|  |  |  |
| --- | --- | --- |
| Taxa | Expected | Pipeline1 |
| A | 22 | 28 |
| B | 340 | 320 |
| C | 0 | 5 |
| D | 0 | 0 |

N expected = 362

N pipeline1 = 348

TP= predict yes, observed as yes

TN = predict no, observed no

FP = predict yes, observed no

FN = predict no, observed yes

[pipeline results, expected results]

A:

TP = 22 (22 times it was truly yes, and pipeline said we would see the yes)

TN = 0, (doesn’t contribute to the category seeing as this requires a no from expected and pipeline to satisfied, and there are observations)

FP = 6 (6 times the pipeline said yes when the expected was no)

FN = 0 (0 times was it expected to be seen and the pipeline said it was not seen

B:

TP = 320 (expected said yes and so did pipeline)

TN = 0 (not relevant)

FP = 0 (none of the pipeline predicted were not expected based off of what we know)

FN = 20 (it was supposed to be seen, but the pipeline said it shouldn’t be)

C:

TP = 0 (it is not supposed to ever be seen, so it can apply to this category)

TN = 0 (also 0 because pipeline predicted it and this category requires double no’s)

FP = 5 (5 times the pipeline said it should be there when expected says no)

FN = 0 (there were 0 times that expected said it should be there and the pipeline said no)

Option A: We do include these situations in the matrix

D:

TP = 0 (it was never observed so can not apply to this category)

TN = 1 (there was one “incident where the pipeline said it would not be seen and we were not expecting it to be seen)

FP = 0 (pipeline said no so this category cant apply)

FN = 0 (it was never expected to be seen so this category can not apply)

|  |  |  |  |
| --- | --- | --- | --- |
|  | Pipeline Yes | Pipeline No |  |
| Expected Yes | 22+320+0+0 = 342 | 0+20+0+0 = 20 | 362 |
| Expected No | 6+0+5+0 = 11 | 0+0+0+1 = 1 | 12 |
|  | 353 | 21 | N = 374 |

Columns and rows sum to same number, this is expected.

Option B: Ignore any taxa that were not expected and were not observed

|  |  |  |  |
| --- | --- | --- | --- |
|  | Pipeline Yes | Pipeline No |  |
| Expected Yes | 22+320+0 = 342 | 0+20+0 = 20 | 362 |
| Expected No | 6+0+5 = 11 | 0+0+0 = 0 | 11 |
|  | 353 | 20 | N = 373 |

Columns and rows sum to same number, this is expected.

I have no idea what these total N values represent or if they are useful/informative