

## Sampling Time Series Rasters

A common step in building species habitat models, characterizing fishing effort, and conducting other analyses involving point data is obtaining the values of images of remotely-sensed environmental variables for the points. For static variables such as bathymetry, this job is simple: you overlay the points on the image and extract the image values beneath the points. This procedure is called *sampling*.

Images of static variables can be easily sampled using the [Sample](#) and [Extract Values to Points](#) tools included in the ArcGIS [Spatial Analyst](#) extension. Both tools take a point feature class and a raster as inputs. The Sample tool outputs a table of the sampled values, while Extract Values to Points outputs a new point feature class with an additional column containing the sampled values.

But many variables are dynamic, with new images published monthly, daily or even hourly. These are called *time series data*. For these variables, the job is much harder: you must match up points and images by date, and then run one sampling operation for each date. For example, you might have observations of animals made over all the months of a three year period, and want to obtain the mean sea surface temperature (SST) for the month each observation was made. This would require 36 sample operations: one for each month of the three years of your study. Implementing this in the ArcGIS model builder using the Spatial Analyst tools requires significant cutting and pasting or some clever logic with the looping capability introduced in ArcGIS 9.2. In this example, we show how to quickly automate this procedure using the [MGET Sample Rasters Listed in Fields](#) tool.

### MGET can now sample certain popular products directly

Since we first wrote this example, we have developed a number of tools that can sample popular products directly. These tools completely automate the steps necessary to extract values of those products for a set of points. When possible, these tools interact with the original data provider's server using the OPeNDAP protocol, minimizing the bandwidth needed complete the sampling. When that is not possible, the tools automatically download, convert, and sample HDF or netCDF files from the provider's HTTP or FTP server.

Before you proceed with manually downloading, converting, and sampling the files, we suggest to check whether MGET already has a sampling tool designed for your product of interest. These tools are available under the Data Products node of the MGET toolbox. Look for tools beginning with the word Interpolate. Here are a few examples:

- [Interpolate AVHRR Pathfinder V5 SST at Points](#)
- [Interpolate NASA OceanColor L3 SMI Product at Points](#)
- [Interpolate PO.DAAC MODIS L3 SST at Points](#)
- [Interpolate Aviso SSH Product at Points](#)
- [Interpolate HYCOM GLBa0.08 Equatorial 4D Variables at Points](#)

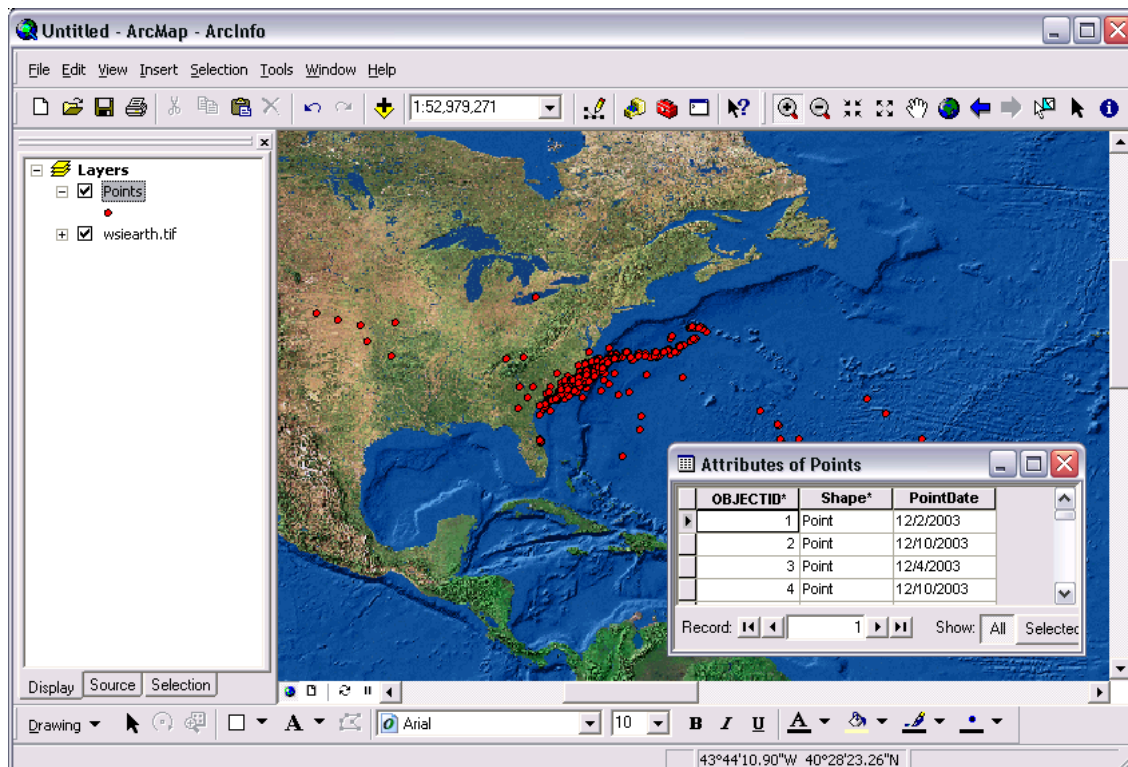
This list is incomplete, so of you do not see the product you're interested in, be sure you check the Data Products node before resorting to manually downloading, converting, and sampling your product. Assuming you need to do that, the rest of this page describes the procedure using AVHRR Pathfinder version 5 SST data (MGET can sample it directly, but we use it as an example because many people are familiar with it).

### The example scenario

In this example, we obtain daily SST values for point locations of sea turtles tracked off the east coast of the United States using the [Argos](#) telemetry system during the month of December, 2003. The turtles were fitted with radio transmitters that broadcast to satellites overhead when the turtles surfaced, and the satellites estimated and recorded the positions of the transmitters. Thanks to Catherine McClellan for making the data available for this example.

### The points

For this example, I deleted all of the fields in Catherine's original data other than **PointDate**, which is the date the turtle was observed at that location by Argos. The points' locations (stored by ArcGIS in the **Shape** field) and dates are the only attributes needed to perform the sampling. The points use the WGS 1984 geographic coordinate system.



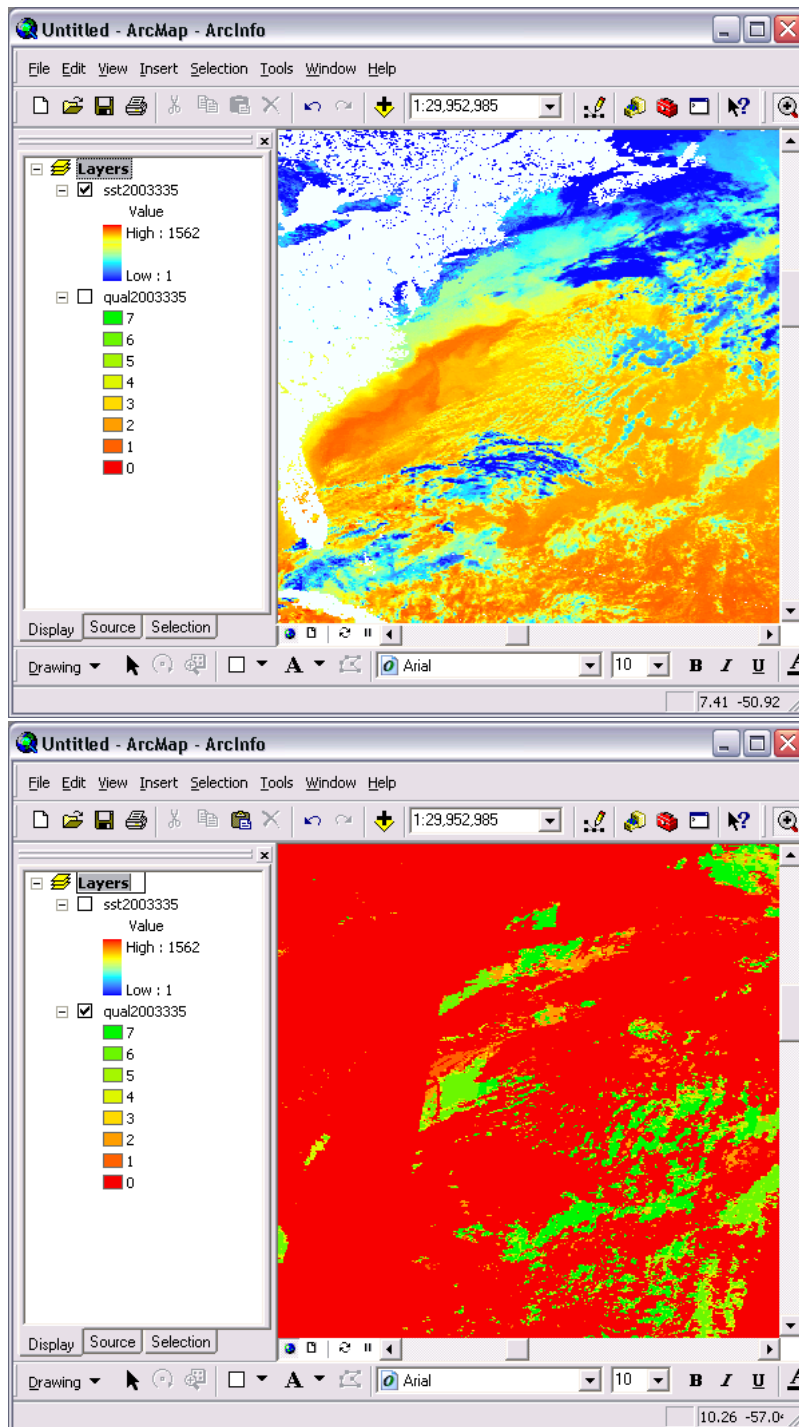
As you can see, some of the points occur on land or far out to sea. These are points where Argos failed to obtain an accurate position of the transmitter. Normally, when working with Argos data, you would "filter" these inaccurate locations from your dataset. For this example, we will leave them in so we can demonstrate what happens when we sample SST for pixels that have no data.

The sampling procedure presented here does not require the date field to be named **PointDate**, but it does require it have the ArcGIS data type DATE. If your date is stored as a TEXT field or some other data type, you must convert it to a DATE field. Contact us if you need help with this step.

### The SST rasters

The sampling procedure presented here assumes you already have SST data available in a raster format that ArcGIS can read. The data must be properly georeferenced, with a projection defined (although it need not be the same projection as your points). For this example, we will use the daytime daily SST from the [NOAA NODC 4 km AVHRR Pathfinder SST](#) dataset, which uses the WGS 1984 coordinate system. For more information about converting this dataset to ArcGIS format, please see the [Converting HDFs to ArcGIS rasters](#) example.

In addition to the **sst** variable, this dataset includes several other variables, including the **qual** variable, which assesses the likely accuracy of each pixel on an integer scale from 0 to 7, with 7 being the highest quality. Pixels that are cloudy or near the edge of the satellite's field of view are assigned low values. For example, the daytime daily sst and qual images for 1-Dec-2003 look like this off the east coast of the U.S.:



As you can see, the **qual** variable alleges that most pixels offer inaccurate SST estimates. When analyzing SST values sampled from satellite data, it is important to know whether a given value is accurate or not. To facilitate this, we will sample **qual** as well as **sst**.

Note that **qual** appears to be a very conservative assessment of accuracy, with very few pixels scoring high on the scale. This undoubtedly has implications for any post-sampling analysis, but that discussion is beyond the scope of this example. If you use this particular dataset in your analysis, I encourage you to read the documentation carefully and consider using the **msk1** and **msk2** variables, which offer a more detailed view of pixel accuracy than **qual**.

You may be surprised to see that the **sst** variable ranges from 1 to 1562. This is because the **sst** variable is published as 16-bit integers rather than 32-bit or 64-bit floating point numbers, to reduce the size of the HDF files. To obtain the actual SST value as a floating point number, you must put the integers through the *scaling equation* documented in the [4 km Pathfinder Version 5.0 User Guide](#):

$$\text{Parameter value} = \text{integer} \times \text{scale} + \text{offset}$$

$$\text{SST} = \text{integer} \times 0.075 - 3$$

There are several ways to accomplish this, including:

1. Before sampling, apply the scaling equation to the entire raster using Map Algebra. This will yield a raster that shows the actual SST values, but it will take up much more disk space. The [Converting HDFs to ArcGIS rasters](#) example shows how to do this with MGET as part of the HDF-to-raster conversion procedure.
2. Sample the integer values and then use the ArcGIS **Calculate Fields** tool to compute the actual SST values and store them in another field of the points.
3. Instruct the MGET sampling tool to apply the scaling equation during the sampling procedure, so that it stores the actual SST values rather than the integers.

This example illustrates the third option.

### Step 1: Add fields to the point feature class

To do the sampling, we must add four fields to the point feature class:

1. The **SST** field will contain the sampled SST, in degrees Celsius. This field should have the data type FLOAT.
2. The **Qual** field will contain the sampled **qual** variable. This field should have the data type SHORT, because it is an integer that ranges from 0 to 7.
3. The **SSTRaster** field will tell the MGET sampling tool which **qual** raster must be sampled. We will populate this field in step 2 below. This field should have the data type TEXT with a length of 250. (The default length of 50 might be too short, depending how deep in a directory structure we keep the rasters.)
4. The **QualRaster** field will tell the tool which **qual** raster must be sampled. As above, we will populate this in step 2, and the field should have the data type TEXT with a length of 250.

Here is a geoprocessing model showing this step:

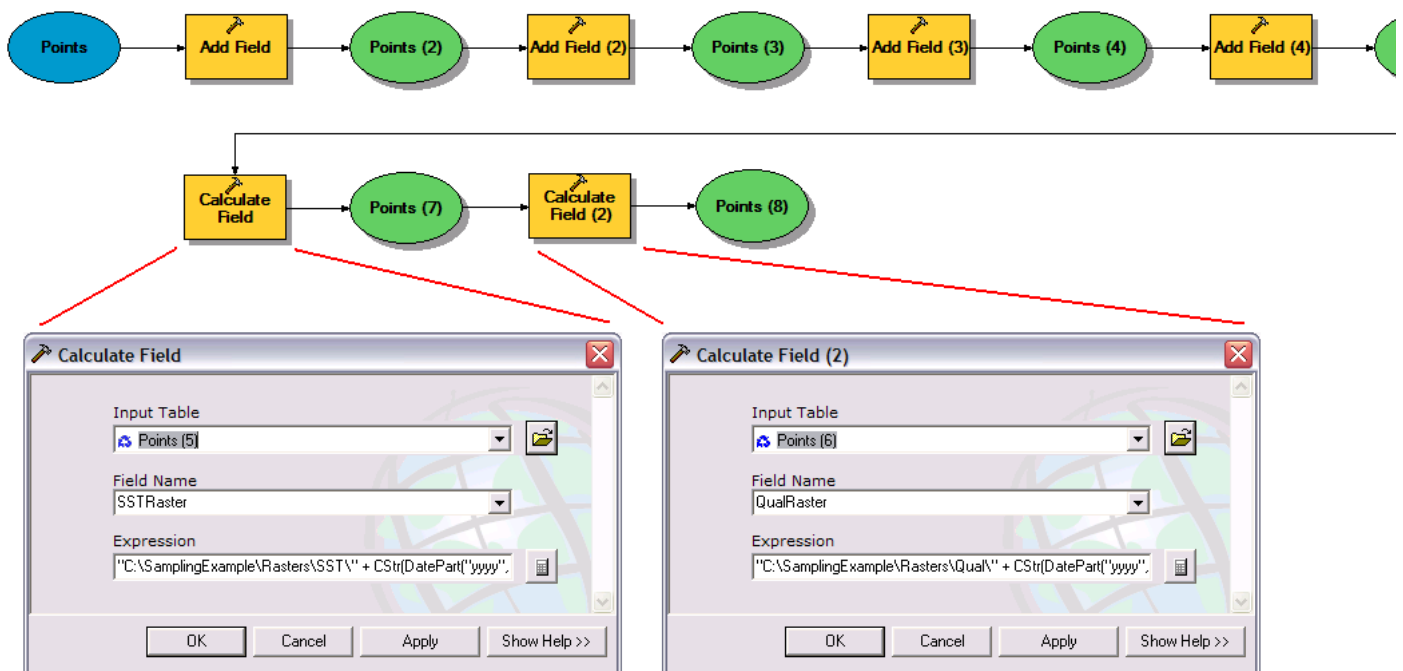


Attributes of Points						
OBJECTID*	Shape*	PointDate	SST	Qual	SSTRaster	QualRaster
1	Point	12/2/2003	<Null>	<Null>	<Null>	<Null>
2	Point	12/10/2003	<Null>	<Null>	<Null>	<Null>
3	Point	12/4/2003	<Null>	<Null>	<Null>	<Null>
4	Point	12/10/2003	<Null>	<Null>	<Null>	<Null>
5	Point	12/17/2003	<Null>	<Null>	<Null>	<Null>

To this:

Attributes of Points						
OBJECTID*	Shape*	PointDate	SST	Qual	SSTRaster	QualRaster
1	Point	12/2/2003	<Null>	<Null>	C:\SamplingExample\Rasters\SST\2003\sst2003336	C:\SamplingExample\Rasters\Qual\2003\qual2003336
2	Point	12/10/2003	<Null>	<Null>	C:\SamplingExample\Rasters\SST\2003\sst2003344	C:\SamplingExample\Rasters\Qual\2003\qual2003344
3	Point	12/4/2003	<Null>	<Null>	C:\SamplingExample\Rasters\SST\2003\sst2003338	C:\SamplingExample\Rasters\Qual\2003\qual2003338
4	Point	12/10/2003	<Null>	<Null>	C:\SamplingExample\Rasters\SST\2003\sst2003344	C:\SamplingExample\Rasters\Qual\2003\qual2003344
5	Point	12/17/2003	<Null>	<Null>	C:\SamplingExample\Rasters\SST\2003\sst2003351	C:\SamplingExample\Rasters\Qual\2003\qual2003351

This is the hardest step of the sampling workflow because it requires a bit of computer programming. Starting with a date, you must compute the full path to the raster for that date. The **Calculate Fields** tool built into ArcGIS provides this capability. Here is the updated geoprocessing model:



Note that this shows the ArcGIS 9.1 Calculate Fields tool. If you use 9.2 or 9.3, you should select Visual Basic as the programming language. Also, this does not show the full contents of the Expression fields. The full expressions are:

```
"C:\SamplingExample\Rasters\SST\" + CStr(DatePart("yyyy", [PointDate])) + "\sst" + CStr(DatePart("yyyy", [PointDate])) + Right("00" + CStr(DatePart("y", [PointDate])), 3)
```

and:

```
"C:\SamplingExample\Rasters\Qual\" + CStr(DatePart("yyyy", [PointDate])) + "\qual" + CStr(DatePart("yyyy", [PointDate])) + Right("00" + CStr(DatePart("y", [PointDate])), 3)
```

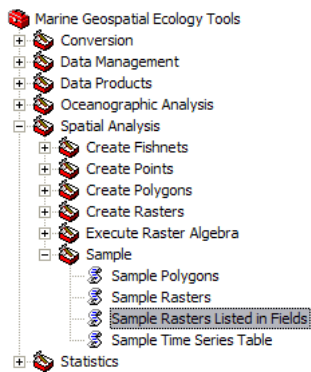
Writing these expressions can be a challenge. Here are some tips:

- DATE fields are represented in Calculate Fields as VBScript Date objects. You must manipulate them with special functions. The DatePart function returns a specified part of a date as an integer. The "yyyy" returns the four-digit year, while "y" returns the day of the year.
- Because you are populating a TEXT field, you must convert integers to strings before you can concatenate them together. Use the CStr function for this.
- When you pass "y" to DatePart, it returns an integer ranging from 1 to 366 (366 is December 31 on a leap year). To form the raster names from this example, you must convert the integers to the strings "001" to "366". If you convert the integer 1 to a string using CStr, you will get "1", not "001". As shown in the example above, you can fix this by prepend the leading zeros and using the Right function to extract the three right-most characters in the resulting string. This must be done because when "00" is prepended to "300", you get "00300", which you need to trim back down to "300".
- When writing these expressions from scratch, always start with the simplest possible piece of the expression and build up slowly from there. Try the first piece, run the tool, and verify that it works with no error and that you get what you want. Then append on the next piece and run it again. Continue building it up until you get your final expression.

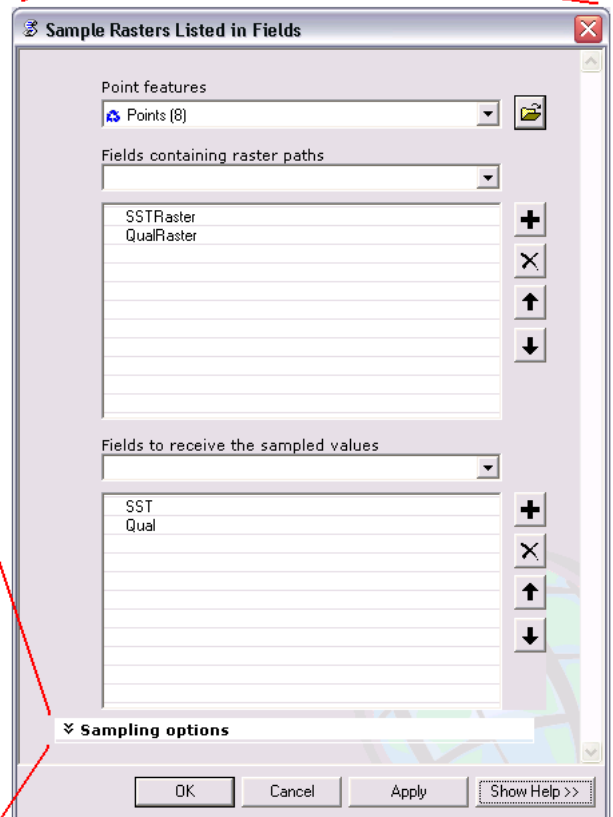
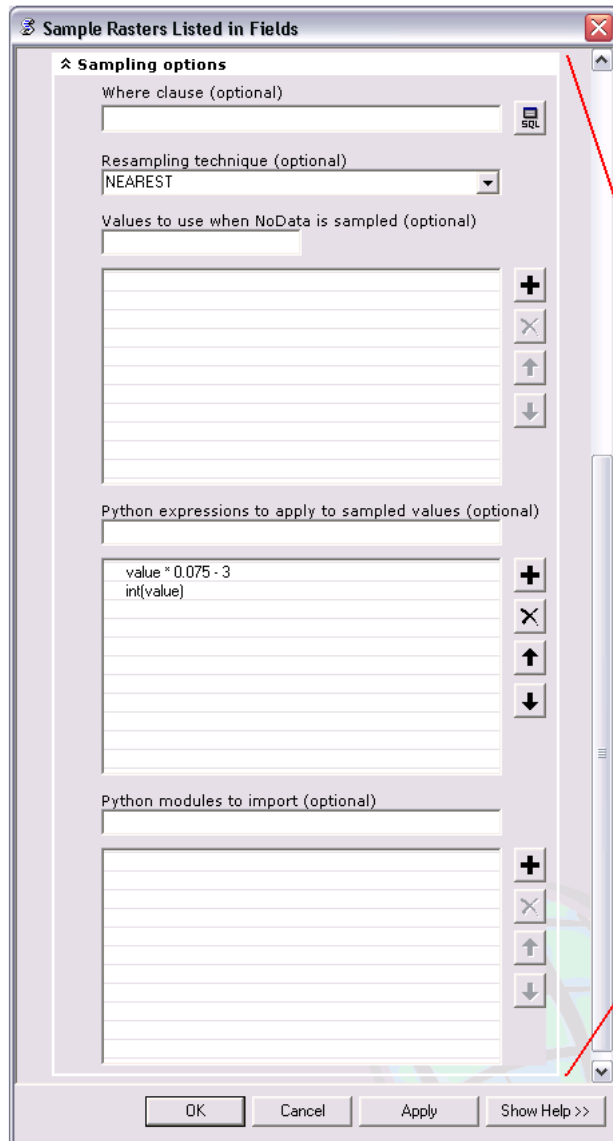
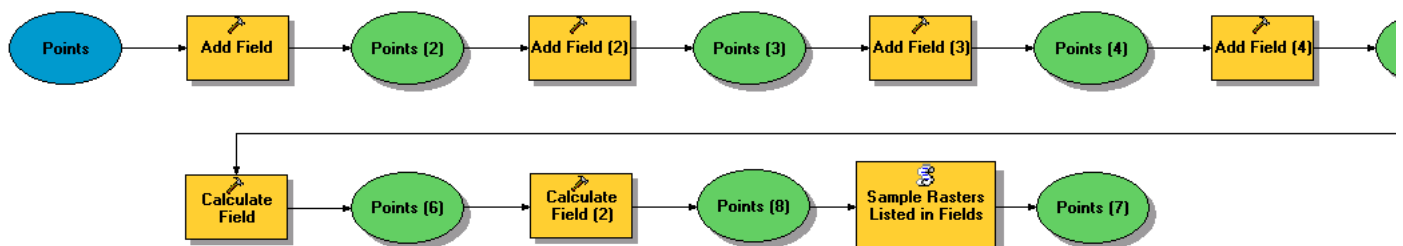
**Important:** After your expression seems to be working, you should spot-check a few of your points to make sure the proper raster path was calculated given the dates of the points. In complicated scenarios, when you are sampling using a temporal lag for example, it is easy to screw up the expression and end up sampling the wrong rasters. The MGET tool you'll use below will fail if you sample a raster that does not exist, but it cannot detect if you are sampling the *wrong* raster. It is up to you to ensure your expression results in the correct raster.

**Step 3: Run the MGET Sample Rasters Listed in Fields tool**

The **Sample Rasters Listed in Fields** tool is located in in the **Marine Geospatial Ecology Tools** toolbox:



As the main inputs to the tool, you provide the point feature class and specify the fields that contain paths to rasters to sample and the corresponding fields that should receive the sampled values:



For all of the points being sampled, the tool obtains the raster paths in the **Fields containing raster paths**, invokes the ArcGIS Spatial Analyst **Sample** tool to perform the sampling, and updates the **Fields to receive the sampled values**. The tool is smart enough to invoke **Sample** the minimum number of times needed to obtain values for all of the points. For example, if several points all reference the same rasters, **Sample** will only be invoked once.

**Sample Rasters Listed in Fields** also has a number of optional parameters. In this example, we leave all of these at their default settings except for **Python expressions to apply to sampled values**. Here, we provide an expression for each of the two types of raster we are sampling (SST and qual):

- For SST, the expression converts the sampled integer SST value (0 to 1562) to the actual SST value, in degrees Celsius, a floating-point number.
- For qual, the expression converts the sampled qual value to an integer so that it may be stored in the **Qual** field that has the **SHORT** ArcGIS data type. This odd maneuver is necessary because the Spatial Analyst **Sample** tool always returns sampled values as floating-point numbers, even if the sampled raster contains integers. So, even though the qual rasters contain the integers 0 to 7, the **Sample** tool returns the floating point numbers 0.0 to 7.0. Because the **Qual** field can only contain integers, we must convert the floating point numbers back to integers. This requirement is cumbersome. A future version of MGET's **Sample Rasters Listed in Fields** tool will automatically address this problem, but for now, you must explicitly convert the sampled values to integers, if your field has the **SHORT** or **LONG** data type.

**Important:** If you supply a Python expression **AND** you have MGET 0.6 alpha 2 (0.6a2) or earlier **AND** your point feature class is not a shapefile **AND** NoData is sampled for one or more points, you will receive an error unless you also provide a value for the **Values to use when NoData is sampled** parameter. This is a bug in MGET in 0.6 alpha 2, fixed in 0.6 alpha 3. All of these conditions must be true for the bug to occur. For 0.6 alpha 3 or later, you can leave the **Values to use when NoData is sampled** parameter blank, and **NULL** will be stored when NoData is sampled.

For this example, the sampled values look like this:



Attributes of Points						
OBJECTID*	Shape*	PointDate	SST	Qual	SSTRaster	QualRaster
1	Point	12/2/2003	19.05	0	C:\SamplingExample\Rasters\SST2003\sst2003336	C:\SamplingExample\Rasters\Qual2003\qual2003336
2	Point	12/10/2003	-2.925	0	C:\SamplingExample\Rasters\SST2003\sst2003344	C:\SamplingExample\Rasters\Qual2003\qual2003344
3	Point	12/4/2003	-2.925	0	C:\SamplingExample\Rasters\SST2003\sst2003338	C:\SamplingExample\Rasters\Qual2003\qual2003338
4	Point	12/10/2003	-2.925	0	C:\SamplingExample\Rasters\SST2003\sst2003344	C:\SamplingExample\Rasters\Qual2003\qual2003344
5	Point	12/17/2003	-2.925	0	C:\SamplingExample\Rasters\SST2003\sst2003351	C:\SamplingExample\Rasters\Qual2003\qual2003351
6	Point	12/29/2003	4.65	2	C:\SamplingExample\Rasters\SST2003\sst2003363	C:\SamplingExample\Rasters\Qual2003\qual2003363
7	Point	12/17/2003	-2.925	0	C:\SamplingExample\Rasters\SST2003\sst2003351	C:\SamplingExample\Rasters\Qual2003\qual2003351
8	Point	12/11/2003	13.95	0	C:\SamplingExample\Rasters\SST2003\sst2003345	C:\SamplingExample\Rasters\Qual2003\qual2003345
9	Point	12/13/2003	12	0	C:\SamplingExample\Rasters\SST2003\sst2003347	C:\SamplingExample\Rasters\Qual2003\qual2003347
10	Point	12/26/2003	<Null>	0	C:\SamplingExample\Rasters\SST2003\sst2003360	C:\SamplingExample\Rasters\Qual2003\qual2003360
11	Point	12/30/2003	-2.925	0	C:\SamplingExample\Rasters\SST2003\sst2003364	C:\SamplingExample\Rasters\Qual2003\qual2003364
12	Point	12/3/2003	-2.925	0	C:\SamplingExample\Rasters\SST2003\sst2003337	C:\SamplingExample\Rasters\Qual2003\qual2003337
13	Point	12/29/2003	12	0	C:\SamplingExample\Rasters\SST2003\sst2003363	C:\SamplingExample\Rasters\Qual2003\qual2003363
14	Point	12/2/2003	15.825	0	C:\SamplingExample\Rasters\SST2003\sst2003336	C:\SamplingExample\Rasters\Qual2003\qual2003336
15	Point	12/30/2003	24.825	7	C:\SamplingExample\Rasters\SST2003\sst2003364	C:\SamplingExample\Rasters\Qual2003\qual2003364
16	Point	12/29/2003	<Null>	0	C:\SamplingExample\Rasters\SST2003\sst2003363	C:\SamplingExample\Rasters\Qual2003\qual2003363
17	Point	12/3/2003	17.7	0	C:\SamplingExample\Rasters\SST2003\sst2003337	C:\SamplingExample\Rasters\Qual2003\qual2003337
18	Point	12/26/2003	<Null>	0	C:\SamplingExample\Rasters\SST2003\sst2003360	C:\SamplingExample\Rasters\Qual2003\qual2003360
19	Point	12/25/2003	18.675	0	C:\SamplingExample\Rasters\SST2003\sst2003359	C:\SamplingExample\Rasters\Qual2003\qual2003359
20	Point	12/9/2003	18.075	0	C:\SamplingExample\Rasters\SST2003\sst2003343	C:\SamplingExample\Rasters\Qual2003\qual2003343
21	Point	12/26/2003	<Null>	0	C:\SamplingExample\Rasters\SST2003\sst2003360	C:\SamplingExample\Rasters\Qual2003\qual2003360
22	Point	12/24/2003	<Null>	0	C:\SamplingExample\Rasters\SST2003\sst2003358	C:\SamplingExample\Rasters\Qual2003\qual2003358

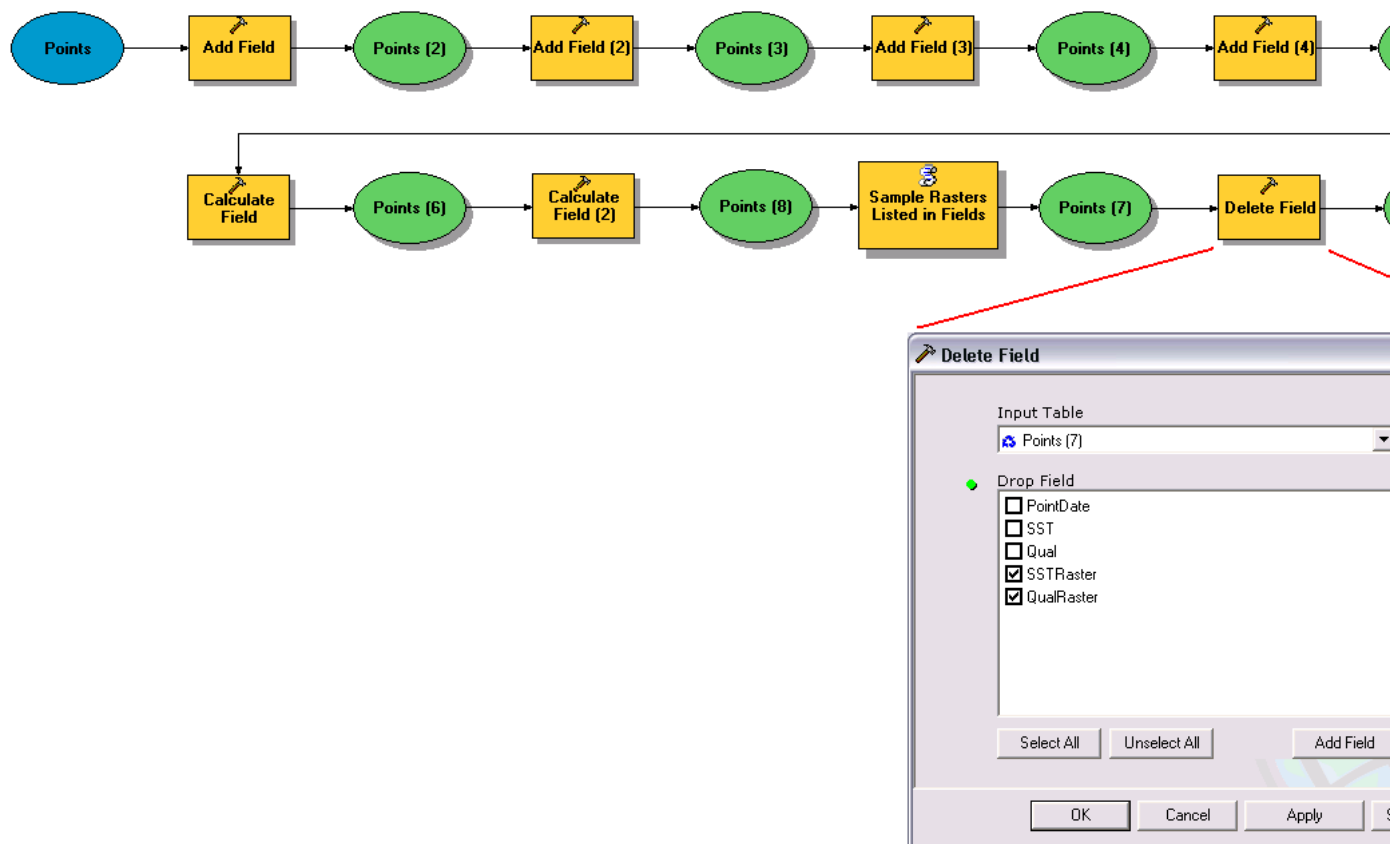
The **SST** field contains NULLs where the SST raster is NoData. In this particular dataset, the SST rasters are NoData only for land. The points with NULL are likely the ones where Argos obtained poor estimates of the turtles' locations.

The **Qual** field does not contain any NULLs because it has data everywhere, including land. You can see that nearly all of the points have a qual of 0. As discussed above, the dataset used for this example takes a very conservative approach to rating the quality of SST estimates. For many of these points with a qual of 0, the estimated SST is either NULL (for land) or -2.925, the minimum allowed value for this dataset. The -2.925 values are probably clouds. But a number of these points have reasonable SST values, such as 18.675, raising the possibility that qual is too conservative for some users. As discussed above, that analysis is beyond the scope of this example, and I encourage you to read the documentation on this particular SST dataset if you decide to use it.

**Important:** I strongly recommend you spot-check a few of your points to make sure the sampled values obtained by the tool match up with what you can obtain manually. To do this, use ArcMap to overlay your points on top of SST and qual rasters and, using the Identify tool, verify that the values of the rasters match what is stored in the points' fields. We do our best to ensure there are no bugs in the MGET sampling tool, but one user did discover a subtle bug that escaped our own tests in an earlier version of the tool. If he hadn't performed his spot checks, that bug might still not be fixed.

#### Step 4: (Optional) Delete the raster path fields

After you are sure that your sampled values are correct, there is no need to retain the **SSTRaster** and **QualRaster** fields in your feature class. If you have many points, these fields can take up a lot of space and slow down ArcGIS's performance. You can delete these using the ArcGIS **Delete Field** tool:



#### Future MGET plans



In the future, we will write a **Sample Time Series Rasters** tool, which would automate the workflow described in this example. We are eager to complete this so that users do not have to struggle with step #2, which is the trickiest part of the workflow. If you have any other suggestions, please do not hesitate to contact us.