Observation Insights Manual

Introduction

Observation Insights is a feature in the Spheer App that helps you create better, more reliable maps. In practice, we see that the quality of a model doesn't only depend on how many observations you add, but especially on which observations you add. Observation Insights makes this visible with a set of clearly interpretable observation quality indicators, expressed in values you can act on immediately. This lets you see more quickly where your model is strong and where small adjustments can yield big gains.

Building a good model is often an iterative process: adding observations, reviewing the result, adjusting, and retraining. Observation Insights gives structure to that process by providing signals about the data quality and coverage of your observations. This allows you to work more purposefully and increases the chance that the model produces a stable and reliable map.

What is the purpose of Observation Insights?

Observation Insights translates your observations into indicators. It helps you understand why a model performs well or less well, and which adjustment is likely to have the greatest effect.

How can you use it?

Use Observation Insights as a checklist while improving your model:

1. Start with an existing model.
2. Review the indicators and watch for warnings or unusual values.
3. Make targeted adjustments: each indicator points you to a specific type of issue in your observations.
4. Retrain and compare: see whether both the indicators and the map quality improve.
5. Repeat as needed until the map is stable and good for your purpose

What should you watch out for when using it?

Observation Insights is a tool, not an automatic "truth." Always keep thinking for yourself and take the following into account:

- **Always look at indicators together.** One indicator may look "good" while another points to a major problem. For example: a large surface area of observations is less effective if all the observations come from a single year or a single class dominates.
- **Observation quality is more than positive indicators.** In addition to the indicators we provide, qualitative aspects also matter, such as:
    - **Purity/consistency of labels:** are the observations truly representative of that class?
    - **Variation within a class:** does your class cover different situations (soil, moisture, management, season)?

    - **Representativeness:** do your observations resemble the areas you want to make predictions for later?

- **Green is not a guarantee.** If all signals show green, that doesn't mean the map is good. Observations can be technically fine but qualitatively impure or contradictory. It's also possible that certain types of nature are simply difficult for the model to distinguish.

So use Observation Insights as a compass: it points you in a direction, but you determine the course based on domain knowledge and the purpose of your map.

Can't figure it out? Our support team is happy to help (support@spheer.ai).

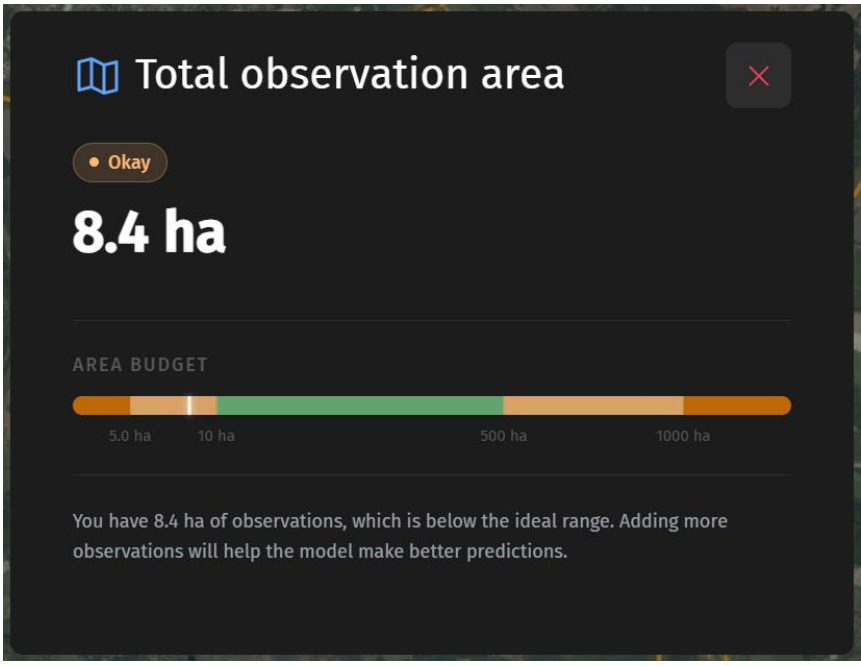
Table of Contents

Total Observation Area

What is it?

The total observation area is the sum of the area of all drawn observations. Observations that apply to multiple years are counted for each year.

Example:



In this example, the total area of the observations is 8.4 ha, placing it in the light orange zone. The green line runs from 10 to 500 ha, which is the area range in which this user model performs best. In the light orange zone, taking action can help the model; in the dark orange zone, acting is recommended.

What is the effect of this?

Spheer uses the area under the observations to learn from. It makes no difference to the model whether that area consists of many small observations or one large one.

- **Too little area** gives the model little information to learn from, which can make it struggle to predict accurately.
- **Too much area** can mean that new observations have little effect on the prediction, since there is little room left to adjust the model.

There is also a technical limit: if you exceed it, Spheer automatically selects a subset of your observation pixels.

How can you adjust this?

Increase a low score by adding more observations.

- **Tip:** use the completeness map for this, which shows where the model still needs extra observations to make a good prediction.

Lower a high score by removing or shrinking (redundant or duplicate) observations.

Als er goede redenen zijn om geen observaties te verkleinen of verwijderen (bijvoorbeeld bij een If there are good reasons not to shrink or remove observations (for example, in a use case with many different classes or many types of the same class that need to be drawn individually), you can choose to use a different "User Model." There are user models available that handle larger amounts of observation data better. For more information about user models, consult the manual or ask our support team for help.

How should you approach this indicator?

Use this indicator mainly as a rough guideline. Our models are good at producing maps from a small number of observations.

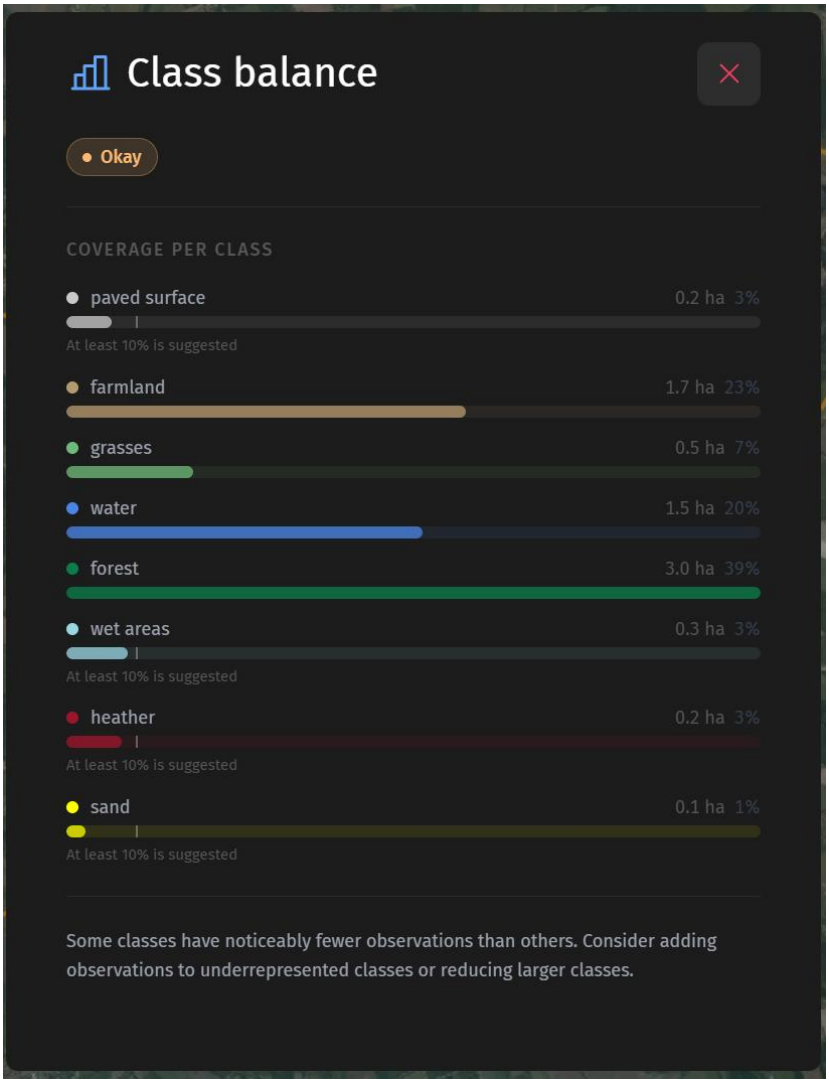
Class Balance (for classification models)

What is it?

Class balance refers to the ratio of (the area of) observations between the classes in the model.

Example:

In this example we see 8 classes. The balance shows that the class "forest" at 39% is much better represented than, for example, the class "sand(path)." This may be because this area has a lot of forest, or relatively little sand. As a result, the model can learn to recognize the "forest" class better and may have more difficulty with "sand."



What is the effect of this?

Spheer creates maps based on the information it is given. If one class has far fewer observations than another, the model learns to recognize the second class much better. This can lead to a lower-quality prediction map.

How can you adjust this?

Add observations to the class with fewer observations, or remove or shrink observations of the class with many observations.

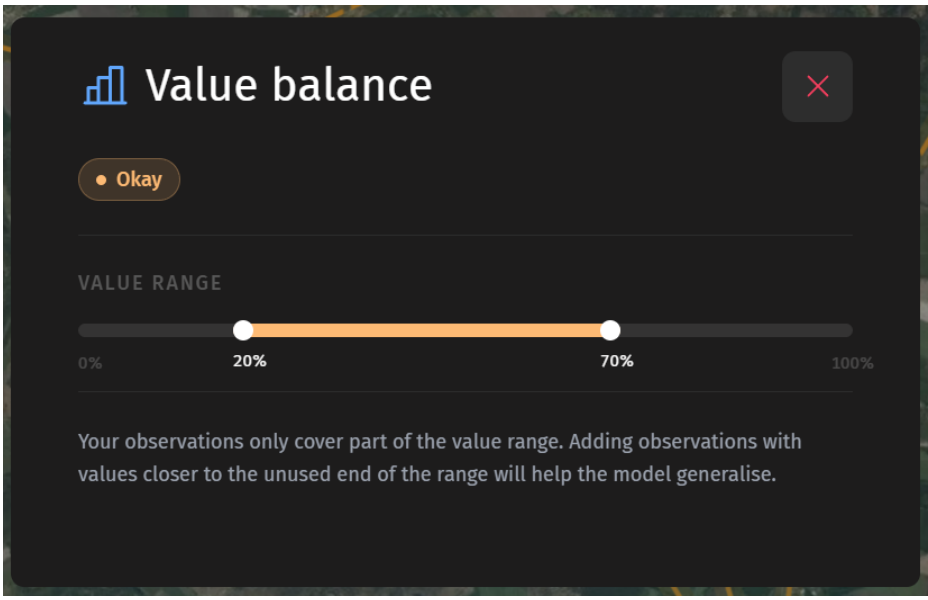
How should you approach this indicator?

It isn't necessary to balance the classes exactly; there are many good reasons to have more or fewer observations between classes. For example, if one class is genuinely much more prevalent than another, or if an "other/miscellaneous" class is used for which many different examples need to be drawn. That said, it can be useful to be aware of large differences. If there is an imbalance without good reason, it's worth adjusting.

Value Balance (for regression models)

What is it?

The value balance shows the value range of your regression observations. With this balance, you can see at a glance between which regression values all your observations fall.



Example:

In this example, the regression values fall between 20% and 70%. The model will mainly predict values that fall within this range.

What is the effect of this?

A regression model mainly predicts values within the value range, since these are the examples the model has seen. The balance lets you see at a glance what your range is, and therefore which regression values will be harder to predict.

How can you adjust this?

If your value range is missing certain values at the extremes, such as 0% and 100%, and you want to predict these too, you can add observations with these values.

If you only have regression values up to a certain level, you can also choose to normalize your range. You then convert your highest regression value to 100% and apply the same multiplication to your other values.

An example of this:

1. Your highest regression value is 50%. You convert this to 100% with the multiplication: 100%/50% = 2.
2. You apply the same multiplication to all other values, for example 0% and 30%:
    a. 0% × 2 = 0%
    b. 30% × 2 = 60%
3. This way, the full regression range can be utilized.

How should you approach this indicator?

If you are only interested in predicting values within the regression values you already have, or you already have a range between 0% and 100%, you can ignore this indicator.

If you are interested in values outside your range, add observations with those values to improve your predictions.

Temporal Coverage

What is it?

Temporal coverage is the distribution of observations across years. Years differ, and this affects nature. Certain types of nature can look different in a dry year than in a wet year. For a good comparison between years, it's important that the model sees examples of the different appearances of that nature.

Example:

In this example, we see a model with observations in every year. The temporal coverage indicator shows "Good," because there is an observation in each year.

What is the effect of this?

Temporal coverage indicates how good the comparison between years is. When observations are only added in a single year, it can be difficult for the model to make a good prediction in a different year. This happens when that year shows an unusual pattern — for example, it was much drier than other years — which gives a poor picture for other years. By adding observations across more years, stability across years improves.

We know from our own research that the biggest jump in accuracy comes from going from 1 year to 2 years, followed by going from 2 years to 3 years, with diminishing gains after that. If you want to make a comparison between years, we always recommend using at least 2 years.

How can you adjust this?

Add observations to the year with no or fewer observations, or remove observations from the year with many observations.

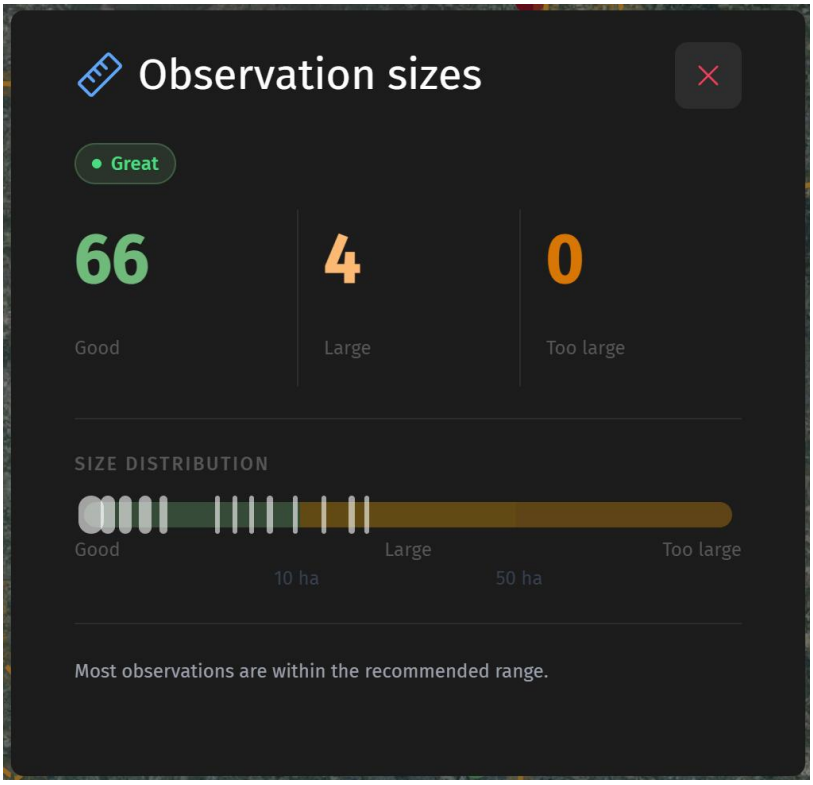
- **Tip:** In Spheer, it's easy to indicate that an observation applies to multiple years. You do this while editing an observation by clicking "select all years" in the year selection.

This means it isn't necessary to create unique observations for every year — in fact, it works well to reuse observations across different years so the model learns variations across years. When doing so, make sure the observation is also accurate for each year, and crop or, if needed, change the class of an observation accordingly.

How should you approach this indicator?

If you want to make a comparison across multiple years, make sure there are observations in multiple years. If you are only interested in the year in which you added the observations, you can pay less attention to this indicator, although in that case too, some temporal coverage can still improve the quality of the model.

Observation Sizes



What is it?

This indicator shows the area per individual observation. This lets you see at a glance how many observations have a good size, are large, or are very large.

Example:

In this example, we see a model with 70 observations in total, many of which are small in order to show enough variation within the classes. There are a few observations that are on the large side, shown in the light orange section.

What is the effect of this?

Very large observations often cover so much area that a lot of the information is redundant. More information isn't bad, but many observations can unnecessarily lengthen the model's training time. Large observations also increase the risk of noise or impurity within those observations.

As mentioned earlier, there is a technical limit: if you exceed it, Spheer automatically selects a subset of your observation pixels.

How can you adjust this?

Match the size of observations to the variation they contain. Large observations that repeatedly describe the same structure or the same type of area add little extra information: split them into multiple observations or make them smaller. Small observations are fine, but make sure that together they cover a wide range of conditions.

How should you approach this indicator?

Use this indicator in combination with the total observation area indicator for the best result.