Recap and Plans for ITS

2013

Feasibility Study Starts

1. Goal: to study the feasibility of integrating the three time series efforts:
    a. Increase the scientific and technological contributions
    b. Decreasing the required resources (maintenance of time series).
    c. Is it possible to combine the data
    d. Identify critical measurements from each time-series that should be sustained in the future
    e. Identify geographic locations, existing or future, that would best suit continued time series studies
    f. Identify critical instrumentation that could be common to each time-series
    g. Identify new instrumentation in concert with the Technology Roadmap that could be deployed at each time series site to increase data collection efficiency while decreasing demain on ship time
    h. Identify external collaborators for formal involvement in integrate time series studies that could contribute expertise not present within MBARI
    i. Write white-paper document summarizing the results of our integration efforts
2. Time series involved:
    a. Monterey Bay Time Series
    b. Midwater Time Series
    c. Pelagic-Benthic Coupling project

2014

Feasibility Continues and Leads to 5 Year Proposal

Results to Date:

1. Agree to continue to occupy the two sites (M1/Midwater and Station M)
2. Merged data sets show trends over 24 years
3. Prototype pass at technology for data integration (extract, transform and load from BOG, VARS/Midwater, Benthic Excel sheet)
4. Generated table of measurements (parameters and frequency) and instrumentation

Goals:

1. Integrate measurements from table across 3 time series
2. Identify platforms and sensors for the future
3. More exploration and prototyping of data integration technologies and architectures
4. White paper:
    a. Describe integration of time series data
    b. Demonstrate why this is a crucial effort
    c. Plan to submit to journal
5. Develop the integrated time series ‘study’ with goal of export
6. Conduct synoptic pilot survey with LRAUV

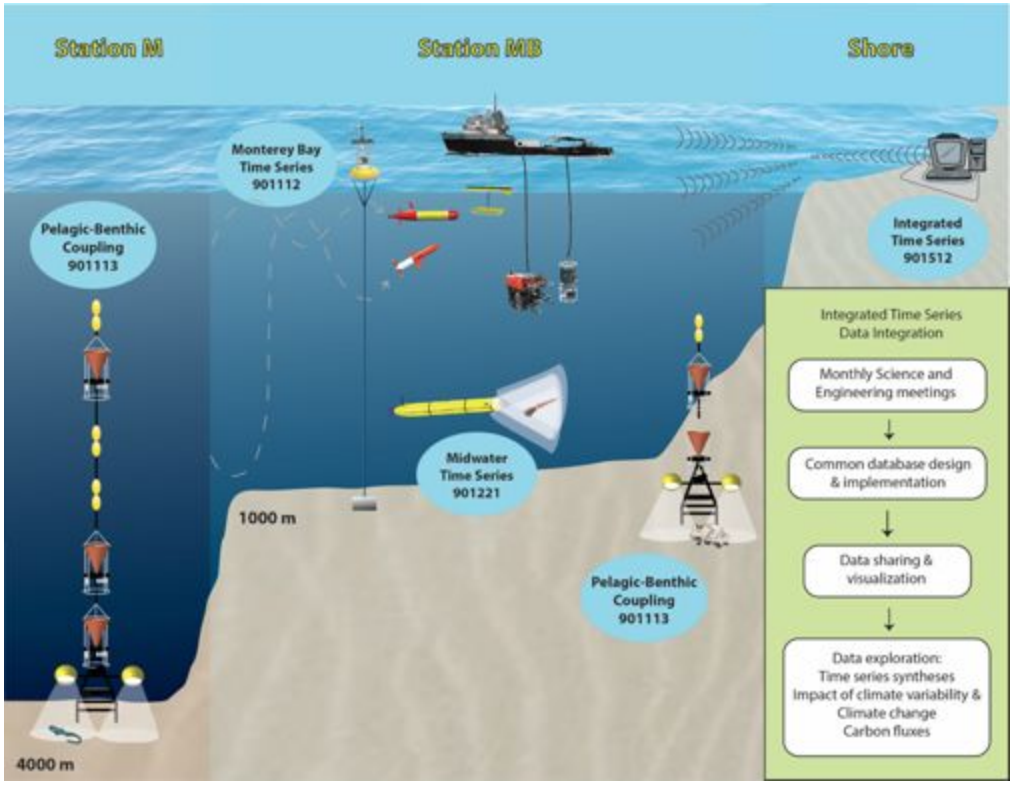
We generated the white paper which was titled “A New Vision of the MBARI Time-Series Program” which was used to prepare the 2015 proposal.

1. Time series are crucial to monitor change
2. Most places cannot fund long term
3. MBARI has 3 time series: Benthic, Midwater, Upper Water Column

4. They all instrument different parts of the carbon cycle
5. If we can combine them, it would produce an unprecedented time series allowing us to interpret the impact of climate change on the oceanic carbon cycle.
6. The goal is to integrate the three time series projects
    a. Goal: Reduce the cost of conducting these time-series studies
    b. Goal: Increase temporal resolution of sampling using autonomy
7. We are uniquely positioned to do this
8. Goal: Integrated the time series data sets into one data source
    a. Enable easy retrieval of data
    b. Facilitate the sharing of data outside MBARI
9. Will require additional instrumentation designed for higher temporal resolution while reducing costs

2015 - Project Kickoff

1. Results of feasibility:
    a. Integration would increase scientific and technological impacts
    b. Also increase efficiency and reduce cost
2. The migration from ship based to autonomy is being done under each respective time series project (there were details in the proposal of how all these efforts tied together and our best guess at schedule)



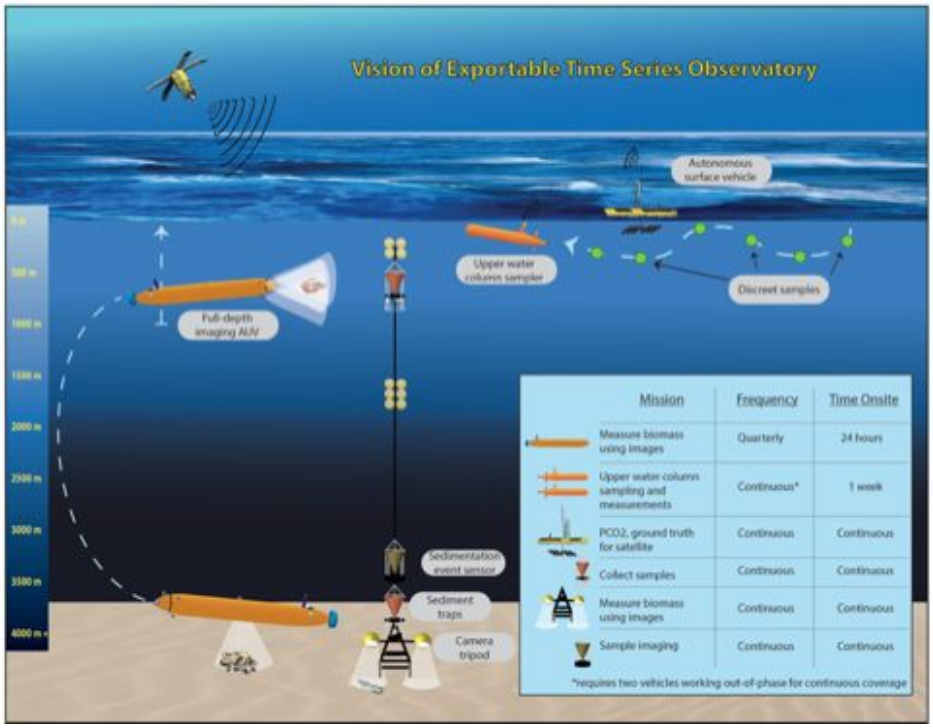
Goals:

1. Integrate data from three time series:
    a. Scientifically
        i. Analysis to identify links between surface, midwater and benthic environments
            1. concurrent climate driven changes across trophic levels
            2. carbon fluxes from from primary production to higher trophic levels and export to the seafloor
        ii. We identify these possible areas for scientific data analysis
            1. Data Synthesis
            2. Synchronous Variability
            3. Prediction

4. State of the Ecosystem
5. Outside datasets
6. Carbon Budget

b. Technically
   i. Propose common data store
   ii. Publicly share data
   iii. Add visualization tools based on scientific analyses

2. Ultimate goal is an integrated time series to achieve:
   a. Affordable and sustainable system
   b. Enable key ecosystem processes to be monitored
   c. Enable tracking of organic carbon production and utilization throughout the water column
   d. Enable assessment of the impact of climate variability
   e. Improve prediction of future states of the system, particularly under climate change
   f. Provide tools to monitor the state of Monterey Bay



Feedback from MT was that this vision was not likely on the timescale we detailed.

2016

Recap of 2015 efforts:

1. Paper published: K.L. Smith, Jr., M. Messie, A.D. Sherman, C.L. Huffard, B.W. Hobson, H.A. Ruhl and A. Boetius. (2015) Navigating the uncertain future of global oceanic time series. EOS 96, doi:10.1029/2015E0038095
2. Meetings
3. Scientific data objectives were focused on data synthesis and state of ecosystem because they share same goal of defining a set of representative variables from each time series.
   a. We defined primary data sets for these efforts:
      i. Surface: plankton concentration (counts per volume)
      ii. Midwater: animal density (counts per volume)
      iii. Benthos: echinoderm density (counts per area) from towed sled and benthic ROV transects
   b. Lots of work on data cleaning, harmonization, and preprocessing
   c.

4. Technical
    a. Several technologies were investigated for common data store.
    b. Extractions based on Monique's work were developed to try to replicate the data extractions used in her analyses.
    c. These extractions were loaded stored in files while technology explorations were occurring.

Goals:

1. Meetings
2. Scientific
    a. Investigating ecosystem indices
        i. Key taxa
        ii. Biodiversity indices
        iii. Indices of common variability
    b. Try to characterize ecosystem variability from 1989 to present. 97-98 El Nino shows clear signal across all three time series.
    c. Deriving carbon fluxes based on respiration rates is a work in progress and will continue this year
    d. Look at advection impacts on time series
3. Technical
    a. During scientific analyses, it was clear extractions need flexibility, will work to make data extractions more configurable
    b. As scientific analyses continue, changes will be made to extraction routine to try and replicate data used during analysis
    c. Start on some visualization components of the combined data
    d. Work with Data TAG to integrate lessons learned into larger data management question facing MBARI (per MT feedback)

2017

Recap of 2016:

1. Meetings
2. Scientific data integration
    a. Formation of smaller 'data synthesis groups'
    b. 2015 and 2016 largely focused on biological (taxonomic) data sets
    c. Integration of phytoplankton cell size and satellite environmental data
    d. Continued work on ecosystem indices
        i. Biodiversity
        ii. PCA
        iii. Multidimensional scaling (self-organizing maps)
    e. El Nino study
        i. Changes in El Nino/La Nina: food supply, does it decrease during El Nino?
            1. Nothing to date shows linkage
        ii. Changes in El Nino/La Nina: timing, is there signal propagation across space time in three series?
            1. No consistent relationship found
        iii. Community structure and change, are communities during El Nino different than during normal times?
            1. Work is still ongoing, no results yet.
3. Technical data integration
    a. Continued development of extraction software
    b. Chose MongoDB as combine data storage as a test of the technology

c. Built prototype visualization interface that integrated three data sets as well as external data (like MEI)

Goals:

1. Meetings
2. Scientific data integration
    a. More focus on carbon flux estimates
    b. More work on ecosystem indices
        i. More work needed to define meaningful taxonomic groups
    c. More work on El Nino
    d. Take a step back and look at higher level relationships. For example:
        i. Strong upwelling should lead to diatom blooms
        ii. Diatom blooms should increase zooplankton concentrations
        iii. Associated midwater predator concentrations should increase
        iv. Should see increase carbon flux to seafloor
    e. Make progress towards carbon budget of the water column. This will be difficult:
        i. Need more carbon content data
        ii. Need bacteria respiration data
        iii. Have discussions to define how we might tackle the carbon budget in following years.
3. Technological
    a. Integration of more data sets (i.e. ROVCTD, etc.), more extraction routines
    b. Start the work of embedding aggregation and analysis into the MongoDB data source (which is why it was chosen)
    c. Look for more overlap in data systems in other projects
    d. Advance visualization to try to keep step with Monique's analyses. This might require different technology on the visualization side of things.
    e. Build mechanism to allow users to integrate and visualize their own data sets with ITS data (this came out of discussions with data synthesis groups.

2018

Recap of 2017

1. Meetings
2. Scientific data integration
    a. Primarily targeted state of the ecosystem category
    b. Experimenting with different techniques: PCA, nMDS
    c. Did find patterns in three systems:
        i. Stable until late 90s
        ii. More variability in all series after the 90s
        iii. 2005-2007 years stand apart in all three series
    d. Investigated how variability itself changes over time (integral).
        i. Surface responds quicker to upwelling than midwater which responds quicker than benthos
    e. Monique taking leave of absence so not starting new threads
3. Technological integration
    a. Requirements moved towards similar requirements for ESP and CANON data systems
    b. For efficiency shifted toward re-tooling to align development efforts and software across the projects.
    c. Changed extraction and data storage to align with catalog efforts in ESP and CANON projects.

Goals:

1. Meetings
2. Scientific

    a. Wrap up current lines of investigation prior to Monique's leave
    b. Not starting new lines of investigation
3. Technological
    a. Develop shared catalog stack of software that will be used across ESP, CANON and ITS projects
    b. Develop visualization interfaces on new software stack that for the ITS project
    c. Discussions about requirements for public data availability which will define needs for authorization and authentication (if any).

2019

Recap of 2018

1. Meetings
2. Scientific
    a. Monique finished study that was presented at Ocean Sciences in Portland, OR. "Temporal Variability and Environmental Forcing of Surface, Midwater and Benthic Communities in the California Current System"
    b. Presentation detailed how upwelling forcing propagates through the three time-series communities, which provides an important understanding of how climate variability impacts ecosystems across the water column
3. Technological
    a. Catalog work continued on software for ITS, ESP, CANON
    b. Not much progress was made

Goals:

1. Monique to refine her analysis and results.
    a. Paper based on her presentation from Ocean Sciences: "Temporal Variability and Environmental Forcing of Surface, Midwater and Benthic Communities in the California Current System"
2. Combined data set that is continually updated with occasional ETL jobs and made available
    a. Question: Is this available to the public?
3. Produce a write up detailing project results and what technology and data gaps exist that prevent us from achieving any of the original goals (listed below) and how we could address those in future years (i.e. a 'roadmap'):
    a. Affordable and sustainable system
        i. Midwater transects are done now with AUVs
        ii. Also integrating acoustics and video data (was not imagined when we started)
        iii. Key species automated identification to help with video analysis workflow (in development in conjunction with Video TAG)
    b. Enable key ecosystem processes to be monitored
        i. Data processing improved translation of data
        ii. We now know how different data from different time series relate and link.
        iii. This is all of Monique's work here.
    c. Enable tracking of organic carbon production and utilization throughout the water column
        i. Still bits and pieces
        ii. Community is broad (species, types)
        iii. Still have work to do in carbon quantification for this broad range.
        iv. Need a geospatially synchronized site for this. Current sites are too different. We tried to ask for this and no-go.
        v. This may be impossible due to variability in processes at different depths (advection).
    d. Enable assessment of the impact of climate variability

    i. Each time series can assess that but because of the variability being able to do it for the entire system is difficult.
    ii. Geographic distance does not help either.
e. Improve prediction of future states of the system, particularly under climate change
    i. Same as d.
    ii. Each individual time series is making progress on this, just not as a whole system.
f. Provide tools to monitor the state of Monterey Bay
    i. i2MAP
    ii. LRAUV
    iii. Rover
    iv. Time lapse cameras
    v. SES
    vi.
g. Other:
    i. Hit the mark on tech development
    ii. Understanding the kinds of data
    iii. 3 time series are comprehensive, but still things that are not covered
    iv. Microbes are something that are not in the time series, this might be something that needs to be added.