Regional geoid-model-based vertical datums – some Australian perspectives

Research Article

W. E. Featherstone1*, M. S. Filmer1, S. J. Claessens1, M. Kuhn1, C. Hirt1 and J. F. Kirby1

1 Western Australian Centre for Geodesy and The Institute for Geoscience Research, Curtin University of Technology, GPO Box U1987, Perth WA 6845, Australia

Abstract:
This article summarises some considerations surrounding a geoid-model-based vertical datum that have to be thought through before its implementation and adoption. Our examples are based on many Australian and some South-East Asian experiences, but these probably also apply elsewhere. The key considerations comprise data quality and availability, politics, and difficulties that users may encounter when adopting quite a different approach to height determination. We advocate some form of new vertical datum to replace the Australian Height Datum, but the exact type (whether using levelling or geoid, or some combination of both) still needs to be decided. We are not specifically opposed to the adoption of a geoid model as the vertical datum, but it is possibly more challenging than appears initially, and may even deter some users that are already well served by levelling-based vertical datums.

Keywords:
Geoid • vertical datum

© Versita sp. z o.o.

Received 14-10-2012; accepted 22-11-2012

1. Introduction and background

Filmer and Featherstone (2012) present some of the arguments for and against three forms of modernised vertical datum (i.e. levelling-only-based, geoid-only-based, and/or their combination), but this is only from the Australian perspective and driven by the peculiarities of the data available.

In summary, the advantages of a geoid-based vertical datum (VD) for Australia comprise: (i) numerous errors in the existing levelling-based VD, the Australian Height Datum (AHD; Roelse et al. 1975); (ii) avoiding the cost of maintaining and upgrading the levelling network; (iii) direct access to the VD at any point via GNSS (Global Navigation Satellite Systems) without the need to transfer heights from existing benchmarks. The perceived disadvantages include: (i) long-occupation GNSS would sometimes be needed to accurately realise the VD, particularly in areas where existing 3D geodetic control is sparse; (ii) the difficulty and cost of acquiring gravity data in poorly surveyed areas (e.g. the outback, mountains and coastal zones); (iii) a lower relative accuracy of height transfer over short distances in relation to classical differential levelling; (iv) surveyors without GNSS equipment may not be able to access the VD; and (v) the legal traceability of heights determined by a different technique.

The principal obstacles to a revised levelling-based VD in Australia comprise: (i) the tilt between the levelling and the geoid as a result of neglecting the ocean’s time-mean dynamic topography (Featherstone and Filmer 2012); (ii) the quality of the levelling data in relation to the long traverses that do not allow proper identification of gross errors from loop closures (Filmer and Featherstone 2009); (iii) evidence of regional distortions and slopes in the AHD with respect to global and regional gravimetric geoid models (e.g. Featherstone and Stewart 1998, Featherstone 2004, Featherstone and Filmer 2008); and (iv) other problems at the small scale (Kearsley et al., 1988). A readjustment of the basic and supplementary levelling used to establish the AHD in 1971 (Roelse et al. 1975) is possible because these data are available in digital format (cf. Featherstone et al. 2011, Featherstone and Filmer 2012). However, excepting the
correction for gross levelling errors along the north eastern coast (Morgan 1992) and new levelling in the south west (Wellman and Tracey 1987), there are many paper-based records in State/Territory geodetic agencies that are yet to be digitised. These additional levelling data - if made available in digital format - could strengthen the levelling network.

All these difficulties with the existing AHD and any revised levelling-based VD in Australia make the prospect of using a geoid model appear to be an attractive alternative, though - as will be discussed in this article - there are several considerations before implementation and adoption (cf. Véronneau et al. 2006, NGS 2007, Kearsley et al. 1993). For instance, geoid models are refined or revised reasonably frequently over time as new data, theories and computational techniques become available, so the VD has the potential to change more frequently over time. This also applies to a levelling-based VD, but perhaps less so. For instance, the AHD has been in use for over 40 years, whereas there have been four national-standard geoid models released since 1990 and even more global models that can be used on a regional basis.

There is merit to having temporal stability in the VD, as follows. Based on Australian experiences with the change to a geocentric horizontal datum, the majority of the user community encountered difficulties from a 200 m position shift. If a geoid model is adopted to replace the AHD, heights in Australia will only change by around 1 m, so will be more difficult for the uninitiated to decipher. Of course, this could also be the case if a levelling-based VD was readjusted. Regardless, either approach has the potential to fragment the spatial data infrastructure if not adopted fully, where some users will rely on heights connected to the AHD and other users will rely on heights transformed using the geoid model. Thus, the traceability of heights may become confusing, particularly if some users opt to use different geoid models on the incorrect assumption that they all deliver the same result.

We also make the distinction between a "pure" gravimetric geoid model, computed from satellite and terrestrial data alone, versus an "impure" geoid model that has been fitted to regional GNSS-levelling data. The discussions below refer only to the "pure" gravimetric geoid model, based on the assumption that all levelling-based AHD heights will be superseded completely by the "pure" geoid model as the VD.

We share our experiences and opinions in a somewhat cautionary way, only so as to flag some of the problems that could possibly arise from the adoption of a geoid model as the VD. However, many of these considerations apply equally to a revised levelling-based VD, so the arguments should not be seen as a dismissal of a geoid-model-based VD because an updated levelling-based VD will be subject to many of the same restrictions. The style of presentation reflects some pluralism of views (also amongst this author team) and does not intend to provide any definitive answers; instead, it aims to keep the debate open. Also, the following sections are not presented in any order of preference or priority.

## 2. Conceptual Considerations

### 2.1. A new approach to height determination

Traditionally, a levelling-based VD is defined point-wise at benchmarks from a national levelling network connected to mean sea level observed at tide gauge(s) over some epoch (e.g. Vaníček 1991). The user of that VD conducts differential levelling from the benchmarks on datum to transfer vertical geodetic control for a particular project or application. Vaníček et al. (1980) review the uses of differential levelling.

A geoid-based VD changes the approach to height determination entirely, where a GNSS-derived ellipsoidal height (in the absolute sense) or ellipsoidal height difference between endpoints of a GNSS baseline are converted to a physically more meaningful height (or height difference) by subtracting the geoid model. This avoids the need to level from benchmarks, but does require a GNSS baseline observed to a 3D control point with an ellipsoidal height.

Akin to the pure and impure geoid models, we make the distinction between pure and impure ellipsoidal heights (cf. Featherstone 2008). If the 3D control point has been observed with GNSS and processed in the ITRF (International Terrestrial Reference Frame) or some other realisation of a geocentric datum, the ellipsoidal height of that control point is "pure". On the other hand, an "impure" ellipsoidal height is derived from the height on the levelling-based VD plus some quasi/geoid height. Because of errors in both the VD and the quasi/geoid model, this derived "impure" ellipsoidal height will not necessarily agree with the "pure" ellipsoidal height as observed by GNSS.

If no nearby 3D geodetic control is available with pure ellipsoidal heights from which to observe GNSS baselines (say, several tens or hundreds of kilometres), then an extended occupation time is required. This could be up to several days in order to get centimetre-precise ellipsoidal heights (e.g. Ebner and Featherstone 2008). However, the required occupation time will probably decrease as GNSS technology advances. Multiple GNSS constellations (GPS, GLONASS, Galileo, BeiDou, etc.) and additional public-access frequencies will reduce the times needed for accurate ambiguity resolution over long distances.

In areas where there is dense 3D control providing pure ellipsoidal heights (e.g. from a CORS network or 3D control network that has been observed with GNSS), then the user can observe GNSS baselines and avoid the need to use long occupations to determine sufficiently precise ellipsoidal heights in the absolute sense. In Australia, however, GNSS networks are not as dense as in most of Europe, Japan and parts of North America, so long occupations will be needed to establish heights on the geoid-based VD in remote regions.
2.2. Geoid or quasigeoid – which do we want?

In theory, the geoid is by far the preferential reference surface for a VD because of its better description of the true figure of the Earth (cf. Vaníček et al. 2012) and best describes the flow of unrestricted fluids. However, geoid computation need knowledge of the density distribution inside the topography (see Section 3.2). On the other hand, the quasigeoid serves as a pragmatic proxy because it does not require hypotheses about the topographic density, but it is not an equipotential surface and does not properly describe the flow of unrestricted fluids. If there is to be a move to a new VD, then it should seek to reflect the true geoid, rather than any approximation of it. Nevertheless, the choice is ultimately up to the nation that is adopting the VD.

2.3. Why rely more on models than observations?

Levelling-based VDs rely primarily on physical observation: differential levelling and mean sea level measured at tide gauge(s). Admittedly, some models are used to apply corrections to these data, but they are generally much smaller in magnitude than the observations. That is, a levelling-based VD is defined primarily by observations.

If a geoid model is adopted as a new VD, then it is defined less by observations and more by modelling. In short, satellite observations are converted to a model based on spherical harmonic coefficients; gravity and terrain observations are convolution-integrated to generate a model. Since there is not yet any universal agreement on the approaches used to compute a regional geoid model (e.g. Sjöberg 2005), the results are inevitably subject to the analyst’s personal preferences. For instance, different groups can generate quite different results from the same input data (e.g. Valty et al. 2012).

Specifically, observations of differential levelling and mean sea level are reasonably transparent and verifiable. It is quite likely that different analysts will generate the same VD from such observations. However, there are many parameter choices to make when modelling the geoid, so the results can be more variable depending on the analyst’s preferences, which can be less transparent and verifiable.

2.4. How frequently will the geoid-model-based VD be updated?

Geoid models are refined over time and as new or refined techniques and data become available. Notable developments over the past decade comprise: (i) the Gravity Recovery and Climate Experiment (GRACE) twin-satellite mission (e.g. Tapley et al. 2004), which delivers static and time-variable external gravitational field models; (ii) the Gravity field and steady-state Ocean Circulation Explorer (GOCE) satellite gradiometry mission (e.g. Drinkwater et al. 2003), which delivers static external gravitational field models to higher degrees than from satellite tracking alone; and (iii) the Earth Gravitational Model 2008 (EGM2008; Pavlis et al. 2012), which adds terrestrial gravity and terrain data to resolve the gravity field down to spatial scales of around 10 km.

As GRACE and GOCE (and their follow-on missions) deliver more data over time, there will be the justification and temptation to update the geoid model that defines the VD. The question is how frequently should this be done? A balance has to be reached between the “half-life” of the geoid model adopted versus the improvements to be gained by updating it. As stated in the Introduction, users of the geoid model as a VD will prefer more temporal stability, rather than having to contend with frequent updates. Fragmentation of a nation’s spatial data infrastructure will occur if users adopt different geoid models. From Australian experience, some confusion occurred when new geoid models were introduced, where users mixed ellipsoidal heights derived from different geoid models, particularly when there were only a few years between model releases.

2.5. How will we verify future geoid models?

If a geoid model is adopted as a VD, then the levelling network will probably be neglected and – over time – there will be fewer ground-truth data to test geoid models. Admittedly, modern geoid models are now detecting deficiencies in some levelling networks (e.g. Filmer and Featherstone 2009, Featherstone et al. 2011), but there are few other options to test geoid models on land, especially if they are to be used principally for the transformation of GNSS-observed ellipsoidal heights (cf. Vermeer 1998, Featherstone 1998).

One option is synthetic gravity field models (e.g. Baran et al. 2006), but these have not gained widespread application. Another option is geoid validation based on vertical deflections (e.g. Hirt et al. 2007), but not all nations have access to high-precision zenith cameras (e.g. Hirt et al. 2010a) and the spatial coverage is sparse in comparison to current GNSS-levelling. In essence, geoid determination will be less “controlled” and users will not necessarily have a sound appreciation of the veracity of the VD, particularly as different analysts can generate different results (Section 2.3).

On a related matter, there is still no universally accepted means to provide geoid model errors with geographic specificity, apart from those provided by least squares collocation (LSC). However, LSC-derived errors are governed strongly by the covariance functions chosen and the assumption of stationarity, so may not always be representative. On the other hand, the variances provided by a least squares adjustment of a differential levelling network do provide some better indication of the quality of the height provided at each benchmark and closure checks on the observations usually reveal blunders that can be re-observed. A good appreciation of the accuracy and precision of the geoid with geographic specificity is one of the remaining deficiencies with geoid models, where users normally have to accept some blanket estimate based on fits to GNSS-levelling on national or regional scales.
2.6. Choice of the permanent tide system?

The effects of the permanent tides have been given some prior consideration for various height systems (e.g. Ekman 1989, Poutanen et al. 1996, Mäkinen and Ihde 2009). However, the tide system embedded in many geoid models is either neglected or not documented very clearly. If GNSS height determination is to be the primary driver for a regional geoid-based VD, then it is logical to provide it in terms of the tide-free system for the sake of compatibility with the tide system often used in GNSS data processing.

3. Data-driven Considerations

3.1. Reducing the omission error in satellite gravimetry

The external gravitational field sensed from space is inevitably restricted to the long-wavelengths due to the attenuation of signal at orbit altitude and the instrumentation used. This is the omission error, but there is also the commission error in the computed coefficients. Terrestrial (land, marine and airborne) gravity observations have to be added to the satellite-only model to provide the shorter wavelengths (i.e. to reduce the omission error). While the GOCE mission is predicted to deliver 10 mm geoid accuracy over distances of 100 km (e.g. Arabelos and Tscherning 2001), the shorter wavelengths still have to be added from the terrestrial data.

This data combination strategy is still open to some debate (e.g. Sjöberg 2005), but the short wavelengths must be provided if the quasigeoid is to provide precise short- and long-range height differences (cf. Gruber et al. 2012, this issue). However, terrestrial gravity data contain long-wavelength errors (Heck 1990) so filtering is necessary to avoid contamination of the satellite data. Also, some regions are not covered by gravity data (or the data are not made available) so alternative techniques have to be used, such as generating the geoid signal from forward modelling of topographic data (e.g. Hirt et al. 2010b, Pavlis et al. 2012). These issues can seriously degrade the quality of any geoid-based VD, so additional observations may be required.

3.2. Lack of 3D topographic density data

If one is seeking to model the geoid as opposed to the quasigeoid (cf. Section 2.2), then information on the topographic mass-density distribution is required. However, global and regional datasets are currently lacking and some simplifications have to be made. Therefore, a balance has to be reached between the reliability of the topographic density model (and the heights encountered in the region of interest) and the provision of a quasigeoid thus suffering the conceptual difference, which may include water appearing to flow from a lower to a higher elevation. As stated in Section 2.2, we advocate the use of a geoid model over a quasigeoid model because of its clearer description of the true figure of the Earth.

3.3. Lack of detailed gravity data in coastal zones

Given that coastal zones are generally more densely populated, particularly in Australia and South East Asia, and thus more susceptible to rising sea levels and extreme weather events, precise vertical geodetic control is needed in these regions (cf. NGS 2007). Unfortunately, however, there is a dearth of gravity data across the coastal zones due to inaccessibility for ship-borne gravimetry and poorer altimeter-derived gravity anomalies (Hipkin 2000, Andersen and Knudsen 2000).

Some advances have been made in altimeter-derived gravity in the coastal zone because of re-tracking (Sandwell and Smith 2009, Andersen et al., 2010), but there is still a fundamental limit on the proximity to the coastline that the altimeter can track the ocean surface (e.g. Deng et al. 2002). Airborne gravity can contribute in this regard (e.g. Hwang et al. 2006, Featherstone 2010), but few coastlines are yet covered by airborne gravimetry, though the USA has embarked on its airborne GRAV-D project (NGS 2007).

3.4. Datums of the terrestrial gravity and terrain data

Terrestrial gravity and terrain data currently refer to local VDs and horizontal datums that can be inconsistent with the geoid. Of these, the VD of the gravity observations is the most critical because of the vertical gradient of gravity. Some attempts have been made to iteratively compute VD offsets to reduce the gravity data to a common level (e.g. Amos and Featherstone 2009, Claessens et al. 2011) and reformulations of the boundary-value problem (cf. Rummel and Teunissen 1988) have gained renewed interest (e.g. Gerlach and Rummel 2012), and this will probably continue.

4. Political and Implementation Considerations

4.1. Data access across national boundaries

The computation of regional geoid models requires gravity and terrain data beyond national boundaries to avoid edge effects. Particularly in South-East Asia, the exchange of gravity data has historically been poor (cf. Kearsley et al. 1993). Also, countries with a history of conflict will probably never agree to share data. As such, a geoid model as a VD is not so feasible for those countries that are unable to access data from their neighbours. However, EGM2008 has made good in-roads by using confidential or proprietary data and the RTM technique to generate geoid models in these countries (Pavlis et al. 2012). Nevertheless, real observational data will always remain the ideal.

4.2. Effects on land titling and any legislation that involves heights

In some countries, heights related to the national VD are embedded in some land titles and legal legislation. The adoption of a new VD – particularly one that has been established using entirely different principles – may have some adverse effects in this regard. In addition, legal precedents will not have been set. Therefore, as well
as considering the technical issues, it will be essential to consider any legal or other ramifications of adopting a geoid-based VD. The options are too many to list here, but as an example, anecdotal evidence from the implementation of the Geocentric Datum of Australia involved the change of around 400 Acts in Western Australia alone.

4.3. Negative heights on dry land – potential for public confusion

Using the Australian example again, a geoid-based VD may yield negative heights on dry land, particularly near the more populated coastal regions. This is due to the difference between the geoid and mean sea level, most of which is caused by the ocean’s tide-mean dynamic topography and other effects (cf. Featherstone and Filmer 2012). This could cause concern and confusion to the lay public (and maritime navigators) if they find that dry land has a negative elevation, especially given the various concerns over sea level change. A public education program would be necessary, but this can be costly if one wishes to reach the majority of a population. There will also be the need to relate the geoid-model-based VD to maritime VDs such as chart and tide datums (e.g. astronomical low tide).

Admittedly, negative heights already occur on dry land for existing levelling-based VDs (e.g. Lake Eyre in Australia, Death Valley in the United States, or around one-quarter of the Netherlands), but this is because they are below mean sea level, and a levelling-based VD is based on mean sea level (e.g. Vaníček 1991). The emphasis of this sub-section, however, is that change to heights that may occur if a geoid model is used as the zero for heights instead of mean sea level. For instance, coastal communities (and particularly property owners in these regions) would be concerned if the heights changes to become negative. This is because most lay understanding is that if a height is positive, then the land is above mean sea level. One possible solution is to add some constant value to the geoid-model-based VD, but this defeats the object and just maintains the status quo where different VDs will remain offset from one another.

4.4. Management and maintenance of the spatial data infrastructure

As with any change, there will be some resistance with preference for the status quo. As stated in Section 2.4, most users of geodetic datums prefer temporal stability. There will also have to be a transition period, during which it is likely there will be some fragmentation of the spatial data infrastructure, with some heights referred to the levelling-based VD and some referred to the geoid-based VD. Because the change in heights is likely to be small - probably around one metre - there will be a real need for proper documentation to maintain the traceability of heights and to avoid fragmentation. There may also be the need to provide a surface that allows conversion of existing heights to the new heights, in very much the same way that geoid models have been distorted to fit levelling-based VD.

5. Concluding Remarks

First, this article is not meant to debunk the adoption of geoid models as VDs; instead, it attempts to bring to the fore some issues that need to be considered, such as non-technical ramifications on legislation and public perception. Nevertheless, many of the arguments for and against adopting a geoid-based VD apply equally to the adoption of a new or revised levelling-based VD.

From the discussions herein, we recommend that (i) a broad-ranging education programme is used to alert all users to any change and to allay concerns, such as the small change in heights and possibilities of negative heights close to the coasts; and (ii) best practice procedures are developed and promoted for the determination of heights from the new approach. The geoid - rather than the quasigeoid - should be modelled because of its better description of the true figure of the Earth, and treatment of the permanent tide should be the same as used in GNSS data processing.

The geoid model used as the VD should not be updated too frequently (e.g. as soon as a new model is available) so as to provide temporal stability. Levelling databases should not be disregarded totally as they may still have utility for testing geoid models, and can have importance in other scientific studies. Gravity data gaps should be filled, most probably with airborne gravity. 3D digital density models of the topography should also be generated for more precise geoid model computations.

Political issues comprise difficulties of acquiring gravity and terrain data from neighbouring countries, and changes to legislation that involves heights. The latter is potentially very time-consuming because of the number of Acts that can include heights, coupled with the slow process of progressing changes to or formulating new Acts. Implementation issues involve promotion and education programmes to avoid fragmentation of the spatial data infrastructure, especially as the change in heights is likely to be small and thus subject to confusion.

Acknowledgements

We thank the reviewers for their extraordinarily prompt reviews.

References


Andersen O. B., Knudsen P. and Berry P. A. M., 2010, The

Arabelos D. and Tscherning C. C., 2001, Improvements in height datum transfer expected from the GOCE mission, J Geod., 75, 5-6, 308-312, DOI: 10.1007/s001900100187.

Baran I., Kuhn M., Claessens S. J., Featherstone W. E., Holmes S. A. and Vaníček P., 2006, A synthetic Earth gravity model designed specifically for testing regional gravimetric geoid determination algorithms, J Geod., 80, 1, 1-16, DOI: 10.1007/s00190-005-0002-z.


Featherstone W. E., 1998, Do we need a gravimetric geoid or a model of the base of the Australian Height Datum to transform GPS heights?, Austral. Surv., 43, 4, 273-280.


Sjöberg L. E., 2005, A discussion on the approximations made in the practical implementation of the remove-compute-restore technique in regional geoid modelling, J Geod., 78, 11-12, 645-653, DOI: 10.1007/s00190-004-0430-1.


