STUDIES OF NPL'S CLOCK ENSEMBLE ALGORITHM

Setnam L. Shemar, John A. Davis, and Peter B. Whibberley

National Physical Laboratory, Hampton Road,
Teddington, Middlesex, TW11 0LW, UK
Email: *setnam.shemar@npl.co.uk*

Abstract

In the last few years the National Physical Laboratory (NPL) has been investigating a Kalman-filter based atomic clock ensemble algorithm as part of its long-term strategy for improving the stability of the UK's national time scale UTC(NPL). To achieve this, Flicker Frequency Modulation (FFM) clock noise has been modeled within the algorithm, using integrated Markov noise processes. Recent results have indicated the performance of the composite time scale generated for a 150-day data-set using an ensemble of two 5071A cesium clocks together with three active hydrogen masers. Here we firstly investigate the performance of the algorithm using a longer data-set of 846 days. We then investigate its performance by comparison against UTC and NPL's cesium fountain, NPL-CsF2. We also consider how data from NPL-CsF2 can be exploited within the algorithm to improve the stability of the generated composite. Finally, we briefly discuss the potential application of such an algorithm for improving the stability of the UK's national time scale UTC(NPL).

INTRODUCTION

Clock ensemble algorithms can be separated into two types. These are often based on the weighted mean of a number of clocks, whereby the weight of each clock is given by a measure of its stability, for example using the Allan Deviation (ADEV). An example of this type is the NIST AT1 algorithm **[1]**. This has been used for many years by NIST with data from an ensemble of active hydrogen-maser (active H-maser) frequency standards and cesium clocks **[2]**. A similar algorithm has also been used with data from cesium clocks as part of the development of a new time scale at LNE-SYRTE **[3]**. Alternatively, ensemble algorithms can be based on the use of a Kalman filter **[4,5]**. In recent years the National Physical Laboratory (NPL) has been developing such an algorithm for consideration as part of its long-term strategy for improving the stability of UTC(NPL). The details of the algorithm have been presented previously **[6]** together with preliminary results using input data spanning a period of 150 days from both active H-masers and cesium clocks **[7]**.

For this paper we have carried out some studies to investigate further the performance of the algorithm as well as an approach to exploit data from NPL's cesium fountain primary frequency standard, NPL-CsF2. We firstly give an overview of the fundamental requirements and properties of the algorithm. We then present results from the algorithm using input data spanning periods of up to 846 days. This is followed by comparisons of the composite time scale against UTC and NPL's cesium fountain, NPL-CsF2. We also consider using data from NPL-CsF2 within the algorithm to improve the long-term stability of the composite time scale. Finally, we discuss a potential strategy for utilizing the algorithm to generate UTC(NPL) as well as some of the further developments required to achieve this.

OVERVIEW OF ALGORITHM

A detailed account of NPL's clock ensemble algorithm has been presented previously **[6]** together with a summary of the key mathematical concepts **[7]**. It requires input data from two or more clocks where one of these is used as a reference clock for the time-offset measurements. The output of the algorithm provides an estimate of the time-offset of each clock from the generated composite time scale. It assumes that there is no measurement noise in the clock input data. The state vector comprises the parameters estimated by the algorithm. These include three components corresponding to the time offset, frequency offset and linear frequency drift offset from the composite for each clock. Flicker Frequency Modulation (FFM) is modeled as a linear combination of integrated Markov noise processes using additional state vector components. The composite has close to optimal performance in the presence of FFM **[6]**.

The linear frequency drift present in an active H-maser may reach levels of $10^{-19}s^{-1}$. Such values will result in a time offset of order 30 ns over a data span of 10 days. Consequently, a significant non-zero value of a linear frequency drift term is used in the case of the active H-masers where appropriate. In the case of the cesium clocks, frequency instabilities originating from stochastic noise processes dominate when compared with those resulting from any linear frequency drift. Consequently, the linear frequency drift parameter for each cesium clock is maintained very close to zero. This approach is used to assist the estimation of the linear frequency drift of the H-masers relative to the cesium clocks. The model also assumes that the linear frequency drift of the active H-masers is not constant.

Separate "noise parameters" are used within the algorithm to model the noise of each clock type. These correspond to the variances of the white noise processes underlying each clock noise type. The Allan Deviation (ADEV) values at short averaging times are very much lower for the active H-masers than the cesium clocks. This is not necessarily the case for long averaging times. However, the Hadamard Deviation (HDEV), which provides a measure of the stability after rejecting frequency drift, can still be lower in the active H-masers. This is because HDEV is a statistic that is insensitive to the linear frequency drift typically observed in these clocks. Consequently, higher weights are given to these within the algorithm than the cesium clocks at all averaging times. Furthermore, for this paper, different weights have been applied to each H-maser as a result of the significant differences observed in their performance. This is achieved by applying different noise parameters in each case.

INITIALIZING THE ALGORITHM

The algorithm requires initialization of certain clock parameters. The performance of the algorithm will depend on the values of the noise parameters set and it is important that these reliably match those of the actual clock noise processes. Ideally these would be determined by comparing each clock against a perfect clock. In practice, these have been obtained by comparison against UTC(NPL) and NPL-CsF2. These parameters are as follows:

(i) noise parameter estimates for each clock;

(ii) time-offset, normalized frequency offset, linear frequency drift offset, and five integrated Markov noise parameters for each clock;

(iii) variance estimates of the time offset, normalized frequency offset, linear frequency drift, and five integrated Markov noise parameters used in the diagonal elements of the parameter covariance matrix, *P*, for each clock.

The frequency drift offset for the cesium clocks is always initialized to zero. Good initialization of the other individual clock offsets is vital to minimize the time required for parts of the algorithm, such as the Kalman Gain, to achieve steady state. For the results presented below, the linear frequency drift of the H-masers have been initialized to non-zero values as appropriate.

The diagonal elements of the parameter covariance matrix for each type of clock should be set to values that are physically realistic. These correspond to the uncertainties in the initial values of the state vector components.

INPUT CLOCK DATA

Data from three of NPL's active H-masers and two commercial cesium clocks have been used as input to the algorithm. The three H-masers are designated as HM1, HM2, and HM4, where the data from HM2 are equivalent to UTC(NPL).

The data for each clock are derived from measurements between the 1 Pulse Per Second (1 PPS) output of each clock and that of a reference clock. Measurements are taken once per hour and correspond to the mean of a series of 10 measurements taken at one second intervals for each clock using a Stanford Research Systems model SR620 time interval counter. For the results presented in this paper, unless otherwise stated, only the first hourly measurement taken in each day for each clock is used as the input to the algorithm, corresponding to one data point per day for each clock.

As will be described later, the three H-masers are found to show significant differences in their performance when compared against NPL-CsF2. In particular, HM2 is found to be more stable than the other two H-masers at least for averaging times of 5 days or longer. Consequently, the data from HM1, HM2, and HM4 have been given different weights within the algorithm according to the variances exhibited by their data at different averaging times. For example, the weights are 4, 1, and 0.5 for HM2, HM4, and HM1 respectively for an averaging time of 2×10^6 s. Furthermore, HM4 is found to show a clear linear frequency drift. Consequently, its linear frequency drift has been initialized to -6×10^{-21} s^{-1} within the algorithm whilst a value of zero has been used for each of HM1 and HM2.

The active H-maser, HM2, that provides UTC(NPL), also referred to as Clock 2 in the results presented below, has been selected as the reference clock and the input data correspond to time interval counter measurements of (Clock 2 – C_i) where C_i is an individual clock.

The occasional frequency steers applied to UTC(NPL), or HM2, have been taken into account within the ensemble algorithm. This is achieved by entering the size of a steer in terms of the fractional frequency change (in units of ns/day) directly into the frequency offset component of the state vector of the clock at the relevant MJD of the steer. Consequently, the effects of the frequency steers on HM2, or Clock 2, have been removed in all of the input data and subsequent results presented in this paper.

RESULTS FOR AN 846-DAY DATA SET

Here we present a brief analysis of the algorithm using input data from HM1, HM2, and HM4 and two of NPL's commercial cesium clocks spanning an 846-day period from MJD 54904 to 55750. Figure 1 shows these data in the form of time offset measurements corresponding to (Clock 2 – C_i) where C_i corresponds to the i^{th} individual clock in the ensemble. Clocks 1, 2, and 3 correspond to HM1, HM2, and HM4 respectively, whilst Clocks 4 and 5 are cesium clocks. Clock 2 is shown as the reference clock. It can be seen from Figure 1 that there is an arbitrary offset in the Clock 2 measurements resulting from the measurement system. It is also clear that Clock 3 is exhibiting a significant linear frequency drift.

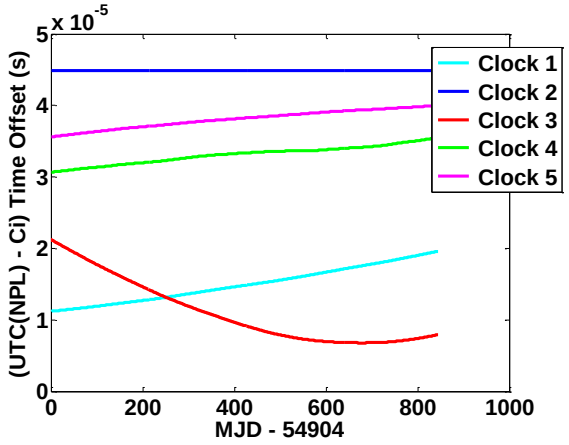

Figure 1. Plots showing the input data used in the algorithm in the form of time offset measurements (Clock 2 – C_i) where C_i corresponds to the i^{th} individual clock in the ensemble. Clocks 1-3 correspond to HM1, HM2, and HM4 respectively, whilst Clocks 4 and 5 are cesium clocks.

The relative clock stabilities of the (Clock 2 – C_i) measurements in terms of Allan Deviation (ADEV) are shown in Figure 2. The very low ADEV results for HM2 are a consequence of this clock being used as the reference clock, resulting in the ADEV in this case showing only the noise in the measurement system. The gradient of the ADEV plot for HM4 for averaging times τ greater than 10 days also suggests the presence of a significant linear frequency drift.

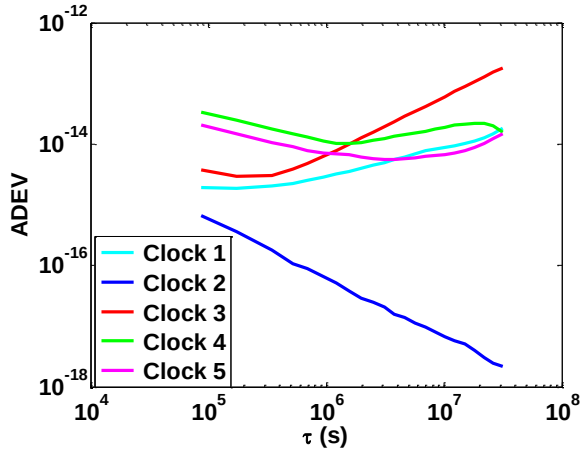

Figure 2. Plots showing the ADEV of the input data shown in Figure 1.

Figure 3 shows plots of the ADEV of the time offset between each clock and the composite time scale generated. As in the case of the input data, it is clear that HM4 exhibits a significant linear frequency drift. It can also be seen that the plots for HM1 and HM2 show ADEV values lower than 10^{-15} for τ values close to 10 days. This indicates that the composite is showing a good stability at these averaging times.

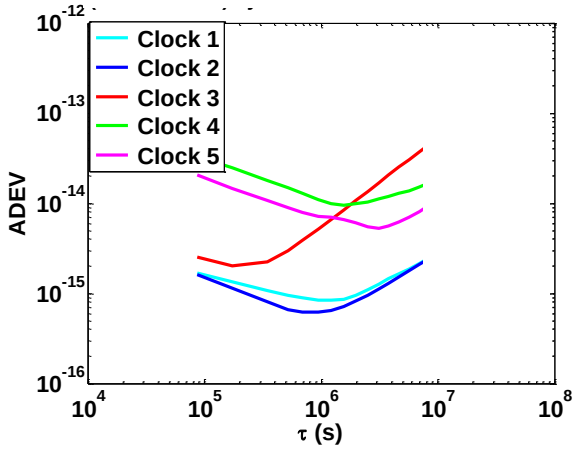

Figure 3. Plots showing the ADEV of the time-offsets between each individual clock and the composite.

Furthermore, Figure 4 shows plots of the HDEV which is insensitive to any linear frequency drift exhibited by the H-masers. This enables a better measure of the stochastic components of clock instabilities and hence the stochastic component of the instability of the composite. It can be seen that the plots for all three of the H-masers show HDEV values lower than 10^{-15} for τ values close to 10 days. This indicates that the composite is showing a good stability at these averaging times. However, in order to better assess the performance of the algorithm it is interesting to compare it against UTC or NPL's cesium fountain NPL-CsF2, both of which have excellent long-term stability. This is carried out in the following section.

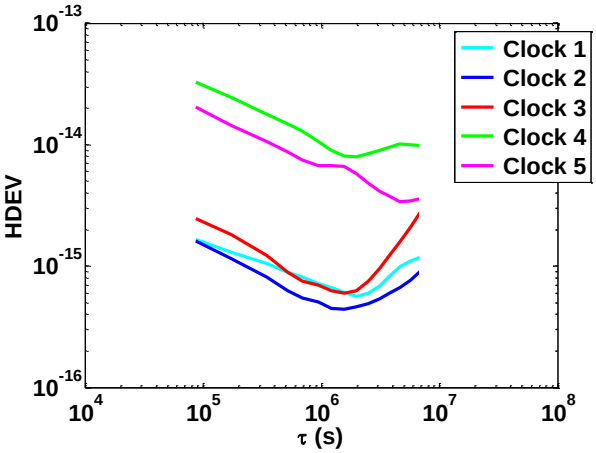

Figure 4. Plots showing the HDEV of the time-offsets between each individual clock and the composite.

COMPARISONS AGAINST UTC AND NPL'S CESIUM FOUNTAIN

Here we present an assessment of the composite time scale generated by the algorithm using firstly UTC and secondly NPL-CsF2. We concentrate on particular periods of data in order to carry out the comparisons in each case.

Comparisons of the composite against UTC are carried out using data spanning MJDs 55164 to 55439. NPL's Two-Way Satellite Time and Frequency Transfer (TWSTFT) system was used to provide traceability to UTC during this entire period. This link has a higher time and frequency transfer accuracy, as well as a greater stability in the medium term, than is available using GPS. Consequently, this data span provides better traceability to UTC for the purpose of assessing the stability of the composite.

Figure 5 shows ADEV plots of the time offsets between UTC and HM2 and UTC and the composite respectively. It is clear that both time scales are showing a similar stability relative to UTC. However, the observed stabilities are also very similar to the known stability of TWSTFT which is of order 10^{-15} at averaging times of one day or greater **[8]**. This raises the possibility that the stability of one or both of these time scales may be better than can be measured against UTC given the limitations of the TWSTFT link. Consequently, a similar assessment was carried out using NPL-CsF2. Data is available between MJDs 55084 and 55294 giving the measured frequency difference between NPL-CsF2 and HM2. In order to carry out an analysis consistent with that described above, these data have been averaged to obtain values at the Circular-T dates. Furthermore, these frequency differences have been converted to time offsets by arbitrarily assigning a time offset of zero between NPL-CsF2 and HM2 at the first measurement MJD. The frequency differences given by consecutive measurements are then used to calculate time-offset values by multiplying by the time interval, which in this case is the Circular-T interval of 5 days. There are some gaps in the data set ranging from durations of two days to seven days. Where necessary an interpolation has been carried out to obtain simulated data points at the Circular-T dates. The above procedure is essentially using data from NPL-CsF2 to construct a time scale.

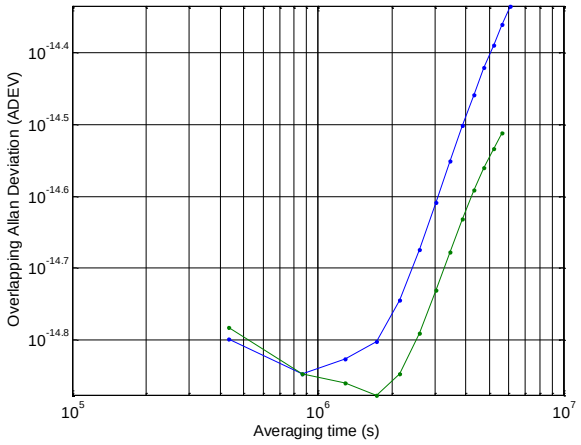

Figure 5. Plots showing the ADEV of the time offsets between UTC and HM2 (dark blue) and UTC and the composite (green).

Before comparing the composite against NPL-CsF2, we first of all compare the frequency stabilities of HM1, HM2, and HM4 against NPL-CsF2. HDEV is used for these comparisons as this is insensitive to any linear frequency drift that may be present in any H-maser, thereby enabling a measure of the stochastic component of the instabilities. Figure 6 shows plots of the HDEV of the time offsets between

NPL-CsF2 and each of the three masers. It is clear that the highest stabilities are observed in HM2, followed by HM4, and finally HM1, with HDEV values for an averaging time of 2 $x10^6$ s given by $5.4x10^{-16}$, 10^{-15}, and $1.4x10^{-15}$ respectively. Weights of 4, 1, and 0.5 are used to take account of the different performances of HM2, HM4, and HM1 respectively at this averaging time within the algorithm.

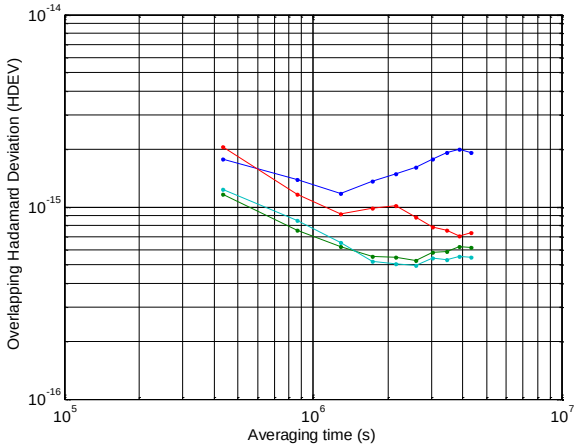

Figure 6. Plots showing the HDEV of the time offsets between NPL-CsF2 and HM1 (dark blue), HM2 (green), HM4 (red), and the composite (pale blue).

Figure 6 also shows a plot of the HDEV of the time offsets between NPL-CsF2 and the composite generated using hourly data for each clock. Its stability is found to be very similar to that of HM2, the most stable H-maser once its linear frequency drift has been removed. Given the weights assigned to the H-masers, the composite would be expected to show a relatively small improvement of approximately 10% compared to the stability of HM2 **[9]**. There is an indication of a marginal improvement consistent with this for averaging times of 20 days or greater. However, the stability at lower averaging times is worse than that of HM2. This has not been observed before using simulated data **[6]** and may suggest further refinements are required to the initialization of clock parameters. This is currently being investigated.

EXPLOITING NPL-CSF2 IN THE ALGORITHM

In this section, a preliminary investigation is carried out of a method for exploiting NPL's cesium fountain to improve the stability of the composite time scale. The aim is to construct a composite time scale with the excellent stability of NPL-CsF2 in the long-term and the stability of the H-masers in the short term.

The algorithm uses input data in the form of time offset values between each clock and the reference clock, in this case HM2. Consequently, frequency differences between NPL-CsF2 and HM2 have been converted to time offsets as described in the previous section. This procedure is essentially using data from CsF2 to construct a time scale, enabling NPL-CsF2 to be treated equivalent to a sixth clock within the algorithm. Noise parameters within the algorithm are used to model the stability of NPL-CsF2 according to a type B uncertainty of $2.3x10^{-16}$ and considering also its type A uncertainty **[10]**. Figure 7 shows the input data used in the algorithm in the form of time-offset measurements relative to HM2. Clock 6 represents data derived for an equivalent sixth "clock" using data from NPL-CsF2. Figures 8 and 9 show ADEV and HDEV plots respectively of the time-offset data between the generated composite time scale

and each of the six clocks represented in Figure 7. It can be seen from Figure 8 that the ADEV values for HM2, represented by Clock 2, and NPL-CsF2 relative to the composite are less than 10^{-15} for averaging times of approximately 10 days.

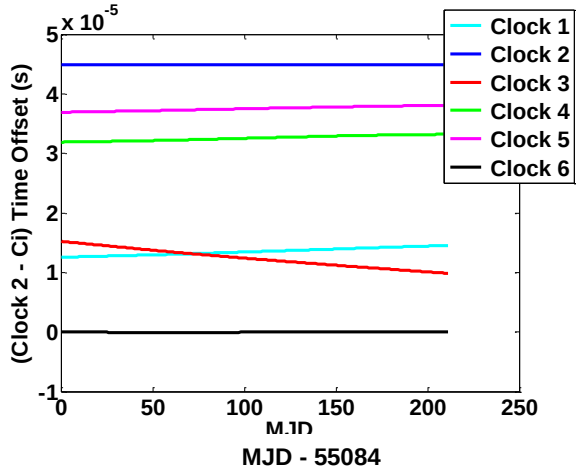

Figure 7. Plots showing the input data used in the algorithm in the form of time-offset measurements (Clock 2 – C_i) where C_i corresponds to the i^{th} individual clock in the ensemble. Clocks 1-3 correspond to active H-masers, HM1, HM2 and HM4 whilst Clock 4 and 5 are cesium clocks. Clock 6 represents data for an equivalent sixth "clock" derived using data from NPL-CsF2.

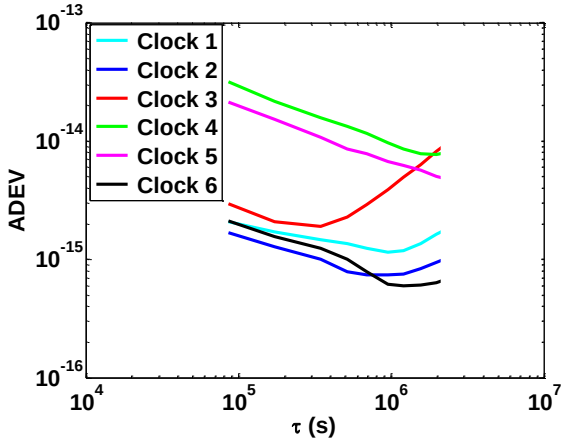

Figure 8. Plots showing the ADEV of the time-offsets between each individual clock and the composite.

Furthermore, Figure 9 shows that the HDEV values for each of the H-masers and NPL-CsF2 are less than 10^{-15} relative to the composite, suggesting that the composite is showing a good underlying stability. As expected, the HDEV values are lowest for HM2 and NPL-CsF2 relative to the composite, suggesting that the underlying stability of the composite is of order $5x10^{-16}$ at averaging times of between 10 and 20 days. The stability is found to be closer to that of NPL-CsF2 for

averaging times of 15 days or greater showing that the composite benefits from the excellent long-term stability of NPL-CsF2.

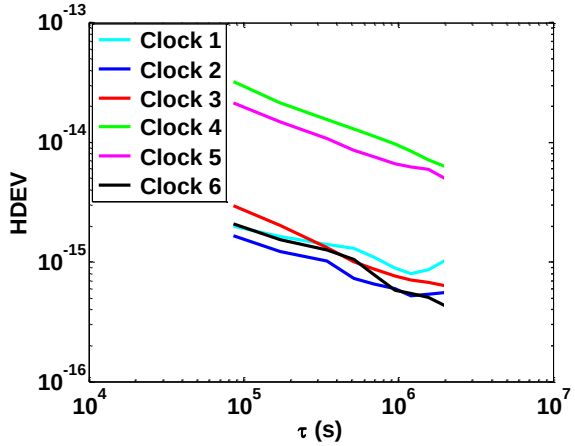

Figure 9. Plots showing the HDEV of the time-offsets between each individual clock and the composite.

A POTENTIAL STRATEGY FOR UTC(NPL) AND FURTHER DEVELOPMENTS

The suitability for generating UTC(NPL) using the approach described in the previous section could be explored further. The benefit of this approach would be that it not only uses an ensemble of H-masers and cesium clocks but also NPL-CsF2 to exploit its excellent long-term stability. The data from an additional H-maser, currently available at NPL, have not yet been used within the algorithm. This data would support further testing and, at least to a small degree, improve the stability of the composite. Furthermore, if in future data become available from a second fountain at NPL then these could also be exploited to further improve the stability of the composite.

Another potential benefit of an ensemble algorithm could arise from a lower latency for generating UTC than available from the present monthly Circular T publication. This is currently being investigated by BIPM. If this was implemented, the typical [UTC-UTC(*k*)] offset of a laboratory might be determined by the predictability of its reference clock over a much shorter period of 8 days rather than the current 40 days **[11]**. An ensemble algorithm might be able to exploit the high short-term stability of a number of H-masers to maintain the [UTC-UTC(*k*)] offset as low as possible between Circular T publications.

Some aspects of the algorithm would require further investigation and development. For example, the results discussed in the previous section were generated using regularly spaced data from NPL-CsF2. The algorithm needs to be developed in order to be able to cope with periods where there are significant data gaps. Furthermore, it also needs to be able to maintain close alignment to UTC through monthly steering of the composite according to the UTC-UTC(NPL) data published in Circular T. An aim of the algorithm is that the stability of the composite should be more robust to failures of hardware such as an H-maser. Consequently, the potential effects of the removal and addition of clocks should also be investigated.

CONCLUSIONS

Results have been presented on the composite time scale generated by NPL's clock ensemble algorithm using input data from three active H-masers and two cesium clocks spanning a period of 846 days. These indicate that the stability of the composite is 10^{-15} or better for averaging times of approximately 10 days. Comparisons of the composite have been carried out against UTC and NPL's cesium fountain, NPL-CsF2. It is found that the stability of the composite is similar to that of the most stable clock in the ensemble, HM2. This is consistent with the much higher weights given to HM2 within the ensemble. There is an indication of a lower performance of the composite than the most stable individual clock for some averaging times and this is being investigated.

A preliminary investigation has been carried out of an approach for exploiting NPL-CsF2 to improve the stability of the composite. In this case, the stability is found to be of order $5x10^{-16}$ at averaging times of between 10 and 20 days. It is found to be closer to that of NPL-CsF2 for averaging times of 15 days or greater demonstrating that the composite benefits from the excellent long-term stability of NPL-CsF2.

A potential strategy for generating UTC(NPL) using data from NPL-CsF2 within the algorithm has also been outlined. BIPM is currently investigating the possibility of reducing the latency between Circular T publications from 40 days to 8 days. If such a change was implemented, an ensemble algorithm might be able to exploit the high short-term stability of a number of H-masers to maintain the [UTC-UTC(*k*)] offset of a laboratory as low as possible between Circular T publications.

REFERENCES

[1] M. A. Weiss, D. W. Allan, and T. K. Peppler, 1989, *"A study of the NBS time scale algorithm,"* **IEEE Trans. Instrum. Meas.**, **Vol. 38**, 631-43.

[2] J. Levine, 2008, *"Steering a time scale,"* 40th Annual Precise Time and Time Interval (PTTI) Meeting, 1-4 December 2008, Reston, Virginia, pp. 205-218.

[3] M. Abgrall, P. Uhrich, and D. Valat, 2010, *"Ongoing improvements at the time and frequency references at LNE-SYRTE,"* European Frequency and Time Forum, 13-16 April 2010, Noordwijk, The Netherlands.

[4] L. A. Breakiron, 1991, *"Time scale algorithms combining cesium clocks and hydrogen masers,"* 23rd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 3-5 December 1991, Pasadena, California, pp. 297-305.

[5] L. A. Breakiron, 2001, *"A Kalman filter for atomic clocks and time scales,"* 33rd Annual Precise Time and Time Interval (PTTI) Systems and Planning Meeting, 27-29 November 2001, Long Beach, California, pp. 431-442.

[6] J. A. Davis, C. A. Greenhall, and P. W. Stacey, 2005, *"A Kalman filter clock algorithm for use in the presence of flicker frequency modulation noise,"* **Metrologia**, **Vol. 42**, 1−10, S154-S164.

[7] S. L. Shemar, J. A. Davis, and P. B. Whibberley , 2010, *"Preliminary results from NPL's clock ensemble algorithm using hydrogen masers and caesium clocks,"* European Frequency and Time Forum, 13-16 April 2010, Noordwijk, The Netherlands.

[8] D. Piester, A. Bauch, L. A. Breakiron, D. Matsakis, B. Blanzano, and O. Koudelka, 2008, *"Time transfer with nanosecond accuracy for the realization of International Atomic Time,"* **Metrologia**, **Vol. 45**, 185-198.

[9] J. A. Davis and B. Rueff, 2000, *"A study of correlated noise within a clock ensemble,"* in European Frequency and Time Forum, 14-16 March 2000, Torino, Italy, pp. 380-384.

[10] R. Li, K. Gibble, and K. Szymaniec, 2011, *"Improved Accuracy Evaluation of the NPL-CsF2 Primary Frequency Standard,"* **Metrologia**, **Vol. 48**, 283-289.

[11] P. B. Whibberley, J. A. Davis, and S. L. Shemar, 2011, *"Local representations of UTC in national laboratories,"* **Metrologia**, **Vol. 48**, S154-S164.