WRT-1003: Data Assimilation, Analysis and Fusion across Modalities and Varying Latencies: Earth Orientation Parameters' Estimation and Prediction

Technical Report SERC-2021-TR-005

September 27, 2020

Principal Investigator: Dr. Mark Psiaki, Virginia Tech

Co-Principal Investigator: N/A

Research Team:

Virginia Tech: Demetrios Matsakis and Mark Psiaki

Sponsor: U.S. Naval Observatory

OSD Defense Research and Engineering

The Systems Engineering Research Center (SERC) is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology.

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense through the Office of the Assistant Secretary of Defense for Research and Engineering (ASD(R&E)) under Contract [HQ0034-19-D-0003, TO#0658].

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Executive Summary

Algorithms and software have been developed to improve the assimilation of Earth Orientation Parameter data for purposes to developing Earth Orientation estimates that need to be disseminated to other military and government entities and to the general public. The work concentrated on modernizing a cubic-spline data fitting technique to express it in terms of normal equations, to enable the simultaneous fitting of function and rate data, and to allow data cross-correlations in the statistical model of noise. The resulting algorithms and software were tested on example Earth Orientation Parameter data, and the results were compared with current best estimates of the parameters based on long smoothing runs. Additional work considered novel means of using GPS data, including GPS-derived USNO Earth Orientation rate of change data, and Atmospheric Angular Momentum data.

PERSONNEL

Research Team: (List all team members, past and present.)

Name	Organization	Labor Category	Contact
Mark Psiaki	Virginal Tech	Principal Investigator	mlpsiaki@vt.edu
Demetrios Matsakis	Virginal Techs	Visiting faculty	dnmyiasou@yahoo.com

Government:

Name	Organization	Role	Contact
Nicholas Stamatakos	USNO	Technical Lead	nicholas.g.stamatakos.civ@mail.mil

BACKGROUND

There are a variety of systems engineering challenges in analyzing data across multiple modalities and latencies, in the presence of uncertainty and data loss. The Earth Orientation Department (EOD) of the US Naval Observatory (USNO) needed to upgrade older data fusion algorithms (many created 25+ years ago) to newer procedures, algorithms, and software to increase robustness, accuracy, and optimality of combination, prediction, and estimation solutions. The USNO sought to conduct research on methods for meeting this general systems need by considering a particular case of interest, that of data assimilation for estimating Earth Orientation Parameters (EOP).

USNO is responsible for measuring and predicting the orientation of the Earth, and reporting the angle and rate of the Earth's rotation, the position of its poles, and other parameters to DoD, the Global Positioning System (GPS) wing, several US government agencies, and the general public.

In order to meet this requirement, the EOD Parameter Combination and Prediction Division (C/P) uses data gathered from a variety of internal and external sources, among them the International GNSS Service (IGS), Very Long Baseline Interferometer (VLBI), Satellite Laser Ranging (SLR), and Atmospheric Angular Momentum (AAM). These data provide complementary information about the several components of the Earth's orientation with differential degrees of accuracy and precision, which can change depending on the age of the data. Some of these data are available on a very irregular basis and none are obtainable with complete certainty. The data are subject to systematic errors, including sudden or gradual change of bias, sudden or gradual degradation or improvement, and outlier contamination.

The EOD algorithms that combine these disparate input data sets were developed over 30 years ago; the primary combining algorithm is called the smoothing, weighted cubic spline (SWCP). While these algorithms and software continue to produce reasonable results, there is growing concern regarding its continued use. The concerns include:

a) the SWCP software is no longer supported by the vendor

b) the software is written in older FORTRAN code, which is hard to understand and debug quickly,

c) older FORTRAN codes seem to raise Information Assurance (IA) concerns, and

d) newer optimal estimation techniques and software have replaced the “in-disuse” SWCP.

In fact, it is hard to find modern, updated literature discussing the use of the SWCP. So, when problems began to arise because of increased input data density, debugging the issues with the SWCP became difficult. Continued use of the SWCP will limit the ability of the EOP CP Division to carry out its mission in the future.

Most of the GPS Data Analysis Division EOPs are computed as the by-product of orbit determination: data from the IGS global network of GPS tracking receivers are processed in order to estimate orbits and clock corrections for the GPS satellites, as well as the desired EOPs and other quantities of interest such as synchronization parameters for ground clocks and estimates of zenith troposphere delay. In addition, this program computes and distributes a unique GPS-based quantity, UTGPS, used to extrapolate UT1-UTC estimates past the most recent VLBI measurements.

The parameter estimates produced and software used are not static. Changes in the underlying scientific theories, advances in technology, new standards and resolutions adopted by international scientific governing bodies, and changing user needs must be continually reviewed, and changes implemented when appropriate, in order to ensure that the most accurate and useful information is provided in the products. This review-and-update process is part of the production cycle.

RESEARCH WORK CONDUCTED BY VIRGINIA TECH

MONTHLY ACTIVITIES AND ACCOMPLISHMENTS

The following section gives a chronological list of summaries of the research that has been accomplished during the monthly reporting periods of this project.

NOVEMBER 2019

Two main items of progress were accomplished from the start of the contract through mid November 2019. The first was the development of a Matlab version of the IMSL spline smoothing function icsscu. This version was shared with Nick Stamatakos and Maria Davis of USNO. It was also tested at Virginia Tech using ICSSCU input/output data as provided by Nick Stamatakos of USNO. The tests showed that the original Fortran IMSL code and the new Matlab code produce the same results to 13 significant digits.

The second accomplishment was a theoretical development that proves that the ICSSCU cubic spline data fitting formulation is the equivalent of a sampled-data continuous-time Kalman filter. This formulation helped guide the development of algorithms. A literature search was also conducted in order to find other potential Kalman formulations that are similar to this formulation.

December 2019

Three main items of progress were accomplished during the period from mid November to mid December 2019. The first was work to complete the literature review. The second was development of an algorithms list. The third was testing of prototype algorithms using USNO UT2S-TT data that had been provided by Nick Stamatakos. The literature review is reported in the next subsection of this report as is the algorithms list.

The third accomplishment tried a preliminary idea of one of the new algorithms. It computed UT2S-TT from standard USNO data using a Virginia Tech implementation of the standard USNO cubic spline algorithm. The second method performed differentiation of the UTGPS data and the Atmospheric Angular Momentum (AAM) data for UT2S-TT. A modified cubic spline fit was then performed that fit these time derivatives directly rather than fitting their synthesized UT2S-TT versions. Thus, this represented a more direct use of the UTGPS and AAM data that has the effect of filtering out their less reliable low-frequency components. Preliminary comparisons with fits of UT2S-TT based on later data sets indicated that the new fitting technique may improve on what is currently being achieved.

January 2020

Two main items of progress were accomplished from mid December 2019 to mid January 2020. The first was to develop an end-to-end technique to test the use of AAM and UTGPS derivative data in place of UT2S-TT integrated data. This testing technique was designed to compare the results with the published finals for UT1-UTC after removal of various modeled effects that turn UT1 into UT2S and after translation from using UTC as the differencing quantity to using TT so that the equivalent of the published finals of UT2S-TT are determined, which is the quantity for which USNO performs its data fitting calculations.

The second accomplishment used the new end-to-end test technique to evaluate the new techniques that fit differentiated UTGPS and AAM data. The evaluations provided a preliminary indication that this new data assimilation approach improves the EOP estimates and predictions.

February 2020

One main item of progress was accomplished during the period from mid January to mid February 2020. It was to finish defining and writing a draft of the math spec for a spline-based data assimilation/smoothing algorithm that can incorporate derivative information. It explains how the resultant calculations are the equivalent of a Kalman filter and how they constitute a normal-equations-based calculation of Earth Orientation Parameters. The draft write-up was sent to Nick Stamatakos of USNO.

MARCH 2020

Two main items of progress were accomplished during the period from mid February to mid March 2020. The first was to review, revise, and finalize the draft math spec. for the new spline algorithm. The second item was to encode the algorithms defined in the new math spec. in Matlab and to perform preliminary tests on them. The revised paper, the Matlab code, and a successful test data file were transmitted to Nicholas Stamatakos of USNO on 26 Feb. 2020. These items are also being transmitted along with this report.

APRIL 2020

Three main items of progress were accomplished during the period from mid March to mid April 2020. They are all related. The first task was the testing of the new spline smoothing algorithm on IGS Length of Day (LOD) data, particularly LOD predictions, in order to evaluate whether these data point might be a good substitute for USNO's UTGPS. Time-integrated LOD data were compared to changes in UT1-UTC from IERS finals2000a.data. It was found that they have always been at least as good as finite-time-differenced UTGPS data and that they have maintained their accuracy after finite-time-differenced UTGPS data started deteriorating in quality in about 2014 or 2015. The second task was to evaluate whether finite-time-differenced UTGPS data quality depends in which subset of the GPS constellation is being used. One test considered just the Block IIR, satellites, and a second test considered the Block IIR and the Block IIR-M satellites. Both tests left out the Block IIF satellites. Neither of these tests showed better accuracy after the onset of the deterioration that started in 2014 or 2015. This seems to rule out the possibility that new GPS satellite designs are the causes of the problem. The third task was to begin evaluating GPS-based LOD predictions from JPL, which is a U.S. sourced data set. These were considered as another possible alternative to UTGPS. Preliminary results indicated that JPL-based LOD data are not nearly as accurate as finite-time-differenced UTGPS, even degraded UTGPS after 2015.

MAY 2020

Two main items of progress were accomplished during the period from mid April to mid May 2020. The first task was to evaluate USNO LOD data in order to determine whether it might be a suitable substitute for UTGPS and just as good a substitute as IGS LOD data. The second task was to evaluate the replacement of UTGPS data with IGS LOD data in the USNO's daily data assimilation cubic-spline smoothing calculations in order to determine whether there is an advantage to IGS LOD.

JUNE 2020

Two main items of progress have been accomplished during the period from mid May to mid June 2020. Both tasks centered around the analysis of LOD data in order to compare it to UTGPS data as a possible substitute for the latter type of data. The first task was to try using IGS LOD data in EOP UT2S-TT spline fit calculations. These spline fits use -LOD as rate data. The task also evaluated the prediction capabilities of such spline fits. These predictions were evaluated by computing their errors relative to "truth" UT2S-TT as derived from finals2000A.data. The prediction capability provided by IGS LOD data was compared with that provided by the existing USNO use of UTGPS and with a modified version that uses finite-differenced UTGPS and AAM derivatives as direct inputs to the spline. The latter two methods produced a superior prediction capability out to 10 days.

The second task was to examine LOD errors, both IGS LOD and USNO LOD, by computing their differences from the "truth" LOD values in finals2000A.data and from the equivalent negative time derivatives of a cubic-spline fit to the UT2S-TT that corresponds to the "truth" UT1-UTC from finals2000A.data. All comparisons were performed after subtracting out the seasonal and tidal effects and after correcting for leap seconds, both in going from UT1-UTC to UT2S-TT before performing a cubic spline fit and taking derivatives and in removing these effects from raw LOD data. Similar analyses were performed for time-differentiated raw UTGPS and time-differentiated pre-processed UTGPS. The results showed that IGS and USNO LOD data are much less noisy than recent time-differentiated raw UTGPS. Old pre-2014 UTGPS was comparable to IGS and USNO LOD in terms of noise power. The results also showed that none of these three GPS-derived sources of rate information yield white-noise errors. Rather, each of these signals has a frequency-dependent power spectrum – see Figure 1. This fact indicates that the errors at different sample times are correlated. This fact has led to a proposal to develop yet another version of the data assimilation cubic spline, one that can account for these time correlations.

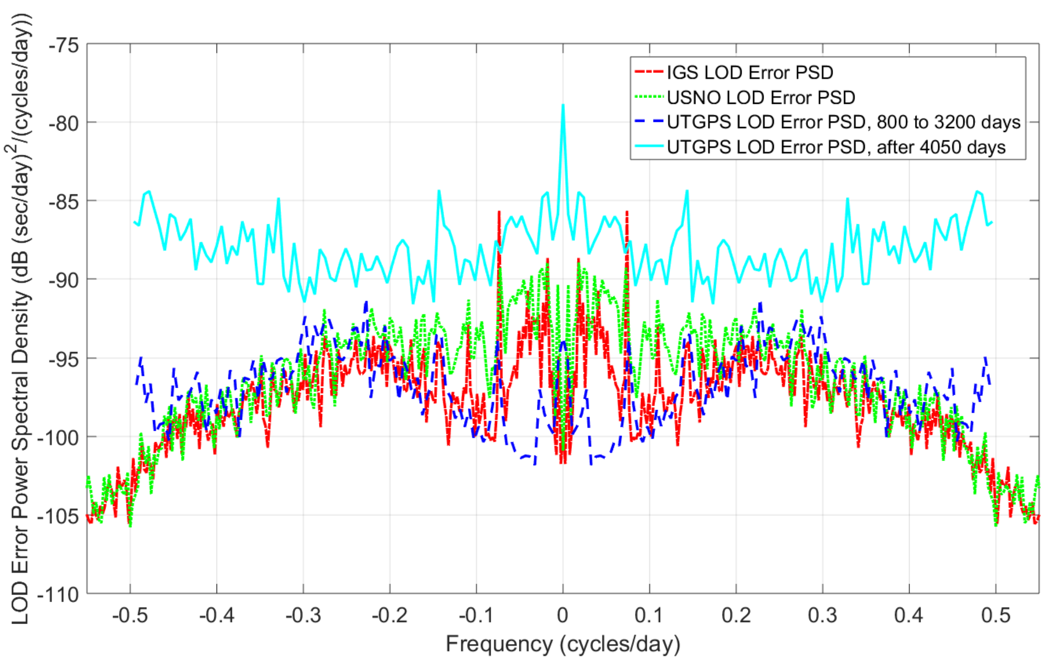

Figure 1 Power spectral densities of (UT2S-TT)-rate-equivalent LOD data errors of various data sources.

JULY 2020

Two main items of progress have been accomplished during the period from mid June to mid July 2020. The first was to complete the development of the new cubic spline algorithm that optimally compensates for time correlations of data, especially LOD rate data. The second milestone was to test this new algorithm on USNO data assimilation data sets in order to evaluate its smoothing and prediction capabilities when using IGS LOD data, USNO LOD data, or time-differentiated raw UTGPS data instead of pre-processed raw UTGSP data. The new algorithm worked well, and all three considered data sources produced good results that are reasonable candidates for eventual use by USNO.

AUGUST 2020

Five main items of progress have been accomplished during the period from mid July to mid August 2020. The first was to complete work needed to use multiple LOD or UTGPS time derivative data sets simultaneously in the latest USNO cubic spline The main work of this accomplishment was to compute cross-correlations between the errors in such data. The second accomplishment was to test the new cubic spline algorithm using multiple rate data sources in order to evaluate whether the use of multiple data sources is a good idea. Counter-intuitively, the use of multiple rate data sources slightly degraded the results rather than improving them. The third accomplishment was to model the error statistics and proper time tagging of time-differentiated versions of the Atmospheric Angular Momentum (AAM) predictions that are used by the USNO to improve the prediction parts of its spline calculations. The fourth accomplishment was to use the new models of Atmospheric Angular Momentum prediction errors and times of applicability in an attempt to improve the cubic spline's prediction capabilities. The attempt appears to have been successful, but an ad hoc scaling down of measurement error covariances and cross-correlations by a factor of (1/1.75)^2 was needed to get an improvement. The fifth accomplishment was to uncover a possible reason why the ad hoc re-scaling was needed: The finals2000A.data UT1-UTC data indicate that the UT2S-TT second derivative time history colored noise rather than white noise. The cubic spline assumes a white noise second derivative. This modeling discrepancy could cause some of the need for the ad hoc factor of (1/1.75)^2.

SEPTEMBER 2020

One main item of progress has been accomplished during the period from mid August to the end of the project at the end of September 2020. It was to study and evaluate the use of raw NOAA Atmospheric Angular Momentum AAM data for use in the forecast part of the UT2S-TT spline calculation. Prediction and analysis data were both studied. The studies indicated that both data sets have large Length of Day (LOD) biases, and that the biases drift over time. They also indicated that the NOAA analysis LOD values are not necessarily helpful for removing any bias in the prediction LOD values. The final take-away from these analyses was that the best manner of using raw NOAA AAM LOD predictions might be to have the cubic spline software estimate an LOD prediction bias as part of its calculation.

SUMMARY OF ACTIVITIES AND ACCOMPLISHMENTS

The activities involved in this project have included literature searching, consideration of algorithms, design of algorithms, encoding of prototype versions of algorithms, testing of algorithms, and consideration of three alternate GPS-based data types and one alternate Atmospheric Angular Momentum (AAM) data type and their use. All of the developed algorithms have the common goal of assimilating EOP data to form the most accurate possible EOP estimates for distribution to EOP users.

One of the most significant accomplishments has been the development of a sequence of modernized cubic spline algorithms for EOP data assimilation. Three important capabilities have been added to these algorithms. The first is a demonstration that all such algorithms calculate optimal solutions of least-squares normal equations for an appropriately defined least-squares problem.

The second capability of the new algorithms and software is that they can assimilate function and rate data, not just function data. The cubic spline software that is being replaced can only assimilate function data. The ability to assimilate rate data is important because GPS-based data and AAM prediction data are fundamentally rate-type data. Currently USNO uses techniques that, in effect, integrate these rate data. These integrations require a number of ad hoc data manipulations in order to develop something that is useful, and there is evidence that the usefulness of such pre-processed data may be transitory. The ability to assimilate raw rate data has been developed in hopes of better maintaining a long-term ability to assimilate such data even as its quality changes over time.

The third capability of the new algorithms and software is an ability to optimally assimilate data whose random errors have correlations. The possible modeled correlations that can be handled by one of the new algorithms can be between different time samples of a given class of data, between different classes of data, and between function data and rate data. This capability enables the data assimilation algorithm to be fine-tuned in a way that optimally adjusts for differing qualities of different data types without any need for ad hoc data pre-processing, or at least, for less such pre-processing.

One final category of accomplishments concerns possible new ways to assimilate two types of data and the possible use of three new data types. The existing data that have been considered for changing their method of assimilation are the USNO GPS product known as UTGPS and the AAM predictions. Both of these data or prediction types are fundamentally of the rate type. In the current assimilation scheme, they must be integrated in order to be used as pseudo-function data. The re-consideration of these data differentiates UTGPS and AAM and with goal of using the more fundamental time derivatives directly in the data assimilation calculations and thereby improve accuracy while potentially eliminating the need for much of the ad hoc pre-processing that goes into generating these two data types.

The three new data types that have been considered are the USNO-generated and the IGS-generated Length of Day (LOD) and the raw NOAA AAM, which can be used to generate LOD predictions. The negative of LOD is effectively the time derivative of UT1-UTC. Therefore, such data can be assimilated directly using the function-and-derivative version of the spline. The goal of considering these three data types is to obviate the need to generate and use the USNO UTGPS data product and to eliminate the ad hoc pre-processing that is needed to generate the current UT2S-TT-like AAM data product that is used by USNO for short-term prediction. The USNO and IGS LOD data products are generated by consideration of GPS signals and GPS satellite orbits in ways that are nearly the same as the way that UTGPS data are generated. There has been some success at trying to substitute LOD for UTGPS, but a final verdict has not yet been reached. The LOD values from NOAA's raw AAM prediction proved challenging to use in an effective way, but there is still some hope that a way will be found to use these data.

RESULTS AND OUTPUTS OF RESEARCH

The following section catalogues the principle results and outputs of this research.

LITERATURE REVIEW REPORT

The literature review consisted partly of reviewing papers that explore the links between splines and Kalman filters. These include:

Harashima, M., Ferrari, L.A., and Sankar, P.V., "Spline Approximation using Kalman Filter State Estimation," IEEE Trans. on Circuits & System – II: Analog & Digital Signal Processing, Vol 44, No. 5, May, 1997, pp. 421-424.

Jauch, J. Bleimund, F., Rhode, S., and Gauterin, F., "Recursive B-Spline Approximation using the Kalman Filter," Engineering Science & Technology, an International Journal, Vol 20, 2017, pp. 28-34.

Kochegurova, E., Khozhaev, I., and Ezangina, T., "Design of Recursive Digital Filters with Penalized Spline Method," In: Nguyen, N., Pimenidis, E., Khan, Z., and Trawiński, B. (eds) Computational Collective Intelligence. ICCCI 2018. Lecture Notes in Computer Science, Vol 11056, 2018, pp 3-12.

Jauch, J. Bleimund, F., Frey, M., and Gauterin, F., "An Iterative Method Based on the Marginalized Particle Filter for Nonlinear B-Spline Data Approximation and Trajectory Optimization," Mathematics, Vol. 7, No. 355, 2019, pp. 1-24.

Another part of the literature review was to read and become familiar the method and conventions of EOP as contained in

Petit, G., and Luzum, B., IERS Conventions (2010), IERS Technical Note No. 26, Verlag des Bundesamts fur Kartographie and Geodasie, (Frankfurt am Main, 2010).

The main result of the literature review is that the spline data fitting and interpolation techniques currently in use at USNO for EOP estimation are the equivalent of a Kalman filter and a smoother. Therefore, the algorithms are already optimal in an important sense. This result also suggests two methods by which the estimation techniques can be improved. The first is to exploit the ability of a Kalman filter to assimilate different types of data, in particular rate data and interval data for the EOP estimation problem. This can be done in the original cubic-spline-fitting context by developing updated versions of the IMSL function icsscu. The second method by which the EOP estimation techniques can be improved is to exploit the recursive nature of a Kalman filter and a Rauch-Tung-Striebel (RTS) smoother. Use of such techniques do not yield a different answer (other than through the effects of machine round-off error), but they operate quickly, in computation time that scales linearly with the number of data points rather than cubically.

ALGORITHMS LIST

The algorithms that have been identified as being promising based on the literature review and other considerations are as follows (with explanations following).

1. Modified icsscu cubic spline fitting algorithm that can process function rate and function interval data as well as function data. This will enable use of direct un-integrated AAM data and differentiated UTGPS data as means of processing these data in a more rational manner that better models the information that they contain and that filters out their less certain low-frequency parts when they are submitted as (UT2S-TT)-like data points.

2. Kalman filter/RTS-smoother version of cubic spline fitting algorithm of 1. This will enable much faster calculation of results.

3. Pre-processing algorithms that directly compute and work with rate data for (UT2S-TT) as given by AAM and UTGPS. These are expected to make better use of these two data types in a way that exploits their strengths while minimizing the impacts of their deficiencies.

REPORT ON NEW TECHNIQUES BASED ON GPS ORBITS (OR THE EQUIVALENT)

Two possible new techniques based on GPS orbits have been identified for use as part of the EOP data fusion process. One possible technique is to use a time-differentiated version of UTGPS, possibly with improvements to the initial calculation of raw UTGPS and other processing of UTGPS that is carried out before it is input to the data fusion cubic spline calculations. The time-differentiated UTGPS would be calculated using finite differencing. These rate data would be input directly to the new version of the cubic spline that allows the input of both function and rate information. The second possible technique is to use LOD rate information in place of UTGPS. The LOD data may come from the IGS or directly from USNO. Again, these data would be input to the new cubic spline directly as rates.

Evaluations of these two GPS-orbits-based sources of information about UT2S-TT have been carried out. Figure 2 and Figure 3 show example results for the EOP cubic spline data fusion processing that would have been carried out on Nov. 3, 2019. Both plots show errors between the particular method's calculated UT2S-TT time history and the best estimate of this time history that is obtained from the IERS data produce finals2000A.data. Note that some additional processing is needed in order to translate from the UT1-UTC values in finals2000A.data to the UT2S-TT values that are used for comparison purposes here. The solid blue curves in the figures show the results using the current USNO cubic spline technique that employs UTGPS and VLBI data. It also employs integrated AAM data. The dash-dotted red curves show the results when replacing UTGPS data by time-differentiated UTGPS data and AAM data by time-differentiated AAM data. The dashed black curves show the results when replacing UTGPS data by IGS LOD rate data and AAM data by time-differentiated AAM data. The dashed black curves also delete the short batch of integrated IGS LOD data points that are used near day 375 for each of the other two cases. Figure 3 is based on the same data as Figure 2, but it zooms in on the final portion of the interval so that the prediction interval after t = 375 days is clearly seen.

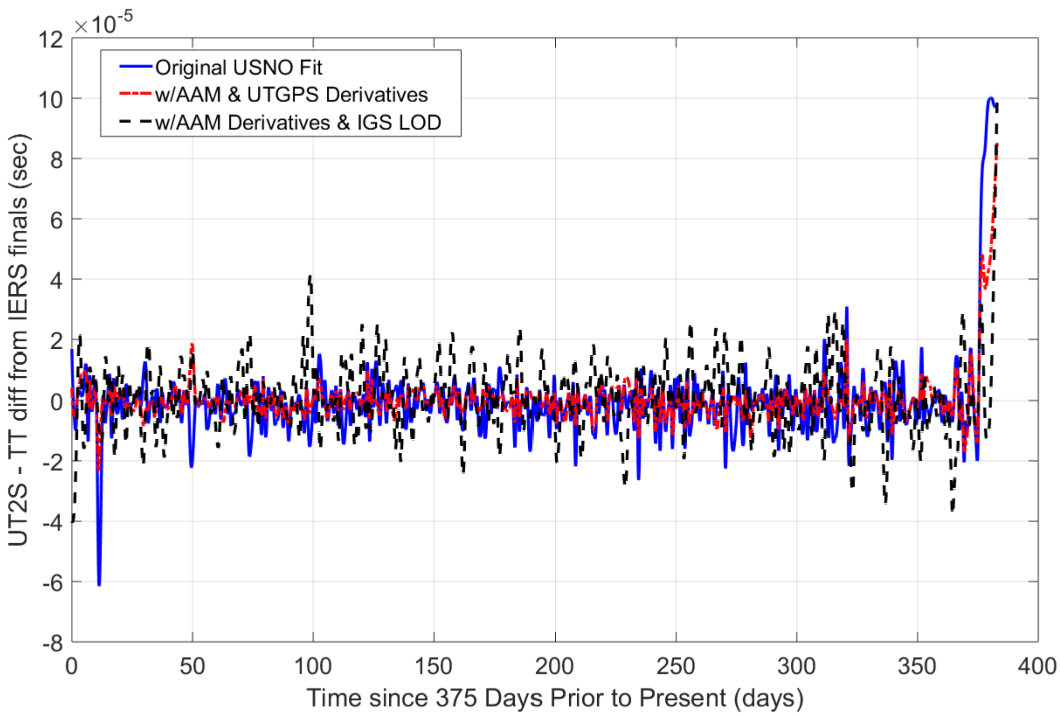

Figure 2 Comparison of errors between cubic-spline-generated UT2S-TT and reference UT2S-TT from finals2000A.data for Nov 3, 2019 (MJD = 58790) present-day solution.

Figure 2, Figure 3, and similar plots for other days' solutions make clear that the time-differentiated UTGPS and AAM case, the dash-dotted red curves, are the best at reducing noise in the smoothing sections of the data fusion cubic spline calculations. That is, the dash-dotted red curves stay closest to zero in an RMS sense at all times before t = 375 days. The prediction interval from t = 375 days to t = 383 days is less clear about which techniques is better, as shown in Figure 3. In this case and in a number of other cases, the two new curves (the dash-dotted red curve and the dashed black curve) do better than the current USNO technique (the solid blue curve) for prediction. More cases need to be considered, however, in order to make a final judgement.

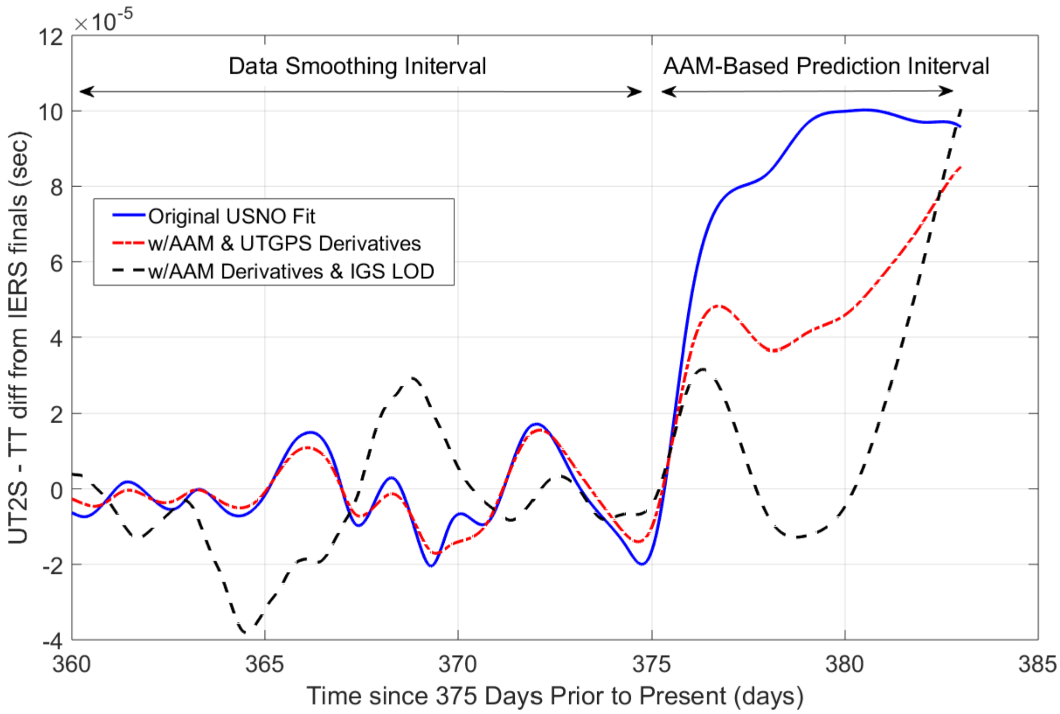

Figure 3 Final portions of plots that compare errors between cubic-spline-generated UT2S-TT and reference UT2S-TT from finals2000A.data for Nov 3, 2019 (MJD = 58790) present-day solution.

In summary, two new techniques have been evaluated for processing GPS-based data in order to aid USNO data fusion to estimate UT2S-TT. One of them involves using finite-differenced time derivatives of UTGPS. The other involves using IGS LOD data in place of UTGPS. In both cases, the new data types are input to the data assimilation cubic spline calculations directly as rates rather than as integrals of rates. The time-differentiated UTGPS method looks like it may be the best, but further study needs to be conducted in order to be sure that LOD is not as good. This study should include consideration of USNO-generated LOD in addition to IGS LOD. If the superiority of time-differentiated UTGPS is borne out, then further work should be done in order to improve the UTGPS data product that gets time-differentiate for input to these data assimilation calculations.

Updated Cubic Spline Theory Paper

One of the main tasks accomplished during this research was to define the math spec for a spline-based data assimilation/smoothing algorithm that can incorporate derivative information and to write a paper about this math spec. It explains how the resultant calculations are the equivalent of a Kalman filter and how they constitute a normal-equations calculation of optimal estimates of Earth Orientation Parameters based on a well-defined optimal estimation cost function. The resulting paper has been transmitted along with this final report.

DOCUMENTATION OF PROTOTYPE ALGORITHMS DELIVERED TO USNO

Eleven prototype algorithms and their corresponding Matlab .m-file functions have been generated during this project. The most important of them have already been delivered to USNO. All eleven are listed below with short descriptions of what they do. Lengthier descriptions of the functions are included in their respective introductory comments sections. Each prototype Matlab function's comments include a complete description of the function's inputs and outputs. Each of these .m-file functions is being sent along with this report.

Table 1 Prototype Matlab Software Delivered to USNO

Function Name	Brief Description of Function
cubicspline01.m	Evaluates the cubic spline and its first and second time derivatives at an arbitrary set of input times. It uses the same cubic spline parameterization that is generated by the functions cubicsplinesmoother01.m through cubicsplinesmoother06.m. That parameterization is given in terms of knot times, function values at the knot times, and derivative values at the knot times.
cubicsplineicsscu.m	Evaluates the cubic spline and its first and second time derivatives at an arbitrary set of input times. It uses the same cubic spline parameterization that is generated by the function icsscu.m. That parameterization is given in terms of the coefficients of the cubic polynomials that apply during the intervals between pairs of adjacent spline knots.
cubicsplinesmoother01.m	Finds the best-fit cubic spline for a given set of weighted function data and a constraint on the sum of the squares of the normalized errors between the cubic spline and the function data. The best fit spline minimizes the integral of the square of the spline's second derivative. This function is equivalent to the original IMSL routine ICSSCU, except that it generates its spline parameterization in terms of function and rate values at the spline knots rather than cubic polynomial coefficients that apply between pairs of adjacent spline knots.
cubicsplinesmoother02.m	This function is identical to cubicsplinesmoother01.m in terms of its inputs and outputs. The only difference concerns the internal nonlinear solver algorithm that it uses in order to determine its Kuhn-Tucker multiplier parameter.

Function Name	Brief Description of Function
cubicsplinesmoother03.m	Finds the nearly the best-fit cubic spline for a given set of weighted function and rate data and a constraint on the sum of the squares of the normalized errors between the cubic spline and the function and rate data. The nearly best-fit spline nearly minimizes the integral of the square of the spline's second derivative. In the case where the data at the two end knots is function data and not rate data, this function yields the cubic spline that minimizes the integral of the square of the spline's second derivative. If rate data are present at either end knot, then this function's solution is slightly sub-optimal because it constrains the second derivative to decay to zero at the corresponding end knot, which the optimal spline would not do.
cubicsplinesmoother04.m	Finds the nearly the best-fit cubic spline for a given set of weighted function, rate, and change-over-an-interval data and a constraint on the sum of the squares of the normalized errors between the cubic spline and the function, rate, and change-over-an-interval data. The nearly best fit spline nearly minimizes the integral of the square of the spline's second derivative. In the case where the data at the two end knots is function data or change-over-an-interval data and not rate data, this function yields the cubic spline that minimizes the integral of the square of the spline's second derivative. If rate data are present at either end knot, then this function's solution is slightly sub-optimal because it constrains the second derivative to decay to zero at the correspond end knot, which the optimal spline would not do. This function's change-over-an-interval input data represent, in effect, finite-difference time derivative data.
cubicsplinesmoother05.m	Finds the best-fit cubic spline for a given set of weighted function and rate data and a constraint on the sum of the squares of the normalized errors between the cubic spline and the function and rate data. The best fit spline minimizes the integral of the square of the spline's second derivative. This function generates its spline parameterization in terms of function and rate values at the spline knots. This function implements the algorithms that have been defined in the math spec paper that has been delivered with this report.

Function Name	Brief Description of Function
cubicsplinesmoother06.m	Finds the best-fit cubic spline for a given set of weighted function and rate data and a constraint on the sum of the squares of the normalized errors between the cubic spline and the function and rate data. The best fit spline minimizes the integral of the square of the spline's second derivative. This function calls cubicsplinesmoother05.m. It wraps an outer iteration around the basic spline calculation that edits out bad data points, both function and rate data points, if their deviations from the spline fit are too large relative to their input error standard deviations. This function is a strong candidate for final use by USNO.
cubicsplinesmoother07.m	Finds the best-fit cubic spline for a given set of weighted function and rate data and a constraint on the sum of the squares of the normalized errors between the cubic spline and the function and rate data. The best fit spline minimizes the integral of the square of the spline's second derivative. This function generates its spline parameterization in terms of function and rate values at the spline knots. This function differs from cubicsplinesmoother05.m in that it allows non-zero cross-correlations between the errors in different measurements, including between measurements that occur at different times. Its constraint on the sum of the squares of the normalized errors between the cubic spline and the function and rate data uses these cross-correlations so that the constrained function is a generalized sum of error squares that has the potential to include cross-term products of errors for different measurements.
cubicsplinesmoother08.m	Finds the best-fit cubic spline for a given set of weighted function and rate data and a constraint on the sum of the squares of the normalized errors between the cubic spline and the function and rate data. The best fit spline minimizes the integral of the square of the spline's second derivative. This function calls cubicsplinesmoother07.m. It wraps an outer iteration around the basic spline calculation that edits out bad data points, both function and rate data points, if their deviations from the spline fit are too large relative to their input error standard deviations. This function is the current best candidate for final use by USNO.

Function Name	Brief Description of Function
icsscu.m	Finds the best-fit cubic spline for a given set of weighted function data and a constraint on the sum of the squares of the normalized errors between the cubic spline and the function data. The best fit spline minimizes the integral of the square of the spline's second derivative. This function is a reproduction of the original IMSL routine ICSSCU. It has exactly the same inputs and outputs as ICSSCU. It generates its spline parameterization in terms of cubic polynomial coefficients that apply between pairs of adjacent spline knots. It has been developed solely for comparison purposes.

CONCLUSION

This project has developed algorithms and prototype Matlab software with the goal of improving USNO's Earth Orientation Parameter (EOP) data assimilation. Data assimilation to generate UT2S-TT has been the focus of this project's efforts. Several algorithms have been developed and prototyped. They are all enhanced cubic-spline algorithms. They have demonstrated an ability to operate at least as well as existing algorithms and software that are currently used by the USNO for its data assimilation operations. They also offer a direct connection to least-squares optimal estimation and will be more maintainable, especially given the math spec for the software that has been generated as part of this project.

This project has used the new algorithms to consider new ways of processing existing data types, UTGPS and Atmospheric Angular Momentum (AAM). In addition, it has enabled the consideration of two new GPS-derived data types: USNO Length of Day (LOD) and IGS LOD. These two data types are candidates for replacing UTGPS in order to obviate the need for all of the ad hoc pre-processing that is needed to generate UTGPS and that seems to be breaking down as the GPS constellation evolves. Consideration of new ways to processing AAM data has two goals: First, improve short-term prediction accuracy for UT2S-TT and second, obviate the need for the current ad hoc pre-processing of AAM predictions. This project's studies of the efficacy of using the various new or modified rate data have shown promise. More work needs to be done in order to arrive at a good decision about how to go forward with GPS-derived and AAM-derived inputs to USNO's data assimilation.

PROJECT TIMELINE & TRANSITION PLAN

1. The long-term transition goal for this research, if continued, is to produce operational code and a new data processing protocol that will be used operationally by USNO to assimilate data for purposes of estimating UT2S-TT.
2. The tools and guides that resulted from this research that might be used by

external sponsors if the long-term transition goals are met are the math spec paper about the new cubic spline and the Matlab functions cubicsplinesmoother05.m, cubicsplinesmoother06.m, cubicsplinesmoother07.m, cubicsplinesmoother08.m, and cubicspline01.m. These tools and guides have been delivered with this final report.

3. All of the useful tools and guides are all being delivered now. None of these is planned to be incrementally delivered in a future research task because this project is ending as far as the SERC is concerned. It may happen that USNO will contract with Virginia Tech using another contract vehicle in order to continue some of this work. In that case, it is possible that additional math specs and software prototypes of algorithms will be generated that might be used by external sponsors.
4. This project's transition partner is also its original sponsor, the USNO. There are no other known advocates or potential adopters of this research.
5. The research was piloted with a potential transition partner, the USNO. There are no other known entities who would conduct pilot use of the research if fully funded.

APPENDIX A: LIST OF PUBLICATIONS RESULTED

None, but the math spec paper may eventually be submitted for publication.

APPENDIX B: CITED AND RELATED REFERENCES [EXAMPLES]

Harashima, M., Ferrari, L.A., and Sankar, P.V. (1997). Spline Approximation using Kalman Filter State Estimation. *IEEE Trans. on Circuits & System – II: Analog & Digital Signal Processing*, 44(5), 421-424.

Jauch, J. Bleimund, F., Rhode, S., and Gauterin, F. (2017). Recursive B-Spline Approximation using the Kalman Filter. *Engineering Science & Technology, an International Journal*, 20, 28-34.

Jauch, J. Bleimund, F., Frey, M., and Gauterin, F. (2019). An Iterative Method Based on the Marginalized Particle Filter for Nonlinear B-Spline Data Approximation and Trajectory Optimization. *Mathematics*, 7(355), 1-24.

Kochegurova, E., Khozhaev, I., and Ezangina, T. (2018). Design of Recursive Digital Filters with Penalized Spline Method. In: Nguyen, N., Pimenidis, E., Khan, Z., and Trawiński, B. (eds) *Computational Collective Intelligence, ICCCI 2018, Lecture Notes in Computer Science*, 11056, 3-12.

Petit, G., and Luzum, B. (2010). IERS Conventions (2010). *IERS Technical Note No. 26*, Verlag des Bundesamts fur Kartographie and Geodasie, Frankfurt am Main..