Quality control at the Current Meter DAC

Current meter records submitted to CMDAC go through several steps before they are placed in the WOCE database. First, they are translated to the OSU format. The local format is a compact binary representation that also incorporates much of the metadata - such as mooring position, seafloor depth, the name of the PI, his/her institutional affiliation, etc. Since data almost always are submitted in ascii format, the translation usually involves a size reduction of 80% to 90%. And because no two PI's employ the same format, it always is necessary to write new software to effect the translation.

Once the data are available in CMDAC's format, they are examined by eye. The quality of the current records we see varies enormously. Anyone with much experience in this field understands that current meters, particularly when they are poorly maintained, do not necessarily present a faithful picture of the fluid medium in which they are placed. There are many failure modes. We have worked with these instruments for over 30 years, and are familiar with the difficulty of interpreting raw current meter data.

A few common problems:

- Sticking compass, revealed by the frequent recurrance, either contiguously or not, of a single direction value or a very narrow range of directions.
- Sticking encoder pins (Aanderaa RCM 4 amd 5): this produces spikes in all parameters, that differ from surrounding values by an additive power of 2.
- Rotorcounter failure, seen as a sudden drop in speed to zero or near-zero levels for several measurement cycles.
- Fouling of the speed sensor, or failure of its bearings, indicated by a gradual reduction of speed.
- Tape glitches, causing spikes in all parameters (for current meters that record on magnetic tape).
- Sensor drift - a change in the response of the sensor. This is hard to distinguish from a true change in ambient conditions, but occasionally is the best explanation for an observed trend.

In general, the most difficult parameter is speed. Perhaps because the sensor is mechanical, most of the problems occur here. If we find a sharp spike or other type of very sudden change in speed we look for a plausible ambient cause, such as tidal oscillations. If no such cause is found, then the spike may be an error. The ability of a fluid in motion to exhibit sudden changes in momentum is limited. A natural explanation for speed spikes becomes less probable as the averaging interval increases in length.

Some of the records we receive are quite clean; in other cases it is clear that the originator either is inexperienced or lacks the resources to clean up the dataset. In some problematical cases we prepare an alternate file that more nearly represents what we believe a problem-free current meter would have recorded. The goal is to provide users of CMDAC's database with current records that they can use with confidence - that are less likely to lead to false conclusions about what happened in the ocean.

At this step we utilize an application that displays each time series on the computer screen in segments several days long. This program allows us to select specific data points or data segments and either remove them or replace them by interpolation. A single-point spike that

has been identified as a probable error will be replaced by linear interpolation. With longer segments we have the choice of linear interpolation or predictive interpolation.

Generally, if a bad segment is a few hours to a few days in length, we will use predictive interpolation. This technique utilizes an algorithm based on the maximum entropy method of analysis. The chief advantage of predictive interpolation is that it introduces no contamination into the spectral makeup of the time series. A linear interpolation does alter the spectrum of the series; for this reason we sometimes insert predictive interpolations in place of linear interpolations supplied by the PI.

The chief disadvantage of a predictive interpolation, as opposed to linear interpolation, is that the data user may not be aware of the interpolated material, since it looks just like the surrounding data. We have attempted to mitigate this by providing a comment file with each modified current record that clearly describes the changes we made. In most cases where interpolations were inserted, the exact location of the interpolation is given.

Predictive interpolations longer than a few days are problematical. The longer the data gap, the greater the liklihood that the ocean did something during the gap that the technique cannot anticipate. Because of this, when a stretch of clearly bad data is longer than a week or so, and we have decided to make an alternate file, we simply remove the bad segment and leave a gap in the time series.

The decision to make an alternate file hinges on whether we believe we can produce a current record that is "truer" than the record that the PI provided. A time series is truer when we have either deleted obvious errors or replaced them by interpolation. In some cases a time series is so poor that that there is nothing we can do to rescue it. Such a series will either be excluded from the alternate file, or drastically reduced in length. In a very few cases a current meter record has been sufficiently poor that the entire record - all of the time series in it - has been excluded from the WOCE database.