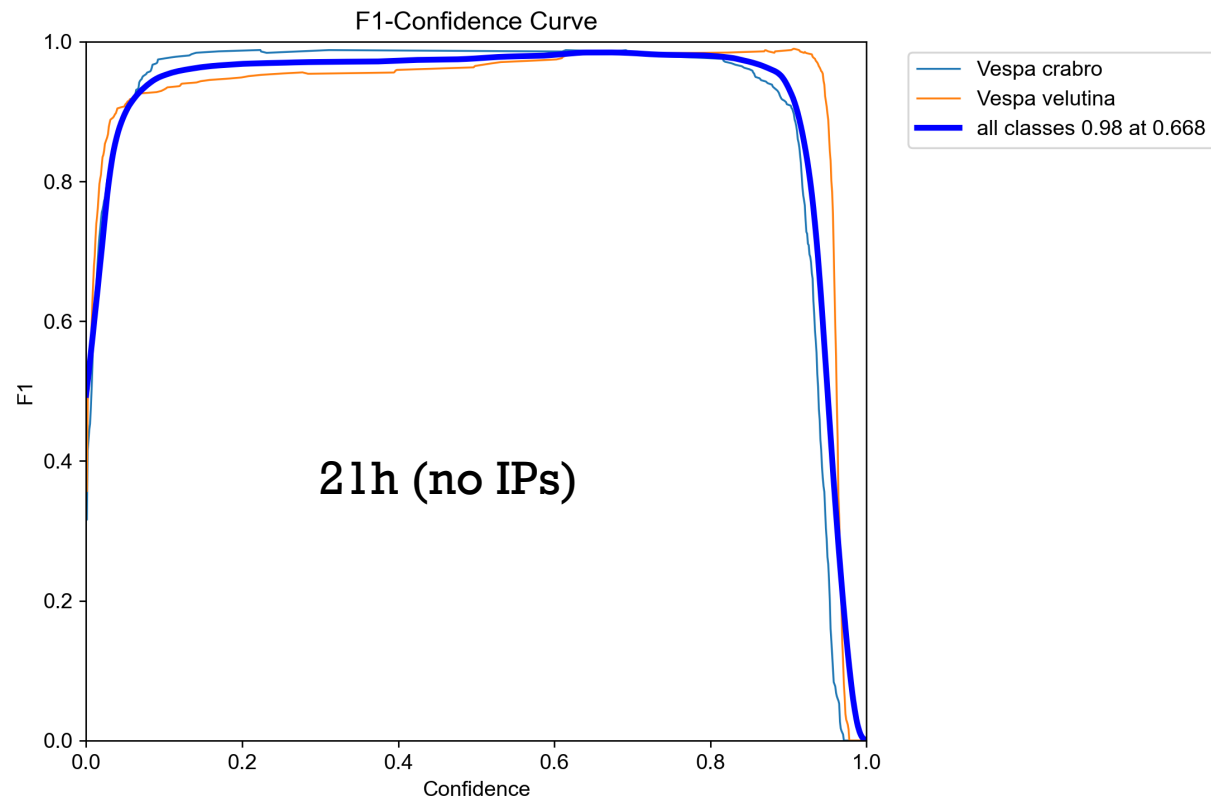


21h0 [2147 images]



21h [2147 images]

2021 Hornets-only on 2022 data



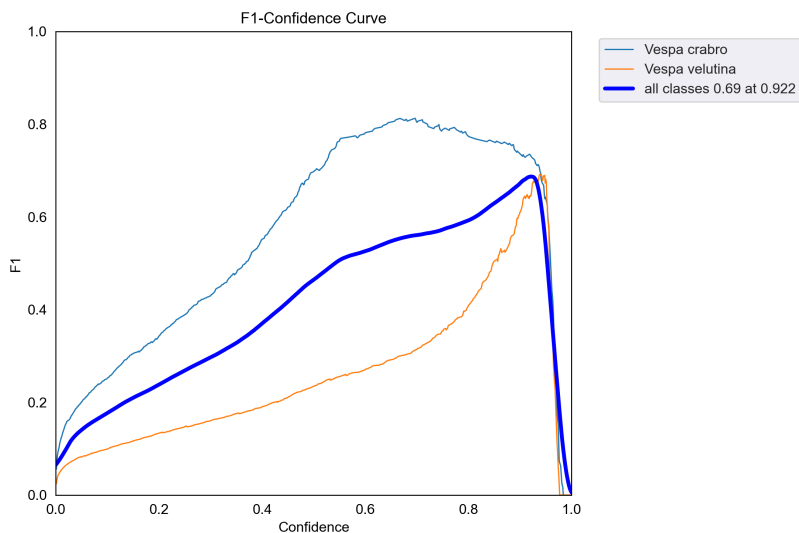
Inclusion of IPs in training data

“Including IPs in the training data set improves precision at lower thresholds, whilst seemingly reduces the model recall at high thresholds.”

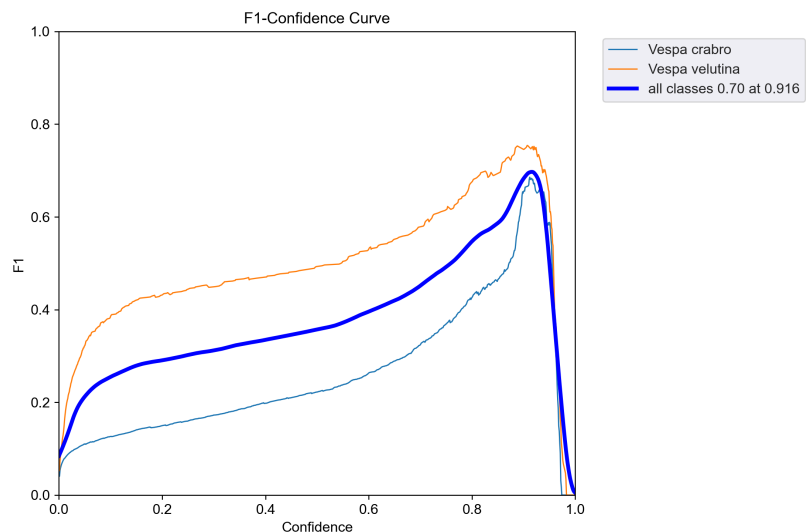
The background of the slide features several thin, curved lines in a light gray color, some solid and some dashed, creating a modern, abstract design.

## Insect Pollutants

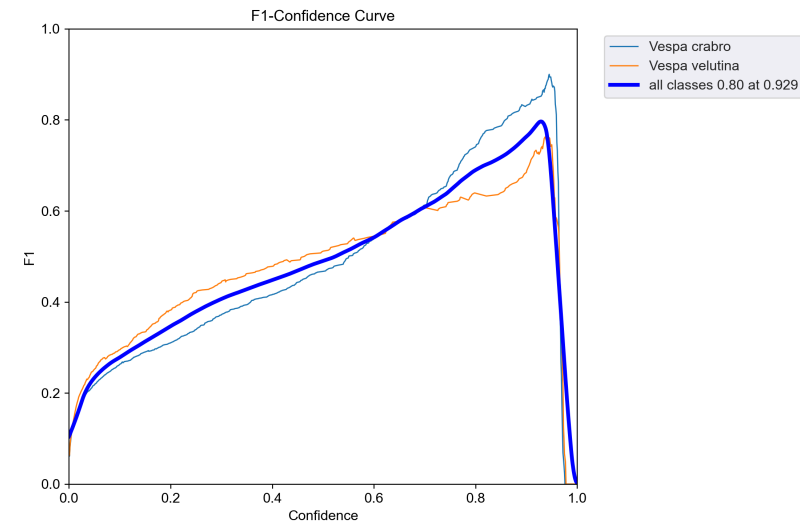
**“Let’s consider the performance of  
hornet-only trained 2021 models on  
unseen insect pollutant data.”**



21h00 [285 images]



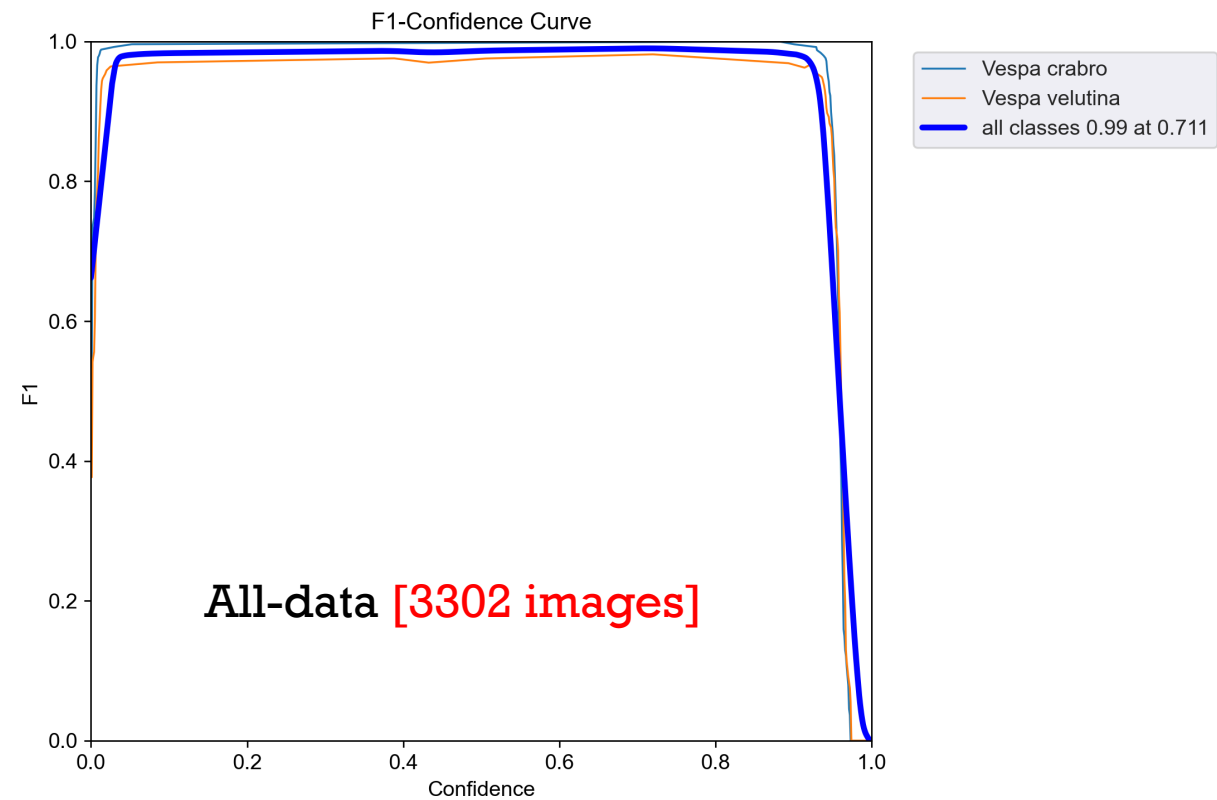
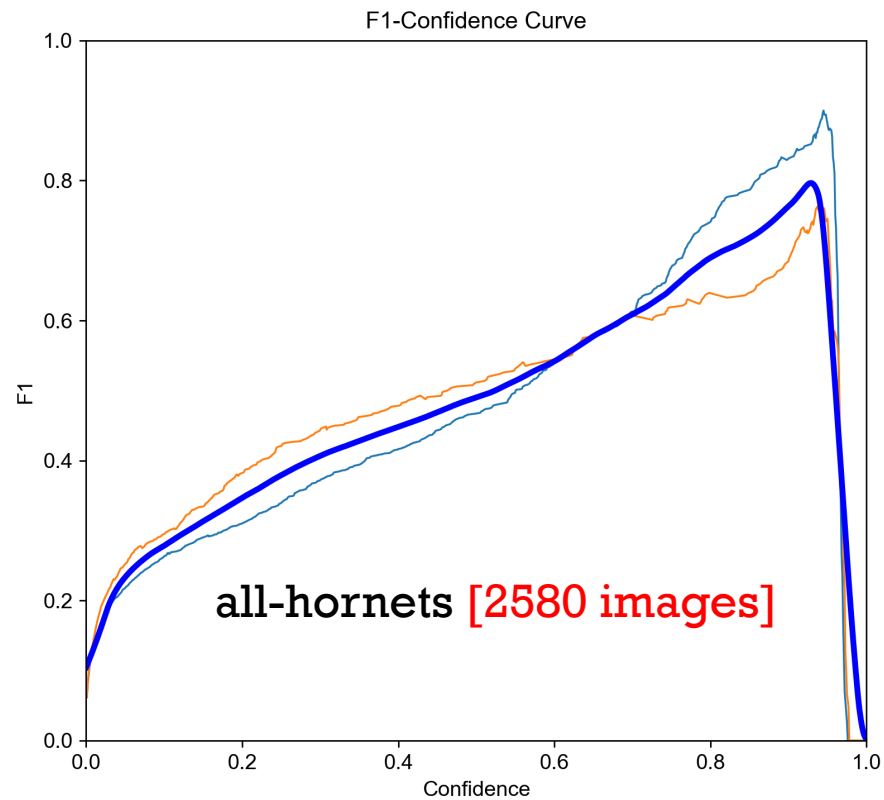
21h0 [2147 images]



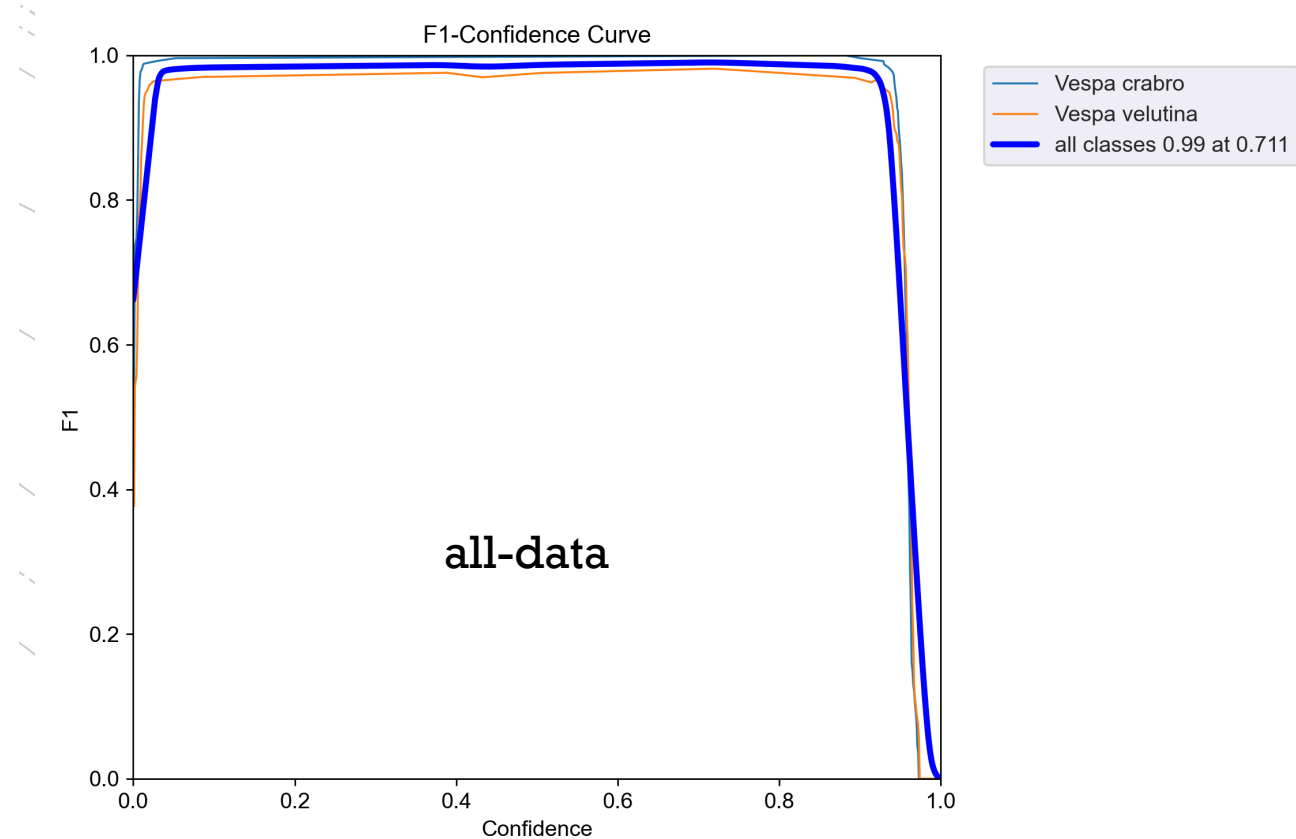
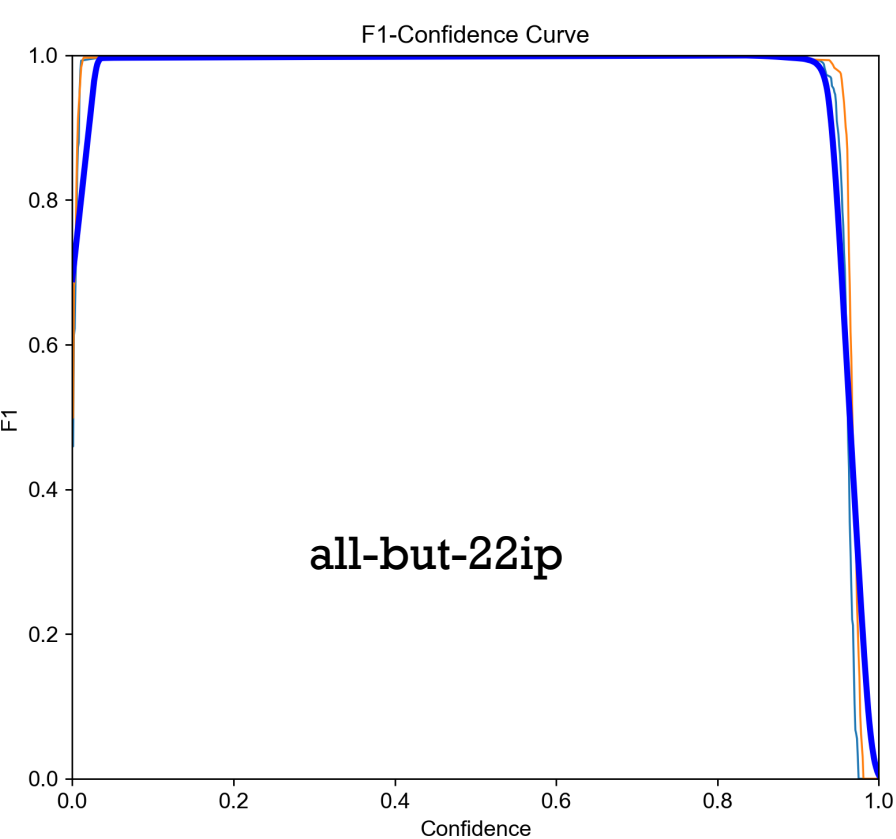
all-hornets [2580 images]

2021 Hornets-only on IP data

“Not very impressive.”



Throw in the IP data into training



Remove the 2022 Ips (mainly blanks/wasps)

“It seems as though the model improves upon removing 2022 IP data. However, the right-hand curve is a more realistic performance curve, indicating a better representation has been learnt.”

A red speech bubble graphic with a white outline, pointing downwards. The word "Gotchas" is written in white text inside the bubble.

Gotchas

**“A note on bait station design.”**



Model: 21h00



Model: 21h0

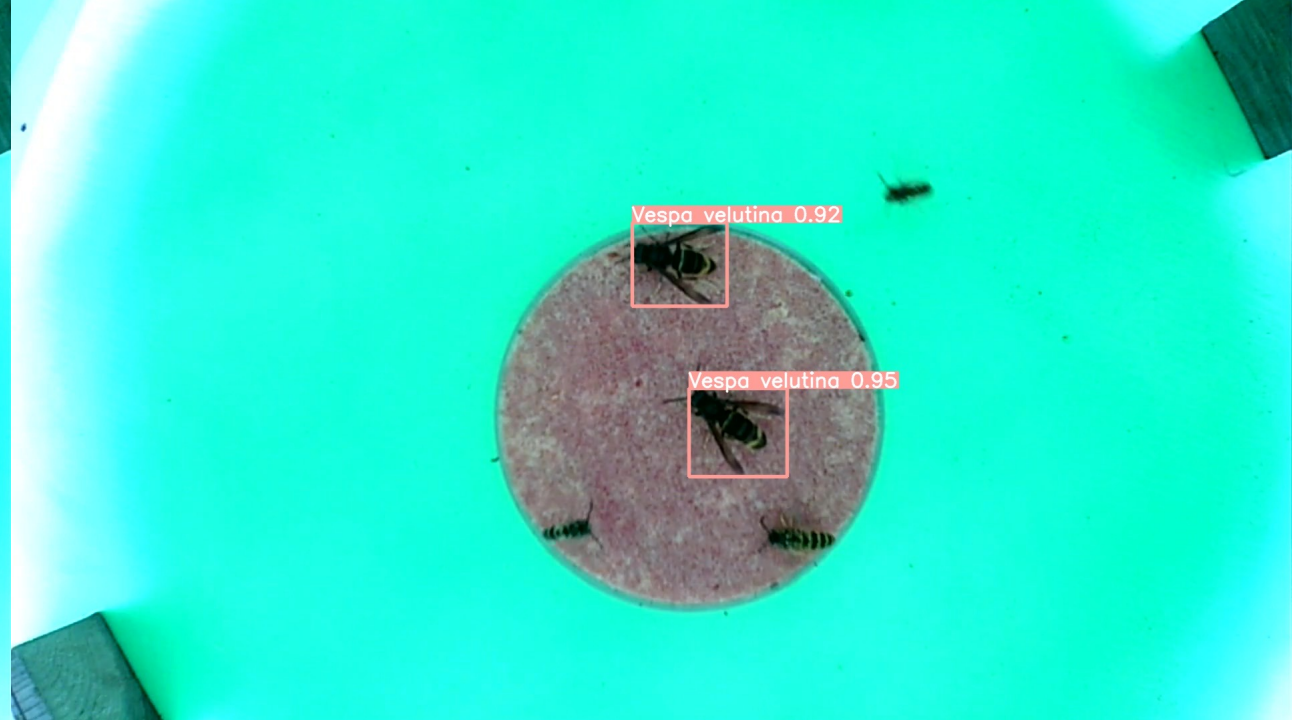
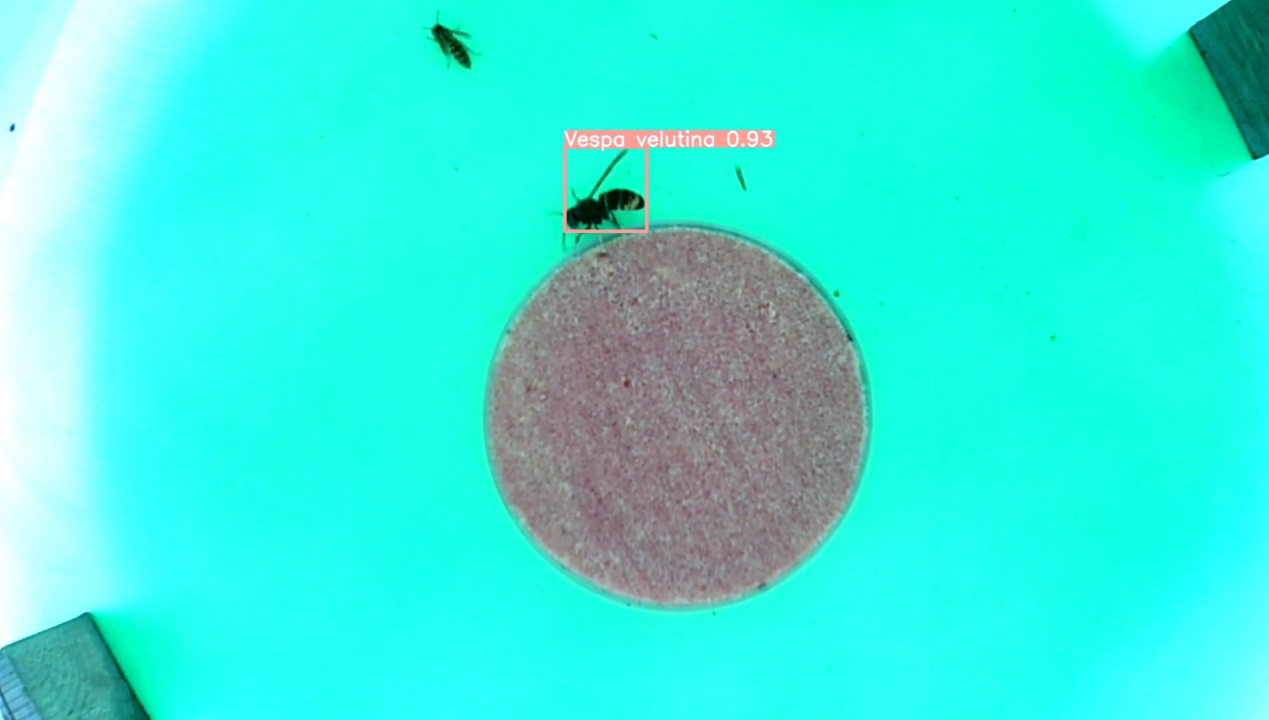
## Example of false positives

- We think the wooden posts are hornets
- None of the 'more data' models mistook this
- Train models on data from new bait stations

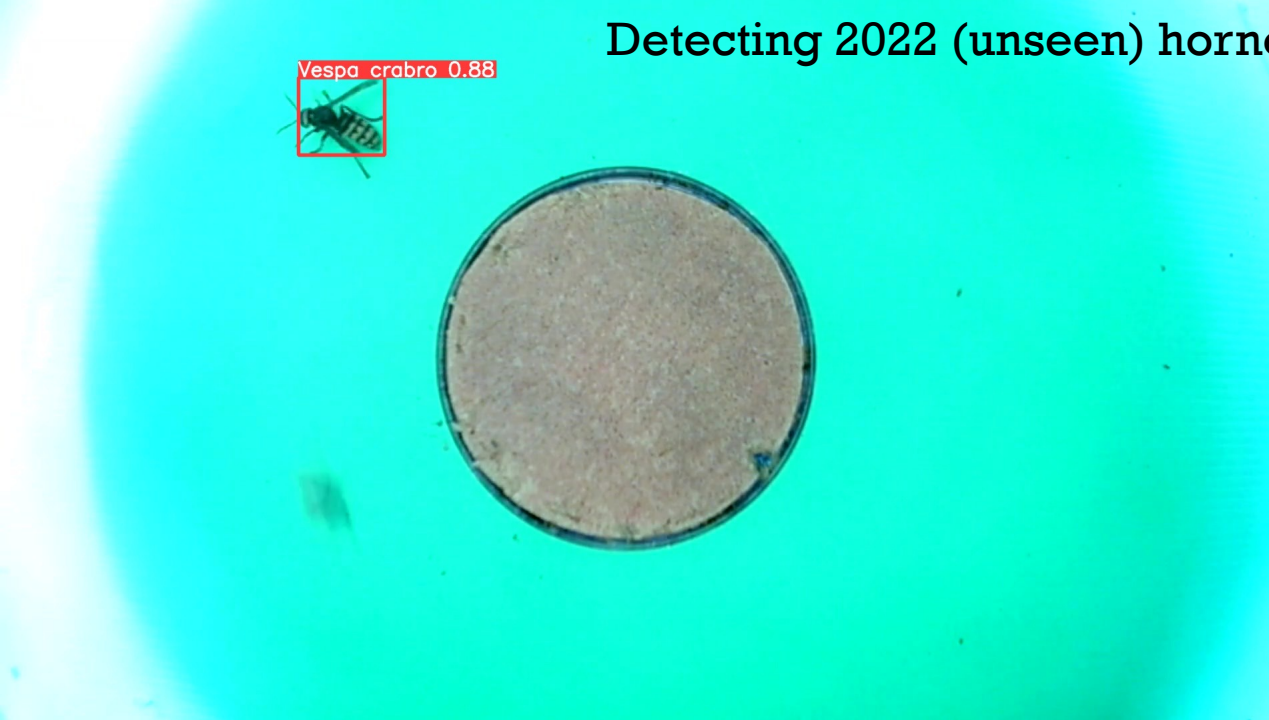




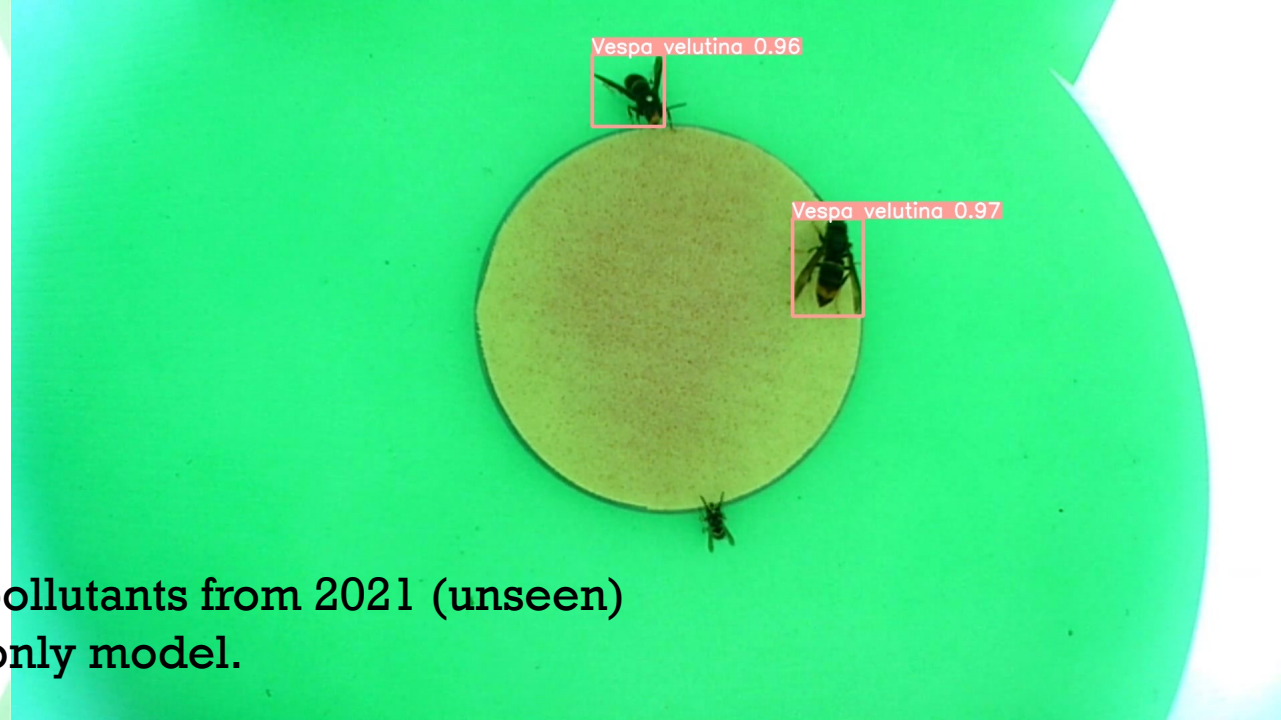
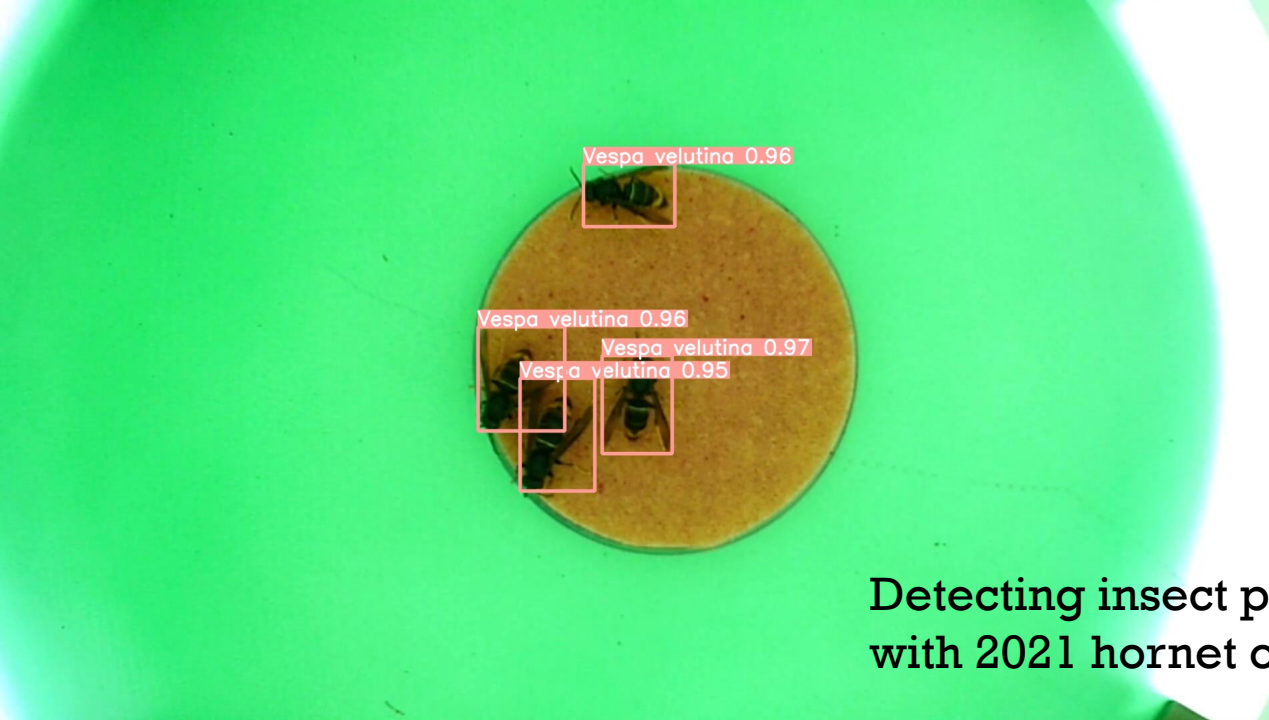
End on some good news



Detecting 2022 (unseen) hornets with a 2021 trained model.







Detecting insect pollutants from 2021 (unseen)  
with 2021 hornet only model.

