The Davis Instruments Vantage Pro2 wireless AWS – an independent evaluation against UK-standard meteorological instruments

By Stephen Burt MSc, PhD, FRMetS

Department of Meteorology, University of Reading, UK

Originally published online v1.0 2009

v1.1 March 2024

This independent review compares the performance of a Davis Instruments Vantage Pro2 Automatic Weather Station (AWS) against UK and Ireland-standard climatological instrumentation at a single site in central southern England over a 12 month period. Whilst other studies have assessed the performance of the air temperature measurements obtained from these and similar systems against US-standard synoptic and climatological measurements (Anon 1999, Karvelot 2007, 2008), to the best of the author's knowledge this is the only detailed and long-period comparative review of an entire system to be conducted against the tighter instrumentation standards of the UK climatological network.

Davis Instruments AWSs are the market leader in their category, which is towards the high end of systems within budget reach of most amateur meteorologists. But are the meteorological data obtained from them reliable and accurate, or are they simply expensive gadgets?

Important note: this review is entirely independent. The AWS was purchased privately from Prodata Weather Systems in Ely, Cambridgeshire, and was a standard 'off the shelf' package without special modifications or calibrations. The author has no connection with Davis Instruments or Prodata Weather Systems (other than as a customer of the latter), and no incentives were offered or sought to influence this review in any way.

Comparisons with UK standard meteorological instruments

The wireless Davis Instruments Vantage Pro2 automatic weather station (hereafter referred to simply as 'VP2') was exposed alongside existing climatological equipment at Stratfield Mortimer Observatory in Berkshire, UK (**Figure 1**, **Figure 2**) and logged data from both systems compared. The site is in central southern England (51.4°N, 1.0°W, altitude 60 m AMSL) and is rural and well-exposed, particularly between east and west through south: the instruments are located in a fenced enclosure. The observing location is an official Met Office and Environment Agency rainfall site, and

is regularly inspected by representatives from both organisations: it is also graded 'A' under the Climatological Observers Link (COL) station grading scheme.

Most of the equipment on site is automated, (cabled) sensors being connected at the time of the review to a Campbell Scientific CR10X datalogger which samples most of the sensors every second, logging data every minute, every 5 minutes and every hour. Daily summary totals, means and extremes are also generated at midnight UTC. All major meteorological parameters are measured, including solar radiation. The datalogging system and sensors closely resemble UK Met Office automated systems in use throughout the UK.

For most of the comparison period the VP2 was set up to log at 5 minute intervals to the interior console [1] (**Figure 3**), the exception being the fortnight commencing 29 July 2008 for which the logging period was 15 minutes. More details of the instruments used to compare with the output from the VP2 are given under each of the element headings. Records from the VP2 commenced on 12 July 2008, supplementing and overlapping an earlier Davis Instruments AWS, a Weather Monitor II model, which remained in continuous use from installation in February 1993 until its 'retirement' in October 2009.

The comparison period used for this review was the 12 months ending July 2009. An indifferent summer in 2008 was followed by a close-to-normal autumn and a colder than usual, and at times quite snowy, winter (the coldest for more than 10 years). The spring and early summer months of 2009 were warmer, drier and sunnier than normal, the later part of the summer being cooler and showery. The 12 months thus provided a fairly varied selection of southern England weather conditions, without any notable extremes which may have resulted in biased results one way or another.

Setup and installation

VP2 setup and installation was very straightforward – it took less than an hour from opening the box to obtaining readings. Documentation is clear, well laid-out and clearly indexed and there were no problems or questions during the 12 months trial that could not be resolved by reference to the accompanying documentation. It is obviously sensible to pre-prepare the installation, mounting masts etc, but wireless functionality makes installing such systems very much quicker and easier than cabled systems.

---

[1] The wireless transmitter for the sensor control unit was only 40 m or so from the display console/logger, located indoors, and thus well within the maximum permitted line-of-sight distance of around 300 m. Few transmission problems were evident during the trial, even in torrential rain or thunderstorm conditions. Anecdotal evidence from other VP2 users suggest that reception remains strong beyond even 400 m distance provided line-of-sight is maintained. Optional repeater units are available to extend the range still further, but obviously they were not required in this comparison.

Air temperature

Standard measurements of air temperature in the UK and Ireland, and in many other countries around the world, are made in Stevenson screens, white-painted wood or (more recently) white plastic double-louvred shelters which permit relatively free ventilation of the contents whilst protecting the instruments inside from both direct and reflected solar and long-wave radiation and from precipitation (Knowles Middleton 1966, Strangeways 2003). For many years liquid-in-glass thermometers located in Stevenson screens have been used to measure both current and maximum and minimum air temperatures, and while these are still in use today (**Figure 4**), the records of air temperature used for comparison with the VP2 were taken from a Vaisala HMP45C temperature and humidity probe (platinum resistance sensor) co-located within the Stevenson screen, sampled every second; maximum and minimum temperatures are taken as the highest and lowest respectively of (at the time) 30 second running mean air temperatures, subsequently amended to 60 s.

Accurate calibrations (to within 0.1 degC) both of backup thermometers and the Vaisala platinum resistance thermometer (PRT) were obtained by comparison with a portable calibrated reference source, a Tinytag TH-2500 thermohygrometer (Burt 2008b). Over a previous 12 month comparison period, the Vaisala sensor was found to reproduce the readings of the maximum and minimum liquid-in-glass thermometers within the screen to a monthly mean accuracy of better than 0.1 degC. It should be borne in mind, however, that calibration errors, although small, may mean that observed differences within about 0.2 degC may simply be within the margin of error of the sensors themselves.

Although the Stevenson screen has been the standard method of housing air temperature sensors for 125 years, it is not perfect: shelters of this design are known to overheat slightly in strong sunshine, particularly with light winds (Burton 2009). However, it remains the 'national standard' to the present day, and that is why it was used as the comparative benchmark.

The VP2 uses a proprietary design of naturally-ventilated white plastic radiation screen consisting of five 'inverted saucers' (**Figure 2**) housing the temperature and humidity sensors. In what is presumably a design based more upon compromise than climatological merit, the (black) tipping-bucket raingauge unit is mounted on top of the radiation shelter: unfortunately the two cannot easily be separated, as the raingauge would be better exposed at or close to ground level whereas it would be assumed that a representative air temperature environment would be better without a bulky black object sitting directly above the radiation shelter [2]. The VP2 sensors were mounted at the same height above ground, namely 1.25 m, as the sensors in the Stevenson screen. No calibration offset was

---

[2] That this has been previously commented upon and noted by Davis Instruments is evident from a sentence in Karvelot's (2008) paper, where he states, and I quote: "This study compared … to the Davis Instruments passive shield … because it was believed that NOAA would be more likely to accept an installation where the radiation shield is separated from the rain collector."

applied to the 'out of the box' reading from the VP2 temperature sensor. According to Davis Instruments' website specifications, the VP2 temperature values are updated every 10 seconds.

Comparisons between the VP2 and the Stevenson screen Vaisala PRT sensor of monthly mean and extreme temperatures over the 12 month period are given in **Table 1**, while **Table 2** details the frequency of the observed differences between the two sets of daily maximum and minimum temperatures. All 'daily' values quoted here relate to the civil day i.e. 00-00h UTC for ease of tabulation; this also avoids any 'observer proximity' terminal hour effects with the Stevenson screen dataset as can happen with comparisons made using 09-09h UTC data. Throughout, a positive (negative) difference is taken to mean that the VP2 was warmer (colder) than the Stevenson screen.

Over the 12 month period, the mean temperature indicated by the VP2 unit was only 0.07 degC different from that observed in the Stevenson screen. This insignificant difference is well within the margins of calibration error of the equipment used. The difference in mean temperature was slightly higher in the winter months (+0.12 degC in each of the three winter months) and lowest in October 2008 (+0.01 degC). This may be partly due to the slight near-linear calibration error evident on close analysis (from around 0.2 degC high at -7 °C, decreasing to 0.1 degC low above 30 °C) rather than any other reason. This derived calibration correction has not been retrospectively corrected for in this analysis; once known and corrected for it would be possible to improve future results still further.

A more useful climatological comparison is that of the logged daily maximum and minimum temperatures. Maximum temperatures tended to be slightly lower, and minimum temperatures slightly higher, than those in the Stevenson screen. This observed reduction in daily range is probably due to lag effects resulting from the relatively large thermal inertia of the combined radiation screen/ raingauge unit and adjacent solar cell/battery pack, than of the sensor unit itself (although it is not possible to ascertain the nature of the temperature/RH sensor surrounds without dismantling the entire unit). The temperature fall on clear nights was measurably slower in the VP2 than in the Stevenson screen, despite the much larger volume of the latter, probably as a result of the slight obstruction to outgoing long-wave radiation from the VP2 screen caused by the presence of the bulky raingauge unit mounted immediately above it. A re-design of the VP2 system to isolate the radiation shield from other components (thereby reducing its thermal mass) would help reduce these effects in future versions.

Differences in mean daily **maximum temperatures** ranged from +0.09 degC in December 2008 to -0.25 degC in July 2009, with a 12 month average of just -0.06 degC. The highest and lowest daily differences during the 12 months were +0.3 and -0.6 degC; over the 364 days with data, 274 days (75%) saw VP2 daily maxima were within 0.2 degC of those observed in the Stevenson screen. This is an impressive level of agreement. The slightly lower VP2 maxima observed in sunnier months may indeed owe as much to the slight overheating tendency of a Stevenson screen in strong sunshine and light winds. Slightly to my surprise, there was no evidence to suggest that the close proximity of the black plastic raingauge to the radiation screen generated spuriously high VP2 maxima – there was no

significant correlation between hours of sunshine, cumulative or peak solar radiation or maximum temperature with the difference between the two maxima. (The radiation screen and raingauge, although connected together, are separated by an air space and an insulation block.) The greatest differences tended to occur on days of broken sunshine and light winds, and thermal lag is likely the main cause. The hottest day of the 12 month period (1 July 2009) saw a VP2 maximum of 31.3°C compared to the screen maximum of 31.5°C.

Mean daily **minimum temperatures** tended to be slightly higher than those observed in the adjacent Stevenson screen – 73% of nights showed a VP2 minimum higher than the Stevenson screen, although the mean annual difference was again small at +0.09 degC. Eleven of the twelve months saw a positive difference in monthly means, although only two (June and July 2009) differed by as much as 0.2 degC. The highest and lowest daily differences during the 12 months were +0.7 and -0.3 degC; over 365 days with data, 291 (just under 80%) saw VP2 daily minima within 0.2 degC of those observed in the Stevenson screen. The coldest night of the 12 month period (7 January 2009) saw a VP2 minimum of -8.9°C compared to the screen minimum of -9.2°C.

These excellent results should be tempered with a note of caution. The 'out of the box' calibration of the sensors, although exemplary for the unit tested in this review, should not be assumed to be accurate for every system (Davis quote a typical accuracy of ± 0.5 degC). Systems can be ordered from reputable suppliers with calibration certificates for the temperature sensors (at extra cost), or if (as here) the system is run alongside existing accurately-calibrated thermometry for at least a few weeks any significant calibration error will quickly become apparent. Calibration checks should be repeated at least every year (preferably in both summer and winter seasons) because slow drift can occur: my previous Davis AWS in use since 1993 showed a slow negative drift of 0.2 degC over 15 years. There is a single-value 'calibration offset' setting on the software to adjust the raw data if required; this would be more useful if two or three calibration points could be allowed for.

It must also be borne in mind that an adequate exposure to the elements is required to provide representative air temperature readings, preferably (as here) mounted in an open location and above short grass (and certainly not in roof eaves as Davis Instruments themselves continue to illustrate in their marketing materials). A sheltered location or non-standard mounting is likely to introduce additional errors owing to microclimate effects. Non-standard exposure effects can quickly exceed minor calibration errors and produce results which bear little resemblance to standard measurements.

Relative Humidity (RH)

The VP2 uses a humidity sensor located within the radiation screen, with a nominal accuracy of ±3% RH, updated every minute. This was compared with the calibrated Vaisala capacitative humidity sensor on the HMP45C unit located in the Stevenson screen, which has a nominal accuracy of ±2% up to 90% RH and 3% between 90 and 100% and is sampled every second (10 second running means are used). Both units derive dew point from software using observed RH and air temperature.

Table 3 compares monthly mean RH and dew point between the VP2 and the Stevenson screen. An exact correlation is not to be expected, as capacitatance humidity sensors are not as accurate or repeatable as temperature sensors, but the agreement is still adequate for most purposes. Throughout the year the VP2 unit tended to read slightly higher, by a monthly average of 2-4% (broadly in line with the manufacturer's specification), while the derived dew point monthly means were mostly within 1 degC of the Vaisala sensor. The performance was not quite linear, the VP2 being less accurate in the lower humidity ranges: for example, on two days in April and May when the Vaisala-indicated minimum RH was 26% and 24% respectively, the lowest indicated by the VP2 was 32%. The slight non-linearity in performance is also evident from **Figure 5**, which plots 5 minute sampled RH for both instruments for a winter and summer month (January 2009 data in red and May 2009 in blue).

Rainfall

Reference has already been made to the VP2 raingauge, which is mounted above the radiation shield (**Figure 2**). The gauge is a black plastic unit with a circular aperture of approximately 165 mm and a reasonable knife-edge rim. Inside, there is a 0.2 mm tipping-bucket unit (the unit as supplied can tip at 0.2 or 0.25 mm, the metric value being selected at setup by a small mechanical adjustment to the unit combined with toggled software selection). Each tip is therefore only 4.3 cm^3 liquid content. When the air temperature/RH sensor is mounted at the standard 1.25 m above ground the rim of the raingauge is located at 1.57 m above ground level (the standard height for the rim of a raingauge in the British Isles is 30 cm).

Over the 12 month period comparison was made with both a standard 'five-inch' copper raingauge read manually at or close to 0900 UTC daily, and a Didcot Instruments 0.2 mm tipping-bucket raingauge located within the instrument enclosure. The rims of these instruments are at standard height and thus considerably lower than the elevated Davis gauge, and increased losses from wind eddying from the higher instrument were expected.

At setup, the calibration of the VP2 gauge was carefully checked and adjusted using the method described in *The Weather Observer's Handbook*, Chapter 15. This is an essential but often overlooked step, as the calibration of these instruments can and does vary significantly from the nominal 0.2 mm: under-reading by 10-20 per cent is not uncommon (the gauge tested was found to be reading approximately 12 per cent low 'out of the box'). All quoted comparison results are post-calibration. It should also be noted that it is vital to ensure the rim is perfectly level, as this can lead to significant under- or over-catch; this can be quite difficult to achieve and maintain for a pole-mounted unit, as this was, and requires regular checking.

Table 4 gives details of the catch of each of the three raingauges over the comparison period, while **Figure 7** shows the scatterplot of the daily values over the 12 month comparison period. (Four days data affected by snowfall in early February have been apportioned to allow for delayed snowmelt; one

day's total had to be estimated because of snowfall blowing out of the funnel.) Of course the VP2 gauge was practically useless in snowfall (**Figure 6**), but in this respect it is not unique [3].

Whilst it can be seen that the period totals agree well, the results from the VP2 for individual months and days are quite erratic, with no obvious or consistent pattern. Compare the VP2 scatterplot, **Figure 7**, with that from the higher-spec (and standard exposure) Didcot tipping-bucket raingauge, **Figure 8**, for an immediate visual comparison between the two.

For the 12 month period as a whole, the agreement with the standard checkgauge was excellent – indeed, much better than expected. For a checkgauge total of 631.1 mm (9% below the then 1971-2000 average annual rainfall at this site), the VP2 logged 642.6 mm, a difference of just 1.8% (higher). If the agreement was consistent, this would be an excellent result (albeit rather surprising, as a gauge at that height should record 5-10% less rainfall than a gauge exposed at the standard rim height of 30 cm). However, a closer look at **Table 4** reveals significant differences in monthly totals, ranging from 19% higher in August 2008 (90 mm vs 76 mm) to 10% lower in January 2009 (63 mm vs 69 mm). In contrast, the Didcot tipping-bucket raingauge total over the year was 635.4 mm, within 1% of the checkgauge total, and monthly totals were within ±2% of the checkgauge for 11 of the 12 months.

Several avenues were followed in an attempt to pin down some of the possible reasons for the relatively wide variation in recorded rainfall totals from the VP2 when compared with the standard gauge. There was no clear correlation with wind speed (as perhaps might have been expected given the height of the gauge). There was more evidence of a link with rainfall intensity – the daily totals on the 10 days with the highest rainfall rates [4] averaged +6% compared to the checkgauge (range -9% to +21%, daily totals 2.8 mm to 24.8 mm), whereas the 30 lowest daily rates averaged -7%. There is however a suggestion of a link with air temperature (**Figure 9**). The density of water varies little over normal air temperature ranges, but its viscosity (and thus presumably surface tension effects) reduces by some 50% between 0°C and 25°C: there is perhaps some evidence for incomplete emptying of the (small) bucket at lower temperatures. This finding would be worthy of further investigation.

In an attempt to identify whether the discrepancies also varied with the amount of rainfall (as might be expected from a device operating in discrete quanta), the variation of total fall by VP2 daily totals was examined, with the results shown in **Table 5** below.

---

[3] An optional tipping bucket raingauge heater is available from Davis Instruments to aid the measurement of frozen precipitation, but for most lowland districts of the British Isles significant snowfall will only rarely present difficulties.

[4] It is not stated in the manufacturer's documentation how rainfall rates are derived, but they were often 50-100% higher than those derived from the adjacent tipping-bucket raingauge.

This table did throw up another interesting question – why is there such a large 'overcatch' with days where only one tip (0.2 mm) was recorded? (Note that all 'spurious tip' events – such as the gauge being nudged whilst mowing the grass, or shaking in gusty winds – were removed from the database prior to analysis.) Only 13 of the 50 days when the VP2 gauge recorded 0.2 mm precipitation (10.0 mm recorded in all during the year) recorded 0.2 mm or more in the checkgauge - although a further 8 recorded 0.1 mm – with the checkgauge recording just 4.0 mm on those 50 days. The VP2 therefore recorded 250% of the checkgauge total in this 50 day sample. Why? I suspect two reasons: firstly, based upon direct observation, the VP2 gauge funnel is more easily wetted by dew than the checkgauge, and a good number of these are probably heavy dewfalls, not showing up in the checkgauge record as anything more than 'Dew trace'. (There is further evidence on this in that 39 of the 50 occurred between September and March.) Secondly, some days with just under 0.2 mm remaining in one bucket from a previous day's fall could therefore record a single tip with < 0.1 mm of precipitation (or dewfall); although it has to be said that there is probably a similar likelihood of any remaining water in the bucket evaporating completely, particularly during the summer months. (The Didcot gauge figures, shown in the table for comparison, do not show this marked 'single tip anomaly'.) At 2 tips or more, the VP2 performance in each class, at an *average* of 1-4% higher than the checkgauge, was extremely consistent up to the highest observed daily fall (27 mm) apart from a rather low 93% between 2.2 and 3.0 mm - although some of this is explainable through snowfall undercatch in February 2009, and over a wider sample it would probably revert closer to the adjacent samples. If consistent across most or all falls, this could probably be dismissed as a minor calibration error. However, even outside the lower daily amounts (where clearly the quantum resolution of the 0.2 mm tipping bucket exerts a marked effect) the daily samples themselves varied occasionally outside ± 20% as can be seen from **Figure 7**. There was no clear weather-related cause or causes that could be identified for either under- or over-catch events, and the phenomenon appears to be an inherent but random variation in the unit's performance.

The 'single tip anomaly' also manifests itself in a greatly exaggerated frequency of 'rain days' (a 'rain day' is one on which 0.2 mm or more of rainfall falls during the period of measurement, usually 0900-0900 UTC). **Table 6** shows the monthly frequency of rain days for each gauge – note the 207 rain days in the 12 month comparison period for the VP2 compared with just 170 from the checkgauge.

Records from the two days with greatest discrepancy in falls between the VP2 and the standard raingauge were examined. On 11 August 2008 the VP2 recorded 14.6 mm while only 10.8 mm was measured in the checkgauge, a difference of +3.8 mm or 35%; on 22 January 2009, the VP2 logged 12.0 mm while the checkgauge showed 14.9 mm, a difference of -2.9 mm or -19%. Results from both days are shown graphically in **Figure 10**. Here the tipping-bucket raingauge record is from the Didcot gauge, which agreed with the checkgauge within 5% on both occasions. On the first occasion, the morning of 12 August, spells of intermittent moderate to heavy cyclonic rain fell 0200-0700 UTC; the second occasion, 22/23 January 2009, saw a long spell of mostly moderate rainfall. Wind speeds were fairly light and cyclonic in nature during the precipitation on both occasions. It can be seen that the

timing record from the VP2 is in perfect agreement with the Didcot gauge, but the relative catches are significantly different.

The comparison shows that the record from the VP2 raingauge, even when carefully calibrated prior to use, cannot be relied upon to provide daily or monthly rainfall totals accurately enough for climatological purposes. Like almost all mechanical autographic gauges, it is best employed for providing records of the time and intensity of rainfall. A standard manual checkgauge (in a standard exposure) should be used both for determining the 'climatological' rainfall and for checking the continued functioning of the tipping-bucket raingauge itself (that the funnel has not been blocked by bird-droppings, for example: although the height of the rim means it is not easy to check for blockages of this nature). This is a shame, as the raingauge performance lags far behind the rest of the automatic weather station sensors on this package. An optional higher-quality separate ground-mounted raingauge unit would be a significant step forward in future designs. However, it is easy enough to substitute a higher-quality tipping-bucket raingauge in place of the one provided with this system; any unit providing a pulsed output can easily be wired in instead.

Barometric pressure

The pressure sensor on the VP2 is integral to the display unit (**Figure 3**), and is thus normally mounted indoors. (The usual qualifications for barometers and barographs also apply to the location of this unit – it should not be placed where sunshine can fall on it, nor above heaters, etc). The documentation is clear on how to set the barometer to mean sea level [5]. For this comparison, I used a Setra Barometric Pressure Transducer (model 278: Campbell Scientific's model CS100, manufacturer's calibration certificate dated 21 January 2008). This sensor is also checked frequently against a co-located Precision Aneroid Barometer (PAB) and a mercury Fortin barometer. The Setra sensor is connected to the Campbell Scientific CR10X logger, also mounted indoors, and is sampled once per second and logged once per minute. Daily (00-00h) and monthly mean barometric pressures were compared based upon 5 minute spot values from each instrument (**Table 7**).

Davis Instruments quote the sensor's accuracy as ± 1.0 mbar, but in practice the sensor, once correctly set, maintained a much higher accuracy level. In the comparison of monthly mean pressures the greatest difference was only 0.3 mbar, and this only after deliberately refraining from 'trimming' and

---

[5] If you don't have an accurate barometer with which to set or check the sensor, a convenient and easy way to check the mean sea level pressure is set out in The Weather Observer's Handbook, Chapter 15, Calibration. It is best to perform checks on days with light winds, and to check the setting regularly (at least once per season). Adjustments should be made only when the values differ consistently by 0.2 mbar or more in one direction, as minor variations from day-to-day can be expected.

resetting the sensor for several months to assess any calibration drift over time. There was little evidence of any drift in calibration. With a calibration check against a reliable mercury or solid-state pressure sensor every month or so, monthly means reliable within 0.2 mbar are certainly achievable. All but a handful of daily means were within 1.0 mbar of the Setra value. Unlike my previous Davis automatic weather station, there was no evidence of a slight pressure-dependent error with the sensor used in the VP2.

On a sub-daily basis, the VP2 sensor exhibited exactly the same degree of sub-millibar variations indicated by the higher-sensitivity Setra sensor, with no evidence of lag or sluggishness in response. Over the 12 month period, the lowest value indicated by the Setra was 967.9 mbar, on 23 January 2009, while the VP2 indicated 967.7 mbar at this time: the highest was 1038.1 mbar, on 26 December 2008, the simultaneous reading from the VP2 being 1037.8 mbar.

The performance of the VP2 barometer tested can only be described as excellent. With regular checks to minimise the risk of a slow or seasonal calibration drift, it can be expected to provide accurate and reliable barometric pressure recordings for many years.

Wind speed and direction

The comparison of logged barometric pressures on this unit with a standard sensor was the easiest of all the comparisons in this trial; the checking of recorded wind speed and direction was the most difficult. In part this is because of the highly-variable nature of this element, and in part because of the difficulty in co-locating sensors.

The anemometer and wind vane on the VP2 are co-located on a protruding bracket (**Figure 2**) which is provided with fixings for either posts or masts. Both instruments are made of black plastic. During operation, wind direction and wind speed are sampled at 2.5 s intervals on the wireless VP2. For the first nine months of the trial, the anemometer/wind vane unit was mounted so that the cups were 1.4 metres above ground level (AGL). On 17 May a longer mast was fitted and the cups raised to 2.8 m AGL (**Figure 2**). As wind speeds increase rapidly with height near the ground, the two sets of data are not comparable but providing the discontinuity is allowed for the record does provide for a useful comparison with standard instruments.

In this case, 'standard' readings – as near standard as are obtainable at this site at least - were obtained from a Vector Instruments A100L2-PC3 anemometer and W200P wind vane exposed atop a dedicated mast on the south wall of a nearby building at 11.1 m AGL and 2.7 m above the apex of the roof-line. This was the best available exposure, but not unobstructed - the horizon is almost clear above mast level. Both sensors are sampled every second and logged every minute (logging vector mean wind direction, scalar mean wind speed and highest gust i.e. the highest 3 second running mean wind speed).

Because the rooftop mast is not easily accessible, the VP2 anemometer/wind vane unit had to be mounted in a position where it could be reached from the ground: a true side-by-side comparison was therefore not possible.

From tests with a sensitive hand-held anemometer in light winds, the starting speed of the VP2 anemometer was determined to be around 1.4 kn when new, the stopping speed close to 1.0 kn. The wind vane was slightly less sensitive, typically needing around 2 kn to respond. The Vector Instruments anemometer starting speed is around 0.3 kn, and stopping speed 0.2 kn; the wind vane responds down to about 0.5 kn. Above its starting speed, the VP2 anemometer record showed a good correlation with that of the standard anemometer, for both mean and gust speeds, although a correction factor is needed to compensate for the height at which the anemometer is exposed; such correction factors (see, for example, Met Office 1982) apply only to mean wind speeds and never to gusts as can be seen from the results obtained.

Table 8 tabulates monthly mean wind speeds, monthly means of the daily highest gust and the highest gust in each month logged by each anemometer during the 12 month period, while **Figure 10** shows the scatter of 5 min scalar mean wind speeds from both instruments over a three month period (June to August 2009), amounting to over 26 000 pairs of observation points.

While at 1.4 m AGL *monthly mean* wind speeds averaged around 35% of those at 11 m, when the anemometer height was increased to 2.8 m this increased to about 50%. As the upper graph of **Figure 10** shows, there is considerable scatter about this figure, which will of course vary with wind speed, atmospheric stability (lapse rate), seasonal variations in upwind roughness length and other factors, but the agreement is good enough to suggest that an approximate value for a 'monthly mean 10 m wind' can be derived from the VP2 records, using tables or equations to estimate a correction factor to be applied to the observed wind speeds at anemometer height. Of course, these will only apply where the anemometer is reasonably well-exposed, and will not be applicable in more restricted exposures with greater eddying or fluctuations in wind speed from one or more directions, particularly in the direction of the prevailing wind.

That the same correction factors do not apply to gust speeds can be clearly seen from both **Table 8** and **Figure 10** (lower graph). While at 1.4 m AGL (2.8 m AGL) *monthly mean* daily wind speeds averaged around 35% (50%) of those at 11 m, the mean daily and maximum monthly *highest gust speed* averaged about 70% (85%) of those recorded by the higher anemometer.

The performance of the VP2 anemometer when new was mostly excellent, but on three occasions during the first 12 months of operation the cups seized up (i.e. the starting speed became much higher than normal). This turned out to be easily corrected by slackening off the cups from their spindle

followed by an application of WD40 penetrating oil, followed by re-assembly and tightening [6], but the high frequency with which this was required (every 4-6 months) is a little concerning, as the records during the time the anemometer cups are seized will be low or even zero, and thus inaccurate. (The problem is not immediately apparent without regular visual checking that the cups are rotating as expected.) Such a frequent maintenance interval mandates that this instrument be mounted in a position where it can be reached for servicing relatively easily – it is certainly not a 'fit and forget' sensor. This echoes experience with anemometer units on my previous Davis automatic weather station, where starting speeds drifted slowly upwards over the years.

The performance of the VP2 wind vane was perfectly satisfactory and no problems were encountered during the trial period. No comparisons with 'standard' wind direction measurements from the 'standard' sensor are presented, because the accuracy of the logged wind direction is crucially dependent upon accurate alignment at installation together with a reasonable instrumental exposure. It is a pity that, although the wind direction can be displayed to a precision of 1 degree on the console (with care, the vane can be aligned within ± 5 degrees), the logged values through the Weatherlink software are stored only to the nearest of 16 compass points (NNE, NE, ENE, E etc), viz. a precision of 22.5 degrees. This is a limitation of the current logging/data archiving system design which Davis have used for some 20 years now, but in a future system redesign a 10 degree precision would be welcome, as would a software toggle on the logged output to switch between compass points (e.g. SSW) and degrees (e.g. 200°). Currently the only way to convert all of the compass points to degrees (for vector mean wind calculations for example) is by means of macro functions in a spreadsheet.

Performance, reliability and durability

There were few problems encountered during the 12 month trial. The console batteries unexpectedly ran down on 19 April, resulting in just over 11 hours data loss: the 'low battery' display is not very obvious and appears not to be displayed for very long before power runs out completely, so a careful and regular check is required to avoid the risk of lost data (although a 220 v AC mains adaptor is supplied as standard). Aside from this, occasional maintenance is all that is required. The radiation screen needs a wipe-over every month or so (for some reason it seems to attract an inordinate amount of spider webs, far more than the other instruments, and while these are unlikely to affect the temperature or humidity readings, they do gum up the raingauge tipping-bucket mechanism and cause the anemometer cups to stick in light winds.) Weekly visual/light maintenance checks are recommended for optimum performance and peace of mind, but the unit can be normally left running and logging happily on its own for long periods.

---

[6] Note however that Davis Instruments advise against doing this, as it can wear away the lubricating grease on the bearings, although no alternative method was suggested by the manufacturer.

Over the 12 month period, the only significant total loss of data was on 19 April as detailed above. In all, data availability was just under 99.9%.

One welcome addition to the Weatherlink software would be a facility to alter the terminal hour from the default midnight; this would greatly simplify the production of 'UK standard' 0900-0900 UTC climatological parameters such as maximum and minimum air temperatures, and rainfall totals. It should be noted however that some elements (such as mean wind speeds) are conventionally summarised by civil day (i.e. midnight to midnight) and that this summary option should therefore be element-specific.

A 12 month trial is clearly not long enough to assess long-term reliability and durability, but with appropriate maintenance there is every reason to believe the system should last for at least 10 years outdoors in southern England. (A more testing climatic regime might prove more challenging.) This system replaced a previous Davis Weather Monitor II AWS, most components of which were still performing well when the system was retired after 16½ years exposure to the elements.

AUTHOR'S NOTE, March 2024: The Vantage Pro2 system described above is still in reliable daily operation, almost 15 years after installation. During this period, the temperature/humidity sensors have had to be replaced twice, while the anemometer performance has deteriorated, with starting speed often over 4 kn. The exterior cosmetic finish of the instrument has deterioted considerably after many years of exposure to ultraviolet in sunshine, although this does not appear to have significantly affected the instrumental readings.

Evaluation summary

The table below summarises ease of use and performance during this trial across 12 parameters between them evaluating the ease and accuracy of providing normal daily and sub-daily climatological data when compared with conventional climatological instrumentation. Of course, an AWS provides very much higher resolution data than standard manual instruments - the VP2 will log down to 1 minute intervals (although 5 minute logging is probably the best balance between resolution and data volume), compared with once-daily for standard instruments. Individual users will have differing priorities by element and may wish to assign greater or lesser importance to some parameters in making an overall evaluation of whether such a system will meet their requirements.

Element	System performance							Against conventional instruments			Notes
	Not available	Very poor	Poor	Adequate	Good	Very good	Excellent	Did not meet requirements	Meets requirements	Surpassed requirements	
Setup and installation							O		O		
Ease of use, system documentation							O		O		
System displays							O			O	
Air temperature measurements							O		O		*Subject to calibration checks*
Relative Humidity measurements				O					O		
Rainfall – daily totals			O					O			*Standard raingauge required for accurate rainfall totals*
Rainfall – timing						O			O		
Barometric pressure							O			O	*Subject to regular calibration/neighbour checks*
Wind speed – mean and gusts					O				O		*Note that comparisons could not be made at the same height AGL*
Wind direction						O			O		
Reliability						O				O	
Value for money						O				O	
SUMMARY COUNTS			1	1	1	1	8	1	7	4	

COL station grading assessment

Subject to calibration checking at or before installation (and regular checks at 2-3 year intervals thereafter), on the basis of the results from this trial there is no reason why air temperature measurements made from a VP2 in a satisfactory exposure should not be accepted as 'A-grade' instruments in the Climatological Observers Link 2008 grading scheme (Burt 2008a). However, the rainfall measurements rate the lower 'C' grade on the basis of this comparison.

Conclusions

The Davis VP2 automatic weather station was evaluated alongside conventional climatological instrumentation at a site in central southern England over a 12 month period. Almost all sensors were found to provide data closely comparable with those from UK-standard climatological instruments. Systematic errors on air temperatures and barometric pressures were very small, while relative humidity/dew point and wind speed and direction data were mostly good. This system is let down by its raingauge, which provided somewhat erratic precipitation measurements (although indications of timing and intensity were excellent): this deficiency is easily rectified by installing a standard five-inch raingauge on the same site and using the latter for 'climatological' rainfall totals.

In the author's opinion the Davis VP2 automatic weather station represents very good value for money, and given reasonable care and occasional maintenance there should be no reason why it should not provide at least 10 years record in all but extreme climates.

Acknowledgements

References

Anon (1999) A study of radiation shield effectiveness. Factsheet issued by Davis Instruments, Hayward, California, USA, dated 30 Nov 1999, comparing active and passively ventilated radiation shields (available from sales@davisnet.com)

Burt, Stephen (2008a) *The new COL station grading system.* Presentation at Royal Meteorological Society - Observing Systems Special Interest Group meeting, Imperial College, London, 14 May 2008. This system was introduced into the COL Bulletin during 2008

Burt, Stephen (2008b) Calibration of air temperature sensors. *Climatological Observer's Link Bulletin* No. 460, August 2008, pp 30-32

Burton, B J (2008) – *Stevenson screen temperature errors – an investigation.* Personal correspondence.

Karvelot, Jason (2007) *New co-operative observer networks and instrumentation data quality.* Davis Instruments presentation at 23rd Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, 23rd, San Antonio, TX, 14-18 January 2007. American Meteorological Society, Boston, MA, 2007

Karvelot, Jason (2008) *New co-operative observer networks and instrumentation data quality – an update.* Davis Instruments presentation at 12th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS), New Orleans, LA, 20-24 January 2008. American Meteorological Society, Boston, MA, 2008

Knowles Middleton, W E (1966) *A history of the thermometer and its use in meteorology*. Chapter X, The exposure of thermometers. Johns Hopkins Press, Baltimore

Met Office (1982) *Observer's Handbook*. HMSO. Formulae and tables for correcting wind speed with height are given in Section 5.2.2, page 83

Strangeways, Ian (2003) *Measuring the natural environment.* Cambridge University Press. Chapter 3, temperature.

Table 1 - *monthly mean and extreme temperatures over the 12 comparison month period. A positive (negative) value indicates that the VP2 was higher (lower) than the Stevenson screen value. Monthly means are given to 2 decimal places to facilitate comparisons, although it should be noted that the accuracy of the sensors is no better than 0.1 degC. Maximum and minimum are for 00-00h terminal periods. Mean temperature is the average of all automatic weather station samples over the 00-00h period, not the average of maximum + minimum. VP2 Data availability was 99.9%.*

TABLE 1 - Temperature comparisons

Year	Month	Mean maximum °C			Mean minimum °C			Mean temperature °C			Highest maximum °C			Lowest minimum °C		
		Stevenson screen	VP2	Diff	Stevenson screen	VP2	Diff	Stevenson screen	VP2	Diff	Stevenson screen	VP2	Diff	Stevenson screen	VP2	Diff
2008	August	20.50	20.46	-0.04	12.47	12.50	+0.03	16.26	16.32	+0.07	26.2	25.9	-0.3	5.7	5.7	-0.0
	September	18.52	18.47	-0.05	8.60	8.65	+0.04	13.12	13.18	+0.05	20.9	20.9	+0.0	2.8	2.8	+0.0
	October	14.96	14.83	-0.13	4.18	4.25	+0.07	9.58	9.59	+0.01	22.9	22.7	-0.2	-2.2	-1.9	+0.3
	November	10.72	10.70	-0.02	3.74	3.79	+0.05	7.48	7.55	+0.07	14.6	14.5	-0.1	-3.5	-3.4	+0.1
	December	7.08	7.16	+0.09	0.85	0.95	+0.09	3.80	3.92	+0.12	12.2	12.2	+0.0	-5.9	-5.8	+0.1
2009	January	6.33	6.38	+0.05	-1.01	-0.88	+0.13	2.71	2.84	+0.12	10.3	10.2	-0.1	-9.2	-8.9	+0.3
	February	7.80	7.81	+0.01	1.11	1.19	+0.07	4.32	4.43	+0.12	14.9	14.4	-0.4	-5.7	-5.8	-0.1
	Match	12.90	12.87	-0.03	1.26	1.30	+0.04	7.02	7.10	+0.10	16.7	16.4	-0.3	-3.7	-3.6	+0.1
	April	15.74	15.71	-0.03	4.64	4.61	-0.03	10.19	10.24	+0.08	20.7	20.5	-0.2	-0.2	-0.3	-0.1
	May	18.16	18.07	-0.09	6.48	6.57	+0.09	12.66	12.71	+0.05	25.4	24.9	-0.5	0.2	0.2	+0.0
	June	21.66	21.45	-0.21	8.60	8.84	+0.24	15.52	15.57	+0.04	29.8	29.5	-0.3	2.6	2.8	+0.2
	July	21.82	21.58	-0.25	11.49	11.69	+0.20	16.44	16.47	+0.06	31.5	31.3	-0.1	8.5	8.6	+0.1
	12 months	**14.68**	**14.62**	**-0.06**	**5.20**	**5.29**	**+0.09**	**9.93**	**9.99**	**+0.07**	**31.5**	**31.3**	**-0.1**	**-9.2**	**-8.9**	**+0.3**

Table 2 - *frequency of the observed differences (irrespective of sign) between the two sets of daily maximum and minimum temperatures over the 12 month period. Maximum and minimum relate to 00-00h terminal periods. A positive (negative) mean value indicates that the VP2 was higher (lower) than the Stevenson screen mean..*

TABLE 2

Percentage of days within these tolerances

	Mean diff (degC)	Days	Within **0.10** degC	Within **0.20** degC	Within **0.30** degC	Within **0.40** degC	Within **0.50** degC	Positive diff	Max diff (degC)	Min diff (degC)
MAXIMUM TEMPERATURE	-0.06									
Number of days		364	166	274	326	353	362	142	+0.4	-0.6
Cumulative frequency %			46	75	90	97	99	39		
MINIMUM TEMPERATURE	+0.09									
Number of days		365	163	291	344	359	362	266	+0.7	-0.3
Cumulative frequency %			45	80	94	98	99	73		

Table 3 – *comparison of monthly mean humidity parameters over the 12 month period: mean daily RH (%), mean daily minimum RH (%), and mean daily dew point temperature (°C). Terminal hours are 00-00h in all cases. A positive (negative) mean value indicates that the VP2 was higher (lower) than the Stevenson screen mean..*

TABLE 3 - Humidity comparisons

Year	Month	Mean daily RH, %			Mean daily min RH, %			Mean dew point temp °C		
		Stevenson screen	VP2	Diff	Stevenson screen	VP2	Diff	Stevenson screen	VP2	Diff
2008	August	81	84	+3	59	65	+6	12.9	13.5	+0.6
	September	82	84	+2	56	62	+6	9.8	10.3	+0.6
	October	85	87	+2	60	66	+6	6.9	7.3	+0.5
	November	87	90	+2	74	79	+5	5.4	5.9	+0.5
	December	88	90	+3	74	80	+6	1.9	2.4	+0.5
2009	January	88	90	+2	72	78	+6	0.8	1.3	+0.6
	February	84	87	+3	66	73	+6	1.7	2.4	+0.7
	March	76	80	+4	50	57	+8	2.8	3.7	+0.9
	April	77	81	+4	50	58	+7	5.9	6.6	+0.7
	May	74	77	+4	48	55	+7	7.5	8.4	+0.9
	June	71	74	+4	42	50	+7	8.6	10.5	+1.9
	July	76	80	+4	48	55	+8	11.8	12.6	+0.8
	12 months	**81**	**84**	**+3**	**58**	**65**	**+7**	**6.3**	**7.1**	**+0.8**

Table 4 – comparison of rainfall observations over the 12 month period. The monthly total and highest daily rainfall in each month are given for the co-located standard 'five-inch' raingauge (rim at 30 cm), for the Davis VP2 tipping-bucket raingauge (rim at 1.57 m) and, as comparison with the latter, from a higher specification (Didcot) 0.2 mm tipping-bucket raingauge (rim at 42 cm). Values are given in millimetres and as a percentage of the checkgauge value. Terminal hours are 0900-0900 UTC throughout.

TABLE 4 - Rainfall comparisons

Year	Month	Total rainfall mm			% vs standard gauge			Wettest day mm			Delta vs standard gauge		
		Standard 5" gauge	VP2 TBR	Didcot TBR	Standard 5" gauge	VP2 TBR	Didcot TBR	Standard 5" gauge	VP2 TBR	Didcot TBR	Standard 5" gauge	VP2 TBR	Didcot TBR
2008	August	75.9	90.0	78.6	100	119	104	14.9	17.2	14.8	100	115	99
	September	53.1	57.2	54.2	100	108	102	13.4	15.6	13.8	100	116	103
	October	59.4	58.0	60.2	100	98	101	17.5	17.4	17.5	100	99	100
	November	75.5	69.4	76.3	100	92	101	18.8	16.8	18.3	100	89	97
	December	30.7	27.8	30.0	100	91	98	11.3	10.2	11.0	100	90	97
2009	January	69.4	62.6	68.0	100	90	98	15.2	14.6	15.0	100	96	99
	February	61.7	62.8	61.0	100	102	99	26.7	27.0	25.8	100	101	97
	March	28.6	29.8	28.6	100	104	100	14.7	15.4	14.4	100	105	98
	April	31.1	30.6	30.9	100	98	99	4.5	4.8	4.7	100	107	104
	May	27.0	27.2	27.6	100	101	102	9.0	9.2	8.5	100	102	95
	June	22.1	23.8	22.5	100	108	102	11.5	12.4	11.8	100	108	102
	July	96.6	103.4	97.4	100	107	101	21.0	20.4	20.3	100	97	97
	12 months	**631.1**	**642.6**	**635.4**	**100**	**101.8**	**100.7**	**26.7**	**27.0**	**25.8**	**100**	**101**	**97**

Table 5 – comparison of tipping-bucket raingauge catch by daily rainfall amount over the 12 months ending July 2009. The total rainfall by each gauge for samples of VP2 daily rainfall totals are presented, together with the mean for the category. The number of observations in each category is indicated by n. Values are given in millimetres and as a percentage of the checkgauge total fall. Terminal hours are 0900-0900 UTC throughout.

TABLE 5 - Rainfall comparisons

Rainfall by VP2	No of 0.2 mm tips	Total rainfall mm			n	Mean daily rainfall mm			Total % vs standard gauge		
		VP2 TBR	Didcot TBR	Standard 5" gauge		VP2 TBR	Didcot TBR	Standard 5" gauge	Standard 5" gauge	VP2 TBR	Didcot TBR
0	0	0	2.1	1.7	158	0	0.01	0.01	--	--	--
0.2 mm	1	10.0	4.2	4.0	50	0.20	0.08	0.08	100	250	105
0.4 mm	2	7.6	7.1	6.9	19	0.40	0.37	0.36	100	110	103
0.6 mm	3	4.8	5.1	4.7	8	0.60	0.63	0.59	100	102	109
0.8 mm	4	9.6	10.2	9.2	12	0.80	0.85	0.77	100	104	111
1.0 mm	5	8.0	8.3	7.8	8	1.00	1.04	0.98	100	103	106
0.2 - 1.0 mm	1-5	40.0	34.8	32.6	97	0.41	0.36	0.34	100	123	107
1.2 - 2.0 mm	6-10	30.0	30.5	30.2	20	1.50	1.53	1.51	100	99	101
2.2 - 3.0 mm	11-15	61.8	67.9	66.1	24	2.58	2.80	2.72	100	93	103
3.2 - 4.0 mm	16-20	58.4	58.9	57.7	17	3.44	3.46	3.39	100	101	102
4.2 - 5.0 mm	21-25	36.0	35.7	34.9	8	4.50	4.46	4.36	100	103	102
5.2 - 10.0 mm	26-50	174.2	171.1	171.2	25	6.97	6.84	6.85	100	102	100
10.2 - 15.0 mm	51-75	96.4	92.6	93.0	8	12.05	11.57	11.63	100	104	100
15.2 - 30.0 mm	76-150	145.8	141.8	143.7	8	18.23	17.72	17.96	100	101	99
	TOTAL	642.6	635.4	631.1							

Table 6 – comparison of rain day and wet day frequencies by raingauge over the 12 month period. The monthly total are also given for the co-located standard 'five-inch' raingauge (rim at 30 cm), for the Davis VP2 tipping-bucket raingauge (rim at 1.57 m) and, as comparison with the latter, from a higher specification (Didcot) 0.2 mm tipping-bucket raingauge (rim at 42 cm). Terminal hours are 0900-0900 UTC throughout.

TABLE 6 - Frequency of rain days and wet days for the three raingauges. Period 0900-0900 UTC

Year	Month	Total rainfall mm			Rain days (≥ 0.2 mm)			Wet days (≥ 1.0 mm)		
		Standard 5" gauge	VP2 TBR	Didcot TBR	Standard 5" gauge	VP2 TBR	Didcot TBR	Standard 5" gauge	VP2 TBR	Didcot TBR
2008	August	75.9	90.0	78.6	23	23	23	16	19	17
	September	53.1	57.2	54.2	14	18	16	9	9	10
	October	59.4	58.0	60.2	14	23	18	9	9	10
	November	75.5	69.4	76.3	19	21	20	12	12	12
	December	30.7	27.8	30.0	11	13	10	5	4	5
2009	January	69.4	62.6	68.0	14	18	17	10	11	10
	February	61.7	62.8	61.0	12	15	13	10	10	10
	March	28.6	29.8	28.6	9	14	11	5	5	5
	April	31.1	30.6	30.9	13	14	13	11	10	11
	May	27.0	27.2	27.6	12	14	13	8	8	8
	June	22.1	23.8	22.5	8	10	8	5	5	5
	July	96.6	103.4	97.4	21	24	22	16	16	16
	12 months	**631.1**	**642.6**	**635.4**	**170**	**207**	**184**	**116**	**118**	**119**

Table 7 – Comparison of monthly mean MSL pressure obtained from the Davis VP2 compared with that from the Setra sensor. VP2 data was missing for 11 hours on 19 April 2009 owing to battery failure, and this period has been excluded from the analysis. Values are given in millibars (hPa). The figures given are the mean of daily means 00-00h. Averages are given to 2 decimal places only to facilitate comparison – the accuracy of the instrument is only ± 0.1 mbar.

TABLE 7 - Barometric pressure comparisons

Year	Month	Mean daily MSLP *mbar*		
		Setra	VP2	Diff
2008	August	1011.01	1010.70	-0.30
	September	1017.05	1016.95	-0.10
	October	1014.89	1014.79	-0.10
	November	1013.76	1013.63	-0.13
	December	1018.85	1018.64	-0.21
2009	January	1011.24	1011.03	-0.21
	February	1015.66	1015.66	+0.01
	March	1015.74	1015.48	-0.27
	April	1013.10	1012.93	-0.17
	May	1018.20	1018.08	-0.12
	June	1017.45	1017.58	+0.13
	July	1012.46	1012.54	+0.08
	12 months	**1014.95**	**1014.83**	**-0.12**

Table 8 – *monthly mean wind speeds, monthly means of the daily highest gust and the monthly highest gust in each month observed by both anemometers over the 12 month comparison period. Values are in knots. The VP2 % column gives the VP2 value as a percentage of the Vector value. The VP2 anemometer was at 1.4 m AGL until 17 May, and 2.8 m AGL thereafter; the Vector anemometer is at 11 m AGL*

TABLE 8 - Wind speed comparisons

Year	Month	Mean wind speed knots			Avge daily highest gust kn			Monthly highest gust		
		Vector	VP2	VP2 %	Vector	VP2	VP2 %	Vector	VP2	VP2 %
2008	August	5.0	2.2	44	21.2	14.0	66	37	24	66
	September	3.8	1.4	38	15.8	11.1	70	29	20	69
	October	4.4	1.8	41	18.3	13.9	76	30	24	79
	November	4.9	1.5	31	19.0	12.6	66	37	27	72
	December	4.2	1.3	31	16.1	11.4	70	27	21	77
2009	January	4.4	1.6	37	17.6	12.3	70	36	29	79
	February	3.6	0.8	23	13.9	9.7	70	29	23	79
	Match	4.8	1.9	40	18.4	13.8	75	35	24	68
	April	3.5	1.1	30	15.2	11.2	74	27	19	71
	May *	4.9	2.2	45	18.9	14.6	77	28	25	89
	June	3.2	1.4	44	14.7	12.6	86	30	25	83
	July	4.8	2.6	55	19.5	16.6	85	29	24	80
	12 months	**4.3**	**1.7**	39	**17.4**	**12.8**	74	**37**	**29**	77

* Instrument height changed from 1.4 m AGL to 2.8 m AGL on 17 May

Figures

Figure 1 – *the trials site at Stratfield Mortimer Observatory in central southern England, looking south-west. The Davis VP2 is located to the left of the Stevenson screen, an existing Davis Weather Monitor II automatic weather station is located to the right of the Stevenson screen. This photograph was taken on 24 May 2009, during the evaluation period.*

Figure 2 (left) – the Davis VP2 at the trials site at Stratfield Mortimer Observatory in central southern England. This photograph was taken on 24 May 2009.

Figure 3 (right) – The Davis VP2 display console

Figure 4 (brlow, right) – Conventional [large] Stevenson screen layout, showing maximum and minimum sheathed liquid-in-glass thermometers (mounted horizontally) together with the Vaisala HMP45C temperature and humidity sensor used in this comparison (grey cylinder, left of screen). This photograph was taken on 14 June 2009.
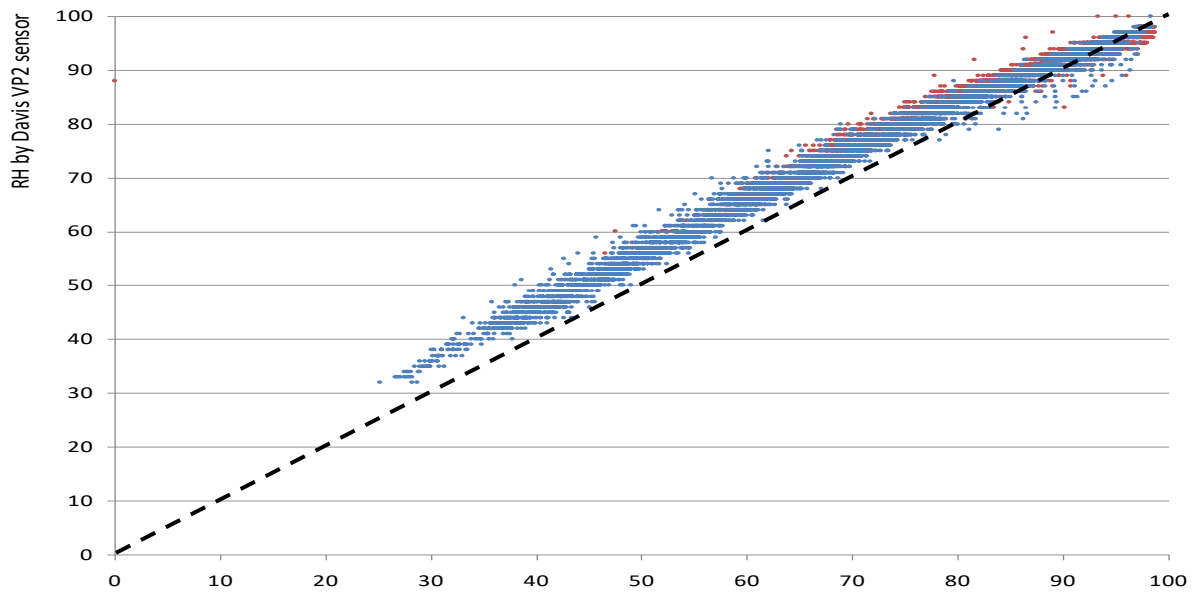
Figure 5 *– 5 minute sampled RH for both instruments for a winter and summer month (January 2009 data in red and May 2009 in blue): in all, 17 780 pairs of observations. The thick dashed line marks the trend line that would be expected with a 'perfect' agreement; points above the line indicate where the VP2 indicated higher than the Vaisala sensor, and vice versa below the line.*

Figure 7 (below) – scatterplot of the daily VP2 0.2 mm tipping-bucket raingauge catch vs the standard checkgauge values over the 12 month comparison period. The midpoint has a slope of almost exactly 1, but there is a large scatter in the daily observations – a few even falling outside the ±20% lines indicated by the orange dashed lines

Figure 6 (above) – the VP2 following a light snowfall (2 February 2009, 3 cm level depth at 0900 UTC)
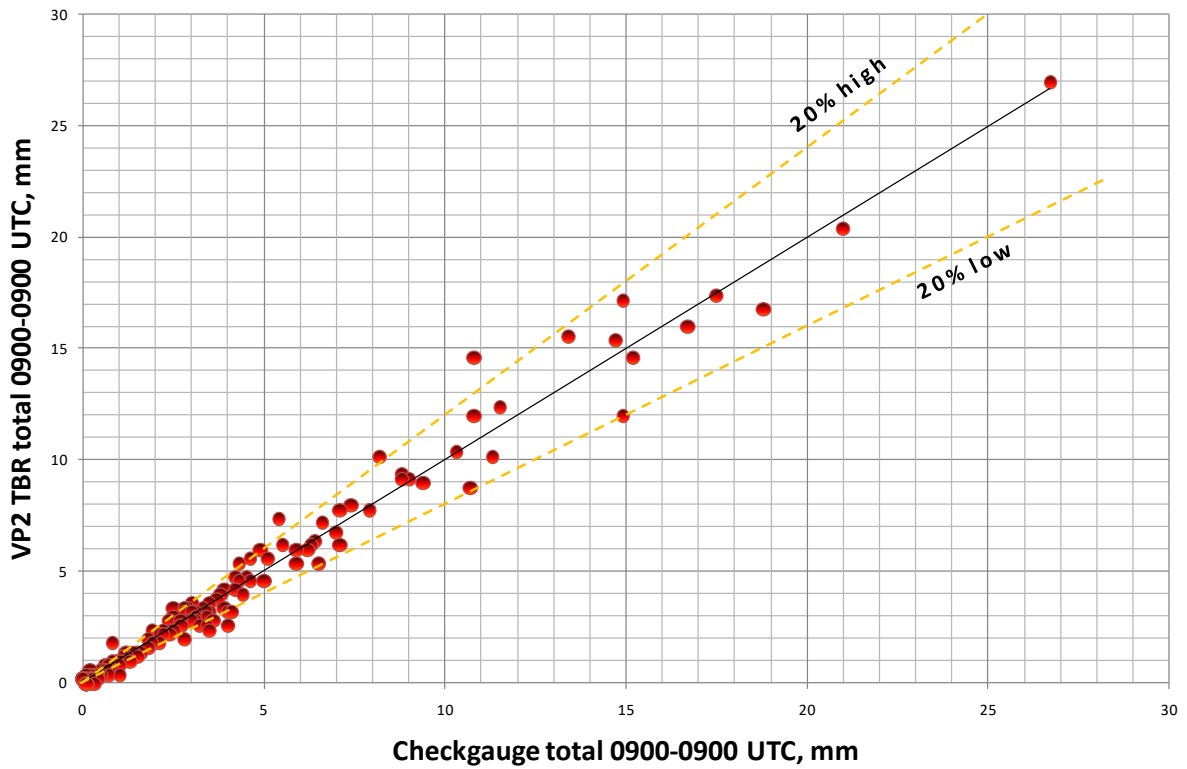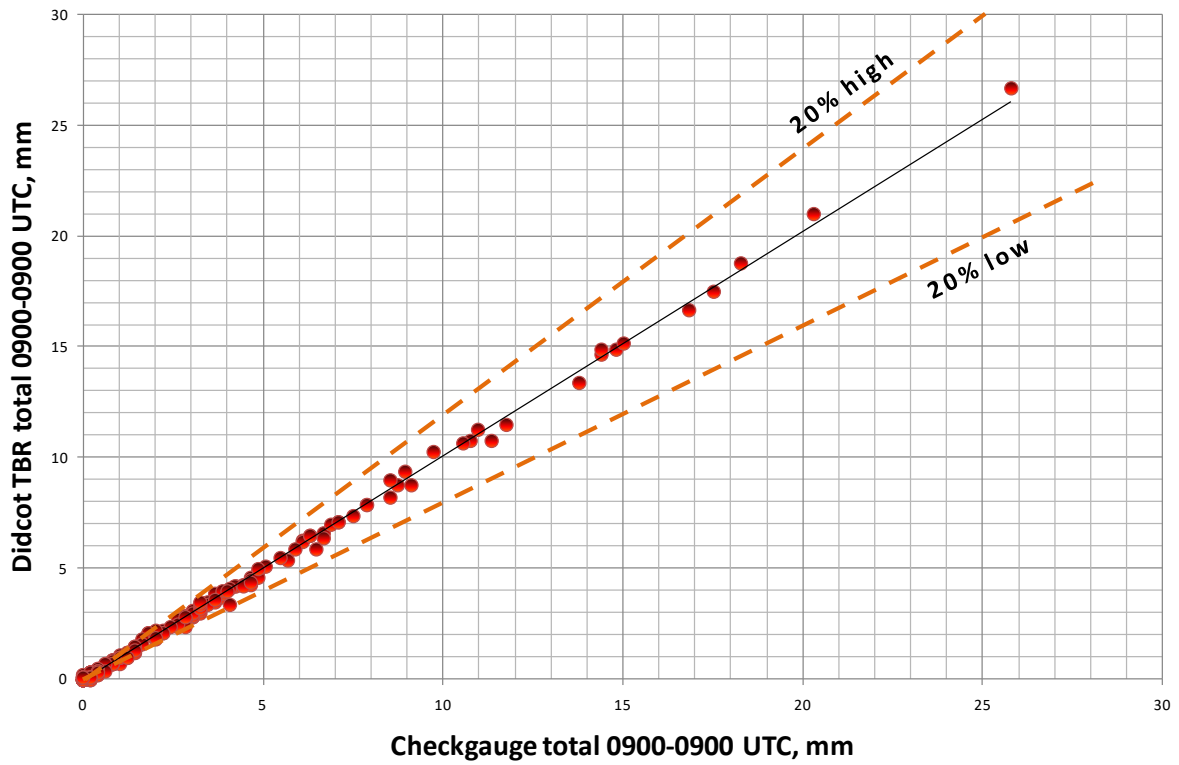
Figure 8 *– scatterplot of the daily Didcot 0.2 mm tipping-bucket raingauge catch vs the standard checkgauge values over the 12 month comparison period. Compare with the VP2 performance over the same period in Figure 7*
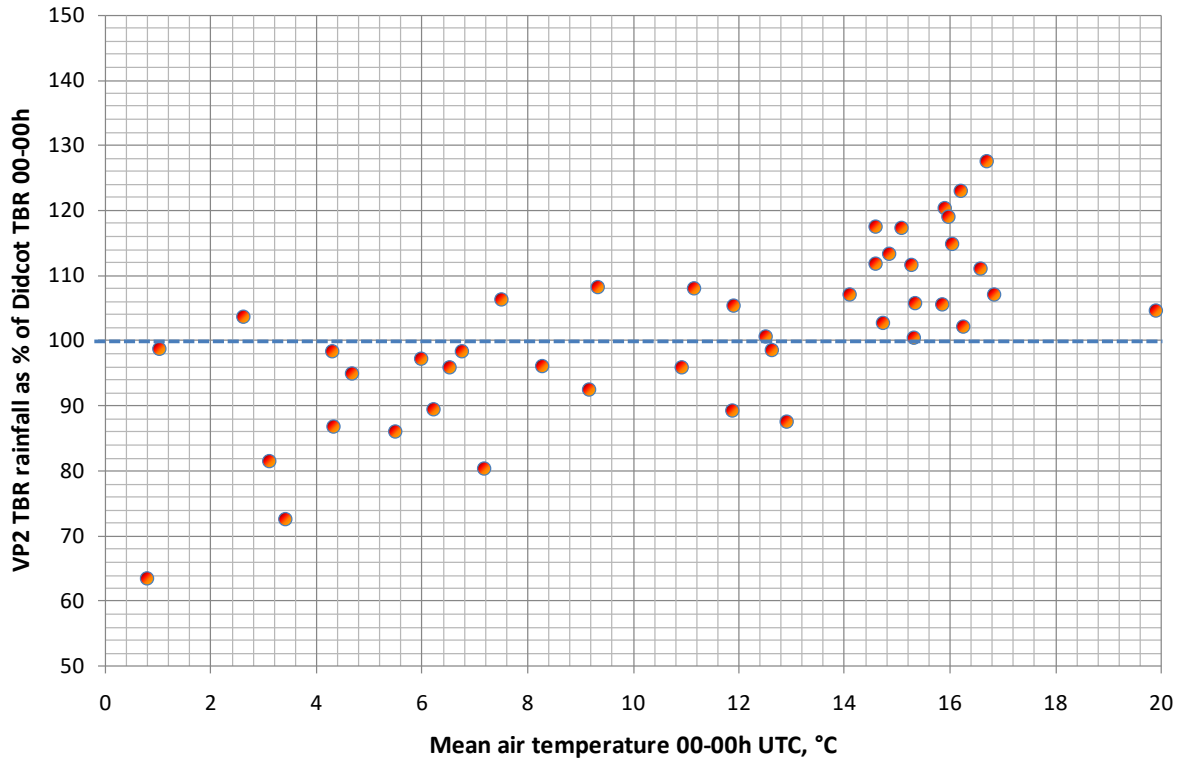
Figure 9 *– scatterplot of the daily VP2 tipping-bucket raingauge catch 00-00h as a percentage of the Didcot gauge values in the same period, plotted against the daily mean temperature 00-00h, during the 12 month comparison period. 100% is indicated by the thicker dashed blue line. VP2 rainfall totals in excess of 5 mm only.*

Figure 10 – *the variation of rainfall amount with time on the occasions during the 12 month comparison period of the greatest overcatch of the VP2 raingauge vs the checkgauge (upper graph, the morning of 12 August 2008) and the greatest undercatch (lower graph, 22/23 January 2009). Rainfall amount is in mm, and time in UTC. The blue line represents the VP2 gauge, and the red line the Didcot gauge. Rainfall values are in mm, and are at 15 minute resolution in the upper chart and 5 min in the lower*
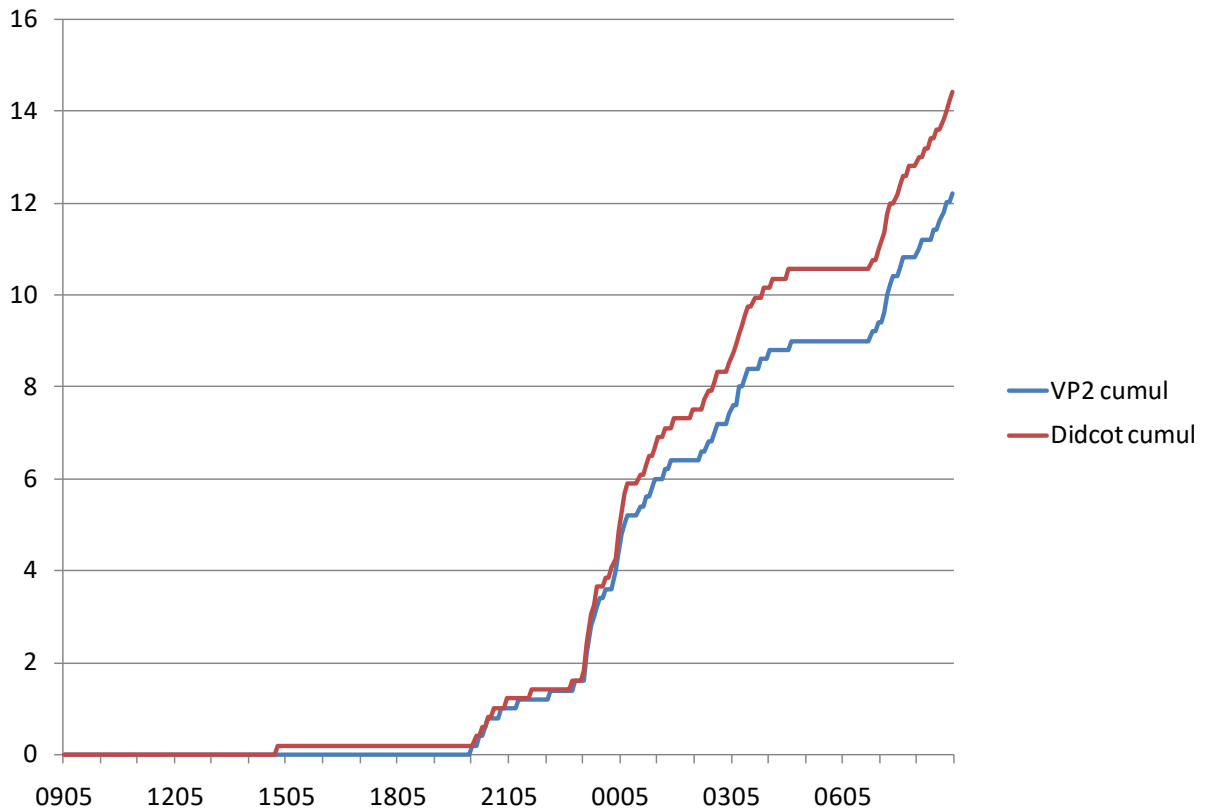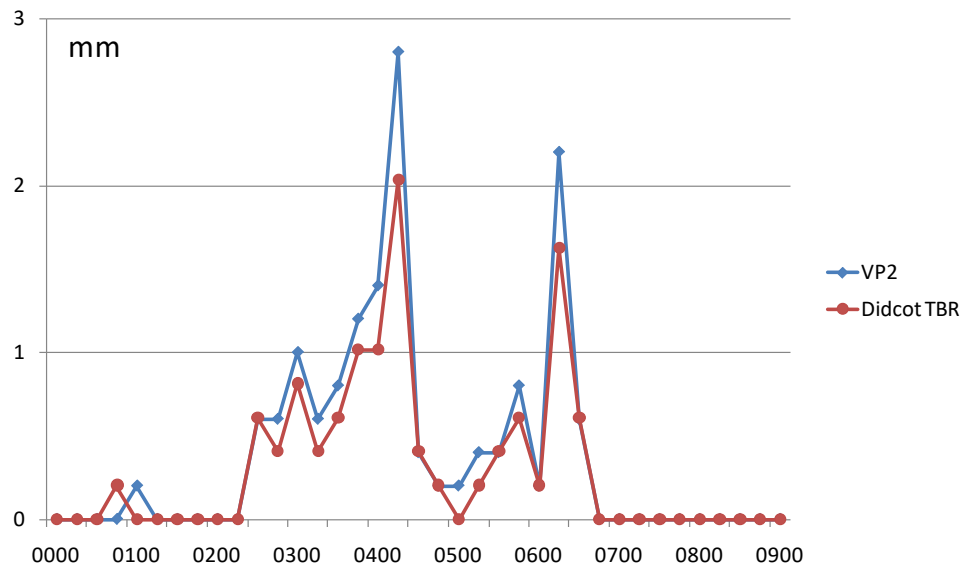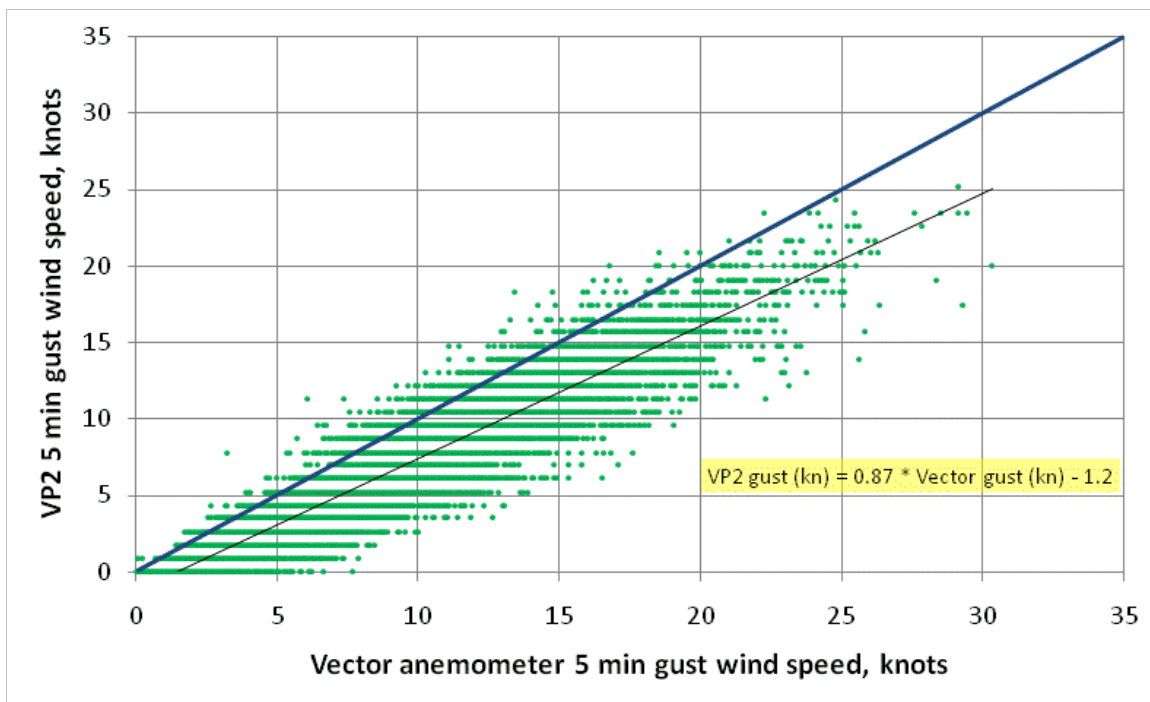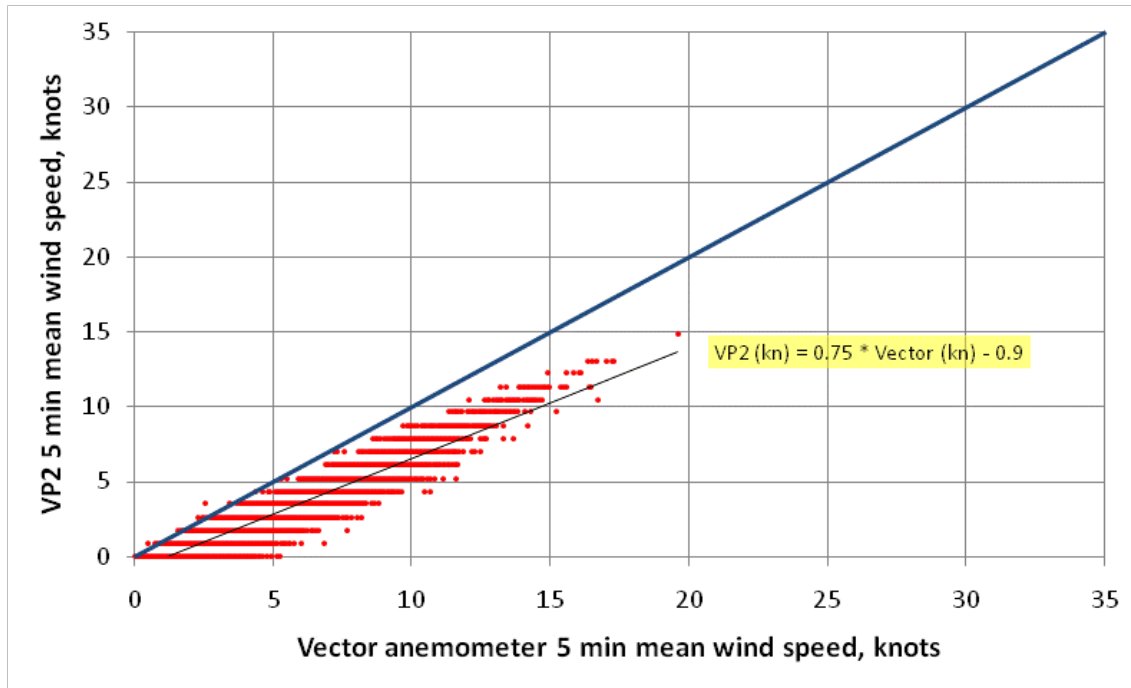
Figure 11 - *scatterplot of simultaneous 5 min scalar mean wind speeds (upper graph, red) and 5 min 'highest gust' wind speeds (lower graph, green) from both instruments over the three months June to August 2009, summarizing in all 26 496 observation pairs. The Davis anemometer's output is given only to 1 mph accuracy (0.9 kn) and thus all Davis data cluster along parallel lines on this plot. The heavy line shows the line of exact correspondence; points below this line indicate that the observed Davis VP2 wind speed (at 2.8 m AGL) was below that of the Vector anemometer (at 11 m AGL). The lighter line shows the derived least-squares linear regression relationship and the inset formula the derived relationship for each parameter*

34