

WORKING WITH GRID COORDINATES

**Richard J. Sincovec, LSI
Edward-James Surveying, Inc.**

January, 2008

Working with Grid Coordinates

“This project was designed in State Plane...”

If you’ve been doing much surveying lately and you hear those words, your immediate response is very likely, “Oh, no... What are we going to run into *this* time?”

The grid coordinate systems were designed in an attempt to make our lives easier. Surveying with grid coordinates is much like standard Plane Surveying, but without the complicated math of Geodetic Surveying. And of course, GIS professionals like State Plane coordinate systems, because they provide a way to tie together adjacent projects over a relatively large area.

Unfortunately, many people have a hazy grasp of exactly what is going on when we use grid coordinates, and that can lead to some rather exasperating errors and problems. These errors can crop up in the design of the project, when an Engineer forgets that distances on the grid are not the same thing as distances on the ground surface. Or they can crop up during construction staking, when the Surveyor uses scale factors incorrectly. Or the software we rely upon may be designed incorrectly, and it may fail to use State Plane Coordinates in the correct manner.

Luckily, most problems can be dealt with, as long as the people involved have a good understanding of exactly what is happening. This paper attempts to sort through and identify the major issues and points where “things go wrong” with State Plane coordinates. At the very least, it may help people “pick up the pieces” when a project designed in State Plane hits trouble.

But before we can get to the issues that arise from using grid coordinates, we first need to have a firm understanding of exactly what is a grid coordinate system. And before we can understand exactly what a grid coordinate system is, it is helpful to understand why we needed a grid coordinate system in the first place.

Plane Surveying

It is pretty much universally understood at this point that the Earth we live on is a giant spherical object. This basic fact has been well-known among Surveyors since at least the time of the Ancient Greeks, and the shape of the Earth has been defined with ever-greater precision over the course of the centuries. And technically speaking, it isn’t exactly a sphere; it’s elongated in the middle, along the equator, and depressed at the poles, somewhat like a peach or an apple, more-appropriately called a “spheroid” or an “ellipsoid”. However, it is such a *big* object that, for small areas of land, we can ignore this fact, and “pretend” the surface of the Earth is a flat plane. This is called **Plane Surveying**.

With Plane Surveying, we use standard Cartesian (x,y) coordinate pairs to identify the horizontal location of points. For the vertical location of points, we typically use the

distance above a reference plane we call “mean sea level”, which is roughly the average elevation of the world’s oceans. This reference plane is assumed to extend perpendicular to the direction of gravity at our job site.

This “approximation” works very well for small job sites. Normal Euclidian geometry rules apply, so we know that a triangle with sides measuring 3 feet, 4 feet, and 5 feet will also contain a right angle. And as Surveyors, we can use basic trigonometry to do almost any calculation we need.

However, once we start looking at large job sites, or jobs like roadways where one end of the project is quite far from the other end, we begin to run into problems with Plane Surveying. Because of the curvature of the Earth, we begin to run into errors in our calculations, and we can no longer use the simplified calculations of Plane Geometry.

Grid Coordinate Systems

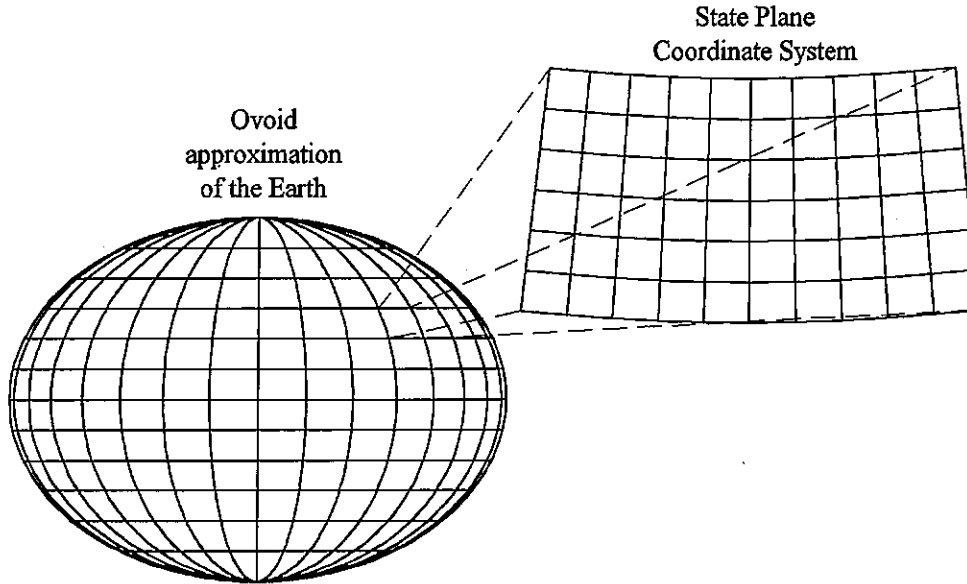
Historically, when it becomes necessary to survey large distances, the principles of Plane Surveying are abandoned, and we enter the realm of **Geodetic Surveying**. Geodetic Surveying is the art of locating and describing objects on or near the surface of spherical objects such as the Earth. In Geodetic Surveying, the distance between two points is measured along an arc, and a rectangle will typically have interior angles that total more than 360 degrees. Surveying is no longer as intuitive, and the computations gain an order of complexity.

Unfortunately, in our modern world, it is getting more and more common to have projects that exceed the limits of simple Plane Surveying. However, it is also undesirable to force all local surveyors to learn and implement the practices of Geodetic Surveying. So in the 1930’s, the United States Coast and Geodetic Survey began developing an “approximation” that came to be known as the **State Plane Coordinate Systems (SPCS)**. Similar systems are now in use throughout the world, although they may use different datum ellipsoids, and the type and orientation of the projection surface may vary. However, all of these grid-based coordinate systems follow the same basic principles.

A grid coordinate system starts with a basic ellipsoid representation of the Earth. The term “spheroid” was used originally, but in recent years, preference has been given to the term “ellipsoid of revolution” instead, often shortened to simply “ellipsoid”. These ellipsoids are regular mathematical figures that are usually defined so that they follow relatively closely to the mean sea level of the Earth. This ellipsoid is then projected onto a curved surface using a conformal projection, usually either a transverse Mercator or a Lambert conformal projection. By using a conformal projection, shapes and angles are preserved, at the expense of throwing most of the error into distances. The error in distances can then be corrected using simple arithmetic.

The State Plane Coordinate System is broken up into a number of smaller areas, or **zones**, and most states contain more than one. The area encompassed by each zone is kept small enough so that the distortion created by the projection is small enough that it was

considered to be within the limits of surveying accuracy, at least in the 1930's. There are a bit more than 120 different zones used to cover the entire United States, which results in a maximum positional error of approximately 1 in 10,000. For the UTM system, another grid-based system with much larger zones, the maximum error is 1 in 2,500.



The boundaries of each zone in the SPCS are generally defined so that they follow municipal lines such as state and county boundaries. This tends to work well for most projects, and in most places. However, if a given project needs to cross the boundary of SPCS zones, there can be additional issues. For this reason, the SPCS is not ideal for work that encompasses truly large swaths of terrain, but it works well for the majority of projects.

When surveying small areas or over short distances – less than five miles or so – these State Plane coordinates can be used with normal Plane Surveying procedures, greatly simplifying calculations. And even when surveying over larger distances, simple arithmetic adjustments can be applied to measurements to account for the error. The procedures used to accomplish these adjustments are outside the scope of this paper, but the net result is that State Plane coordinates remove much of the difficulty involved in surveying moderately-large sites. It achieves the original goal of providing a way to survey large projects without getting bogged down in the details of Geodetic surveying.

Using State Plane coordinates also yields another benefit by providing a “universal” coordinate system for large areas. This is particularly useful for GIS applications such as urban planning, because many disparate projects can be easily linked together, even if the projects are being done by completely different companies that have no communication with each other. There is still a problem for projects that cross over the boundary between zones, or for large metropolitan areas that straddle multiple zones, but it simplifies many things. This benefit accounts for the popularity of State Plane coordinate systems among the various governmental bodies, particularly city planners.

The Datum Ellipsoid

All grid coordinate systems are defined using a particular ellipsoid, and it is very important that you know what ellipsoid your grid system uses as its datum.

When the SPCS was originally created, it used the **North American Datum of 1929 (NAD 29)**, which is based on a spheroid called the **Clarke 1866 Spheroid**. The Clarke 1866 Spheroid is simply a mathematical construct, specifying a geometrical figure that is approximately the same shape as the Earth. In other words, it is defined as a spheroid with a specific radius and flattening coefficient, and that is all that is contained in the definition of the Clarke 1866 Spheroid. We therefore call it simply a **reference ellipsoid**, not a datum. In order to become a **datum**, this reference ellipsoid must be fixed in space, and given a definite location and orientation. The Clarke 1866 Spheroid was tied to a base station in Kansas, along with an azimuth to another nearby base station, creating the NAD 29. This datum was then used as the basis for creating the original SPCS. Because of the way it was defined, the NAD 29 is not what is called a **geocentric** (“Earth-centered”) datum – in fact, the center of the ellipsoid was quite some distance from what is now considered to be the “center of the Earth”. But it served its purpose for a number of years, until the rapidly-improving technology began to illuminate its weaknesses.

In the 1980’s, the National Geodetic Survey switched to the **North American Datum of 1983 (NAD 83)**, which is a datum based on another reference ellipsoid called the **Geodetic Reference System 1980 (GRS 80)**, a slightly-different ellipsoid than the Clarke 1866 Spheroid. The GRS 80 has a slightly-different radius and flattening coefficient than the Clarke 1866 Spheroid. This ellipsoid was then placed so that its center is very near the actual center of the Earth, creating the NAD 83. Most of the SPCS grids were then updated to use the NAD 83. It is critical that when dealing with grid coordinates such as State Plane coordinates, you know which datum is the basis for your grid. Similarly, when working with latitude and longitude, it is also critical that you know which datum is being used. NAD 27 latitude and longitude for a point are not the same as NAD 83 latitude and longitude for the same point, and NAD 27 State Plane coordinates are not the same thing as NAD 83 State Plane coordinates.

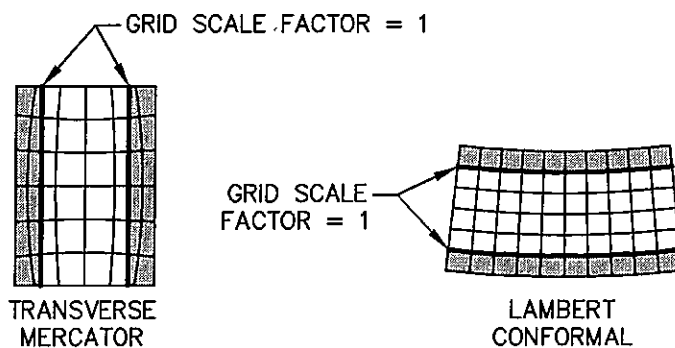
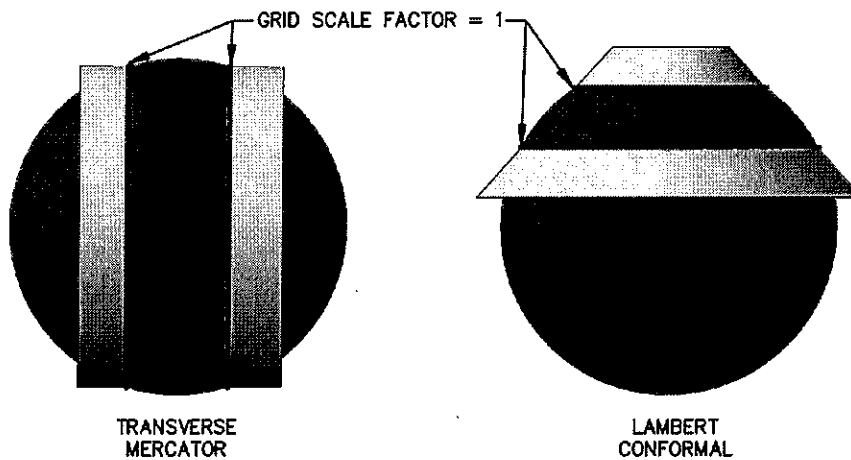
The US Department of Defense has also created another datum, called the **World Geodetic System of 1984 (WGS 84)**, which is typically the datum used by GPS equipment (although modern GPS equipment can typically convert this data to other coordinate systems, as desired). The WGS 84 datum uses a reference ellipsoid that is almost (but not quite) identical to the GRS 80 ellipsoid. In addition, its location is fixed using orbiting satellites, and the methods of determining this location have periodically been refined. As a result, although the WGS 84 and NAD 83 datums were originally intended to be identical, there is some variance between them. In addition, the NAD 83 datum was adjusted using a variety of additional measurements, including conventional traverses, while the WGS 84 was not. The net result is that there can be some variance when inverting between coordinates in NAD 83 versus coordinates in WGS 84. Most of the time, this difference is minimal. But complications may arise if you need to combine NAD 83 and WGS 84 reference points in the same project. When such issues arise, it

becomes necessary to apply the art of Surveying, and use knowledge of the Law, experience, and common sense to arrive at a sensible solution.

The Grid Scale Factor

The terrain of the Earth is very irregular. This complicates any effort to attempt to place a regular grid upon the surface of the Earth. Grid coordinate systems are designed with every effort to minimize the conformal distortion of the projection, at the cost of maximizing the linear distortion. This means that angles and shapes are almost identical between the real world and the projection, at the cost of throwing large error into distances. The amount of error depends on how far our projection surface lies from the ellipsoid surface. The **grid scale factor** is used to correct for this distortion.

The choice of projection method depends upon the shape of the area, or **zone**, to be covered by the system. For a State Plane coordinate system, zones that extend primarily in a north-south direction use a transverse Mercator projection, and the zones that extend primarily in an east-west direction use a Lambert conformal projection. The State Plane coordinate system is a grid on this curved projection surface. Specifying points on this curved surface using State Plane coordinates yields results that are almost identical to using Geodetic coordinates.



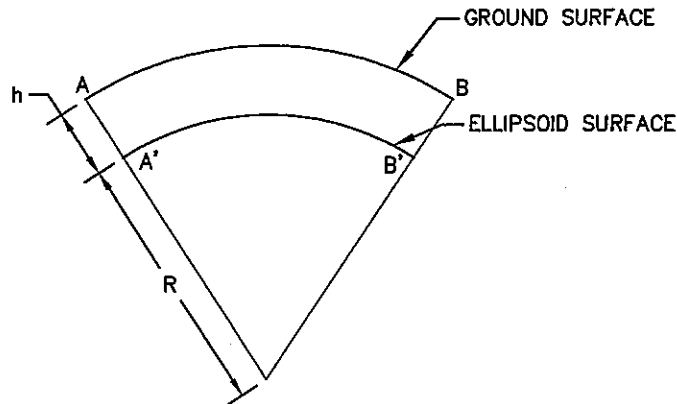
WHITE AREAS: GRID SCALE FACTOR < 1
 GREY AREAS: GRID SCALE FACTOR > 1

The preceding diagram illustrates how a State Plane Coordinate System would be created from a projection. Notice that the projection surface intersects the surface of the ellipsoid. Wherever the projection surface lies inside the surface of the ellipsoid, the grid scale factor for that point is less than 1; whenever the projection surface lies outside the surface of the ellipsoid, the grid scale factor for that point is more than 1. The grid scale factor is exactly 1 at all points where the projection surface intersects the ellipsoid surface. This diagram is fairly typical of State Plane Coordinate Systems, in that the grid scale factor is less than 1 for most of the grid.

The important aspect of this is that the grid scale factor varies depending upon our location in the State Plane. For a Lambert conformal projection, the grid factor varies depending upon the Latitude of the point. For a transverse Mercator projection, the grid factor varies with the distance east or west of the central meridian, as measured on the State Plane ellipsoid.

The Elevation Scale Factor

The Grid Scale Factor helps to compensate for distortion caused by the projection of the ellipsoid onto the State Plane. However, the elevation of the terrain also has an impact, as illustrated in the diagram below.



The ellipsoid used to create our grid is usually placed near mean sea level, e.g. at an orthometric elevation of 0. For most of the US, the actual ground surface is higher than this ellipsoid. In the above diagram, the ground surface is higher than the ellipsoid surface, so the distance between points **A** and **B** is greater than the distance between points **A'** and **B'**. If our ground surface happened to be lower than the ellipsoid surface, then the distance between **A** and **B** would be less than the distance between **A'** and **B'**.

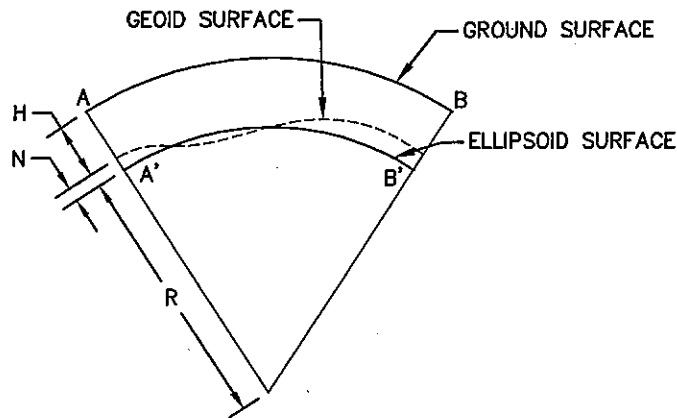
In the above diagram, the value for **R** is the average radius of the ellipsoid over the area encompassed by the SPCS zone. (Typically for the value of **R**, we simply use the average value for the radius of the Earth – the NGS recommends using 20,906,000 feet or 6,372,000 meters – which yields an approximation that is good enough for most purposes.) The value of **h** is the height of our point above this ellipsoid surface. The elevation scale factor r_e for any given point is determined by the following equation:

$$r_e = \frac{R}{R + h} \quad (E-1)$$

As with the grid scale factor, this definition implies that the elevation scale factor is usually less than one. When the elevation of our point is the same as the average ellipsoid elevation for our grid system, the elevation scale factor is exactly 1. When the elevation of our point is below the ellipsoid elevation, the elevation scale factor is greater than 1.

This method of calculating an elevation scale factor is good enough for many purposes. But with modern Surveying techniques, we can actually apply a better elevation scale factor. We still typically use the same values for **R** as before, 20,906,000 feet or 6,372,000 meters. However, we can determine a better value for **h**.

The Earth is an irregular blob of material. Because of this, gravity varies slightly across the surface of the Earth. And since we historically use orthometric levels to determine elevations, and orthometric levels depend on the direction of gravity, this fluctuation in the gravitational field can cause an undulation in the surface we consider to have Elevation = 0. If we imagine the Earth completely flattened to Elevation = 0 everywhere, we get the surface of something we call the **geoid**. Because of fluctuations in gravity, the surface of the geoid is not a regular surface, like the surface of our ellipsoids. In some places, the geoid is above the ellipsoid, and in some places it is below the ellipsoid, and the separation can be as large as 300 feet (100 meters) or more.



In the diagram above, **R** is still the average radius of the ellipsoid over the area in question, and is still usually approximated to 20,906,000 feet or 6,372,000 meters. However, **H** is now the orthometric elevation, and **N** is the distance between the geoid and the ellipsoid, called the **geoidal separation**. Note that **N** may have a negative value. In fact, for the entire continental US, the geoid lies below the NAD 83 and WGS 84 ellipsoids, and **N** is negative. We now have the following equation for our elevation scale factor r_e :

$$r_e = \frac{R}{R + N + H} \quad (E-2)$$

Sometimes an average value will be determined for N for an entire grid zone. Or we can determine a more-accurate value that pertains to our site or project, and use that value for the entire site. This is relatively easy to do if we are using the SPCS, because the SPCS grids typically use the NAD 83 datum, which is almost identical to the WGS 84 datum. Typically, coordinates published for either of these datums will also include the geoidal separation at that point, so the value for N can generally be determined by looking at the data sheets for some of the published control points (such as NGS monuments) in the area.

In some cases, the best choice is to use a “weighted average” for the elevation scale factor. This weighted average is then used as the elevation scale factor for all points in the project. To determine a weighted average, examine the project site, and attempt to discern how much of the site is at each elevation. These numbers do not need to be exact, but they should be generally on target. Then to come up with a weighted average elevation for the site, combine the various areas as follows:

$$elev_{average} = \frac{elev_a area_a + elev_b area_b + \dots}{area_a + area_b + \dots} \quad (E-3)$$

For example, say we had a site with 20% at about an elevation of 5000 feet, 50% at about an elevation of 5200 feet, and 30% at an elevation of about 5500 feet. We could come up with an average site elevation as follows:

$$elev_{average} = \frac{(5000)(20) + (5200)(50) + (5500)(30)}{20 + 50 + 30} = 5250 \text{ feet}$$

Calculating Distances from Grid Coordinates

When we inverse between two grid coordinates, we come up with a distance called the **grid distance** between the two points. If we want to convert this value to a distance on our ellipsoid, called the **geodetic distance**, we must correct for the distortion in our grid projection. Since the grid scale factor varies across the grid, the real-world distance between any two points varies depending on where in the grid the two points lie. Over very small areas, we can ignore the changing grid scale factors. However, as the distance between the two points gets larger, we may need to correct for the changing scale factor.

The approximate geodetic distance between two points can be calculated by taking the grid distance between the two points, then dividing that value by the effective distance scale factor. The effective distance scale factor, r_{eff} , can be determined from the following equation:

$$r_{eff} = \frac{r_a + 4r_{ab} + r_b}{6} \quad (E-4)$$

In equation E-4, r_a is the grid scale factor for one of the points, r_b is the grid scale factor for the other point, and r_{ab} is grid scale factor for the point midway between the two

points. Then we can convert our grid distance d_{grid} to a geodetic distance d_{geo} using the following formula:

$$d_{geo} = \frac{d_{grid}}{r_{eff}} \quad (E-5)$$

Note that this is the *geodetic distance*, i.e., the distance on the ellipsoid, which is roughly the distance at sea level in most cases. If we then need to convert this value to the approximate surface distance, i.e. the value we would physically measure by traveling between the two points on the surface of the Earth, we may need to apply an additional elevation correction. If the entire route is roughly the same elevation, we can simply use a single number for the average elevation of the route, and apply an elevation factor as described in equation E-3.

Using Grid Coordinates

There are two distinct and separate ways of using grid coordinates for a project. Many of the issues that arise with grid coordinates are caused when someone fails to understand this fact, or treats a project designed in one way as if it were designed in the other.

The first method is to use the grid coordinate system in the way it was originally designed. In this method of usage, the grid and elevation scale factors are calculated individually for each point in the project. If the grid coordinates for a point are divided by both the elevation and grid scale factors for that point, we get something called a **surface coordinate** (also called a **ground coordinate**). Over moderate distances or over areas with a consistent slope, we can inverse between the surface coordinates, and we will get the actual ground distance between the points. However, all aspects of the project are performed using the grid coordinates, including the initial survey of existing conditions, the engineering and design, and the subsequent construction stakeout. In this type of project, all distances specified in the plans are *grid* distances, and must be appropriately-scaled in order to determine the actual distance on the surface of the ground. For the rest of this paper, I shall refer to this type of project as a **project that uses grid coordinates**, or as a **grid-based project**.

The second method involves using an average value for the grid and elevation scale factors for the entire project. The average grid scale factor for the project is multiplied by the average elevation scale factor for the project, yielding a **combined scale factor**. The entire project is then tied to the grid system by taking the grid coordinates of a published control point and dividing those coordinates by the combined scale factor. These new coordinates then become the basis for the coordinate system used for our project. The project itself is designed in this new coordinate system, which I shall refer to as **localized coordinates**. For the rest of this paper, I shall refer to this type of project as a **project that uses localized grid coordinates**, or as a **localized project**.

Make sure you do not confuse localized coordinates with surface coordinates. While they are very similar, they are not the same thing. Surface coordinates are generated by applying a floating scale factor to grid coordinates. On the other hand, localized

coordinates involve using what is basically an assumed coordinate system, where a single point in the assumed coordinate system is tied to the grid system using a combined scale factor.

Let's take a detailed look at each of these types of projects. Because they are more-common and more-familiar than grid-based projects, let's start out by looking at the localized projects.

Localized Projects

A localized project uses what is essentially an assumed coordinate system. This assumed coordinate system is tied to the grid coordinate system by a base point and a combined scale factor, creating our localized coordinate system. All survey tasks are performed "normally" in the localized coordinate system, with no scale factors anywhere. All engineering is also performed in the localized coordinate system, and all distances on the plans are surface distances. If it says 500 feet in the plans, it's 500 feet on the ground.

With this system, we can take any localized coordinate in our project, and multiply by the combined scale factor to come up with something that is relatively close to the grid coordinate for the point. There is some error, and the error gets larger the farther we get from our project base point. However, as long as the project does not cover a very large area, and the project is roughly constrained to the same elevation, this error tends to be very small.

And I'll repeat this point, because it is so important: **when we create a project in a localized coordinate system, we do all work in the localized coordinate system.** For this type of project, the grid system really only comes into play when we go to combine different projects in the same general region, i.e. when the GIS professionals go to use it. We'll return to this point later. But first, let's look at how we should work in a localized project.

Tying the Grid to the Localized Coordinate System

To create a localized project, we must first define a base point and a combined scale factor for the project. The goal here is to determine values such that when we multiply our localized coordinates by the combined scale factor, we get points that are roughly in the correct place on our grid coordinate system.

If there is already a published control point near the center of our job site, then we are in good shape. We can probably use that control point as the base point for our project. Otherwise, we can pick any point near the center of our project to use as the base point, and use some other method for determining the grid coordinates for our base point. If we are using GPS equipment, we probably can simply get the grid coordinates directly from the equipment. Otherwise, we may need to traverse to the point from known control points, and assign grid coordinates to our control point in that fashion.

Once we have a base point, we then need to determine the grid scale factor. If we are using a published control point, we can probably just use the published grid scale factor

for that point, assuming it is near the center of our site. Or, if we have determined the grid coordinates for one of our control points near the center of the site, and we have a software program that will let us key in the coordinates and will calculate the grid scale factor for us, we can use that. Otherwise, we may be able to determine the grid scale factor by averaging several control points around our site. And of course, we can always do it the "hard" way, using the tables that are designed for this purpose and are published along with our grid system.

Once we have determined the grid scale factor at our base point, we should also come up an average elevation scale factor for our project, using the technique described in Equation E-3. We then multiply the grid scale factor by the elevation scale factor to determine the combined scale factor for this project.

The combined scale factor may then be used to scale the entire project down to grid, so it can be placed on a map with other projects in the same area. When we do this, the edges of the projects will not tie together precisely on the grid coordinate system, because of the way each project is scaled by a different combined scale factor, and using a different base point. But if we stake out a point from the control for Project 1 and using the localized coordinates for Project 1, and stake out the same point from Project 2 and using localized coordinates for Project 2, the two points should hit exactly the same place on the ground.

Using Project Coordinates

Frequently in a localized project, we do not actually use the localized coordinates. For several reasons, the coordinates in a grid system tend to be very large, and since localized coordinates are tied to the grid coordinates, the localized coordinates tend to be very large, as well. Also, if our combined scale factor is close to 1.0000, it can become difficult to tell the difference between grid and localized coordinates.

For this reason, it is common to truncate some of the leading digits from the localized coordinates. These "truncated localized coordinates" are then used for the project, and are called **project coordinates**.

And really, it is not a good idea to actually "truncate" the digits. It's better to use a horizontal transformation to translate the coordinates between localized and project coordinates. As an example, say we have a project where all northings are in the 3,510,000 range. So, since every Northing starts with the digits "351", we might decide to truncate the first three digits from the northing. If our localized coordinate is 3,512,168.50, then our project coordinate would be 2,168.50. But now something unexpected comes up, and we need to go tie down a point a fair distance south of our project. The northing for this new point is 3,509,895.55. If we truncate the first three digits from that coordinate, we get 9,895.55, which would be way, way to the north of our project, and not to the south of our project. On the other hand, if we apply a horizontal transformation, we end up with a project northing of -104.45. Negative coordinates aren't very palatable, but these days, I believe all survey equipment can handle them with no trouble. So, if we apply a coordinate transformation, we get results that work. If we simply truncate digits, we may get a mess.

It is also possible to use a completely independent assumed coordinate system for our Project Coordinates. In this case, we can come up with a horizontal translation and a rotation to convert between our project coordinates and our localized coordinates. Conversion to grid would then use the localized coordinates. In other words, it becomes a two-step process to get from project coordinates to grid coordinates – first, convert the project coordinates to localized coordinates using the horizontal translation and rotation, then convert the localized coordinates to grid coordinates using the combined scale factor.

Surveying a Localized Project

Surveying in a localized project is very straight-forward. There are no scale factors, and the localized coordinates (or, more often, the project coordinates) are simply used “normally”. All procedures of basic Plane Surveying hold, and everything is very simple. All distances in the plans are surface distances, distances in all legal documents tend to be surface distances, and all field measurements are surface distances, so there is no confusion there. All in all, it is the best way to survey a project on a small site, because there is comparatively little room for error or confusion.

Engineering a Localized Project

Engineering in a localized project is also very straight-forward. Again, there are no scale factors, and inverting between features in the existing ground survey yields the precise ground distance between the points. All distances in the engineering plans are ground distances, and any coordinates should be localized (or project) coordinates. Again, it’s basically “business as usual”, with no complications, so this is the best way to engineer a project on a small site.

GIS and Localized Projects

In order for a GIS professional to use a localized project, the entire project must be scaled to grid using a combined scale factor. If the project is using project coordinates, the project must be translated/rotated to the localized coordinates first, then scaled down to grid. Since GIS professionals are typically only interested in “the big picture”, the fact that this scaling introduces slight errors is irrelevant, as is the fact that the lengths of lines no longer match the distances on the construction plans. In order for the GIS professional to be able to accomplish this task, the plans must include the combined scale factor, as well as any transformation information necessary for getting from project coordinates to localized coordinates.

Grid-Based Projects

Grid-based projects are a bit strange at first, because of the fact that this type of project has a **floating scale factor**. In other words, every point in the project gets its elevation and grid scale factors determined separately from every other point in the project. It can take a little while to get a grasp of this concept.

But it goes back to the fact that we basically want to use Plane Surveying techniques, but our job site is large enough that the curvature of the Earth impacts the project. Since the scale factors change for each point in the project, the curvature of the Earth is basically compensated for in the scale factors. When using the grid system for design, we can largely ignore the fact that the surface of the Earth is really curved, and design our project normally. Then, during construction stakeout, the floating scale factor gets applied to stakeout points, and compensates automatically for the shape of the Earth.

This has some important implications, though. Since we have a floating scale factor, things become very irregular if we try to design in surface coordinates – in fact, it is pretty much impossible. So for this type of project, **the project must be designed in the grid coordinate system**. The design is then scaled up to surface coordinates during stakeout.

When designing in a grid system, it is vital to remember that everything will be scaled during construction. That means that if it is critical that a certain distance measure 500 feet exactly on the ground, it may need to measure something else (probably less than 500 feet) in the plans. Also, for something like a very long roadway or tunnel, the surface distance between 100-foot stations will vary across the project, and will probably never be exactly 100 feet. This is something that often throws construction crews (and especially inspectors) into fits if they happen to notice it, and it can even have the same effect on Surveyors if they do not understand what is happening and why. We will talk more about these points later. First, let's take a closer look at the inner workings of this type of project, starting with how the data is collected.

Surveying a Grid-Based Project

When working with grid-based projects, we must work with grid coordinates. The grid coordinates are the only thing that is regular. Since our surface coordinates are determined by applying a floating scale factor to each point in our project, the surface coordinates are “fuzzy” and cannot be used for survey calculations. Instead, we basically survey “on the ellipsoid”, and convert all our survey observations from surface observations to ellipsoid (grid) observations before performing any calculations. But since our grid system uses a conformal projection, we can use our measured angles with no adjustment; we only need to adjust measured distances, converting them from surface distances to grid distances. Similarly, during stakeout, all calculations are made “on the ellipsoid”, with grid coordinates and distances, and then the distances are scaled up from grid to surface during stakeout. Luckily, most of these calculations occur automatically in our data collector, and the process becomes largely transparent to the surveyor. It is still critical to understand what is happening, though.

A lot of people get confused by the floating scale factor, and think of it as a way to get back and forth between surface and grid coordinates. That is incorrect, however. It is a one-way thing, where we can convert grid coordinates to surface coordinates by dividing by the floating scale factor. However, we typically do not do the reverse, and use the floating scale factor to convert ground to grid coordinates.

Instead, we use grid coordinates for our control points. Then, when we use conventional equipment to measure an angle and shoot a distance to an unknown point, we reduce the measured horizontal distance to a grid distance, and then use the grid distance to determine a grid coordinate for the unknown point.

In most cases, it is perfectly acceptable to just use the occupied point for determining the floating scale factor for this calculation. If the unknown point is a large distance away from the occupied point, or is at a dramatically different elevation, then this may or may not be precise enough. If it is not precise enough, then we can use the grid and elevation scale factors for our occupied control point to come up with a “guess” at the unknown coordinate. We can now use our “guess” as the coordinates for the unknown point, and determine approximate grid and elevation scale factors for the unknown point. Then, by applying Equation E-4 to determine a new effective scale factor for the measurement, we can repeat our calculations with the revised scale factor. The result is the grid coordinate for our unknown point.

If we are using GPS equipment, then things are a bit simpler. The GPS equipment resolves the measurements from the satellites into a latitude and longitude, which it then converts directly to our grid coordinates.

When performing the design survey of existing ground for a grid-based project, the surveyor must be sure to configure his equipment properly for performing these calculations, so that all control points and all coordinates for collected data use the grid coordinates. Other than that, surveying in a grid-based project is much like surveying a localized project.

There are two main drawbacks to surveying in this manner. The first is that all legal documents, including property descriptions, are typically written using the surface distances. It can therefore be very tricky doing boundary work in grid coordinates. The second problem is that inverting between plan coordinates yields a grid distance, and a pulling out a handtape or using an EDM to measure a distance at the job site will yield a surface distance. This complicates the process of checking our work. It is relatively simple if we are using our equipment and known coordinates, because our equipment will give us a delta – i.e., it will tell us that we are 0.04’ off of our calculated point, or whatever. But if we are not using our equipment, we need to remember that we will not measure exactly the same distance as what is called out in the plans. Similarly, the distances called out in the plans are all grid distances, and the actual distance we measure on the ground will not be exactly the same. Usually, the difference is not significant for construction, but there are exceptions. We’ll talk about some of those exceptions in the last portion of this paper.

Engineering a Grid-Based Project

In most respects, the workflow involved in engineering a grid-based project is identical to that in a localized project. The only significant difference is that for a grid-based project, all project design is performed in grid coordinates, and all distances on the plans are grid distances.

Again, the only significant difference between a localized project and a grid-based project is that distances on the surface will not be the same as the distances in the plans. So if it is critical that a certain feature measure 500 feet on the ground when the project is built, it will probably need to be designed somewhat smaller in the plans.

This is not usually a problem if everyone involved understands what is going on with a grid-based project. But it can really confuse people who don't understand the process.

GIS and Grid-Based Projects

In the minds of the GIS professional, a grid-based project is ideal. The GIS professional can simply take the linework for the project as it is, with no adjustments, and it should fit together extremely well with the surrounding projects, assuming they were also done in the same grid system.

What Can Go Wrong?

Now that we've learned what a grid coordinate system is and how it is used, let's take a look at some of the issues that can arise from a misapplication of this technology.

Surface Distances Used in Grid-Based Projects

This is the most-common error that arises when using grid-based projects. We've touched on this already in this paper. The problems occur when an engineer is designing a feature with a critical surface length, and the engineer uses that length in the plans, forgetting that all distances in the plans must be grid distances.

As an example, let's take a recent prefabricated building we worked on. This building was rather large, approximately 500 feet long. The building itself was composed of very large, prefabricated metal panels. There is some give in these panels, and some room for error, but not a lot.

This project was designed "in State Plane". The engineer even provided coordinates for each corner of the building in the plans. However, we are located in the middle of Colorado, at nearly 7000 feet above sea level, and in the middle of our State Plane zone, where the grid scale factor is the smallest. The net result was, if we had staked out the coordinates the engineer gave us, we would have staked a building that was nearly 500.3 feet long – more than three inches too long.

Three inches isn't a lot, but it's enough that it could've created complications for the guys erecting the building, because the modular pieces are designed to be exactly 500 feet long once constructed. And if the construction guys measured the distance themselves, they would find something that is obviously quite-a-bit longer than the 500 feet they were expecting. If we were to let this happen, then the construction crew would likely start to question our competency as Surveyors, since we lay out a simple building with so much slop in it.

In this particular case, it was quite simple for us to catch the error, realize that the engineer actually wanted a building that measured exactly 500 feet, and recalculate the coordinates appropriately. But this is the sort of problem that should not arise if the engineer is cognizant of what it means to work in a grid-based project.

Grid Distances Used in the Field

This is a similar error to the last one, but one that can catch the field surveyors.

As an example of this one, let's take an airport runway we worked on recently. This project was designed using the State Plane grid. The runway itself was approximately 2½ miles of concrete, laid out in 20'x20' panels.

The interesting thing about this job is that it was an FAA job. The FAA has created a set of tolerances that, frankly, are highly-questionable. All features must be within 0.04' of their plan location, horizontally, including the joint lines in the concrete. This is a rather ridiculous tolerance for joint lines, especially when there is 21" of concrete laid on top of 8" of asphalt laid on top of 12" of concrete-treated base course. But this is the tolerance that the FAA has given, and it is the tolerance that we were required to meet.

Now let's take a look at these joint lines. Since their horizontal location had so little room for error, the contractor naturally wanted the surveyor to mark the pavement where the sawcut lines should go. And naturally, since we were marking multiple paving lanes, each of which was 2½ miles long, and the paving was being performed at break-neck speeds to keep up with the strict schedule, we sought to make this task go as quickly as possible.

We had some 300-foot chains, so we decided the best course of action would be to use our instruments to mark periodic stations, say one every 600 feet. Then we could use the chain to mark out 20 foot panels, measuring 300 feet each way from our "good" points.

However, once again, we are at high altitude, and in the middle of a State Plane zone where the grid scale factor is the smallest, and we had a combined scale factor for this project of approximately 0.9995. This means that we needed to measure 20.01 feet for each panel. Had we failed to correct for this variation in distance and marked out panels that measured exactly 20 feet, we would have been 0.05' off by the time we went only 100 feet. We would already be outside our specified tolerances for the project, after only 100 feet! And with the extremely-strict inspectors that work for the FAA, that would have created a huge hassle for us. At best, it would involve spending lots of additional time fighting red tape, and essentially begging for special exceptions for out-of-tolerance areas. At worst, it can mean tearing out and replacing entire sections of completed (and expensive) runway.

So, as we can see, while it may *seem* like the difference between a grid distance and a surface distance may not seem like a whole lot, it can become critical, especially when a project has narrow tolerances and inflexible inspectors.

Grid Distances used to Pre-Fabricate Structures

This issue is related to the previous two. It can arise when the engineer is fully-cognizant of the fact that a project is being designed on the grid, but being built on the ground. That 500-foot prefabricated building should be exactly 500 feet long in the architectural and structural plans, which are done using the actual distances that the building should measure once it is built (i.e., surface distances). However, assuming that 0.9995 combined scale factor mentioned earlier, distances in the civil site plan are *grid* distances, so that building gets labeled as being 499.75' long in the site plans. The manufacturer of the building panels should be sure to use the *ground* measurements when prefabricating the building components, and not the grid distances. And of course, everyone involved in the project should understand why the building is 500.00 feet long on the structural plans, but only 499.75 feet long on the civil site plan.

Mixing-Up Grid and Project Coordinates

I have learned that some people like to adjust their project coordinates so that they are as close as possible to grid coordinates. This is a variation of using localized or project coordinates. However, it is done in a very specific way. So to differentiate this method of using grid coordinates from the others, I will call this method a **stretched project**, using **stretched coordinates**.

With this method, a localized coordinate system is created as detailed earlier in this paper. We then use a "base point" to determine a combined scale factor for the project, then divide the grid coordinates for our "base point" by the combined scale factor to get to the localized coordinates. But then, to get to our *stretched* coordinates, we transform the localized coordinates so that the project coordinates for our base point are identical to the grid coordinates for the base point.

The following table illustrates the difference between these three. The base point for the project in this example is at N 1,500,000 and E 3,500,000, and our combined scale factor is 0.9996. Therefore, our localized coordinates are equal to our grid coordinates divided by 0.9996. To get to the stretched coordinates, we then subtract 600.24 from the northing and 1,400.56 from the easting, so that the coordinates for our base point are the same as the grid coordinates. The following table illustrates grid coordinates, localized coordinates, and stretched coordinates for several points in the project:

Grid Coordinates		Localized Coordinates		Stretched Coordinates	
Northing	Easting	Northing	Easting	Northing	Easting
1,500,000.00	3,500,000.00	1,500,600.24	3,501,400.56	1,500,000.00	3,500,000.00
1,500,500.00	3,500,500.00	1,501,100.44	3,501,900.76	1,500,500.20	3,500,500.20
1,499,200.00	3,501,100.00	1,499,799.92	3,502,501.00	1,499,199.68	3,501,100.44
1,502,500.00	3,497,500.00	1,503,101.24	3,498,899.56	1,502,501.00	3,497,499.00
1,495,000.00	3,495,000.00	1,495,598.24	3,496,398.56	1,494,998.00	3,494,998.00

This actually makes some measure of sense. The net result is that, for small projects, the project coordinates and grid coordinates are almost identical. The amount of error gradually increases as we get further from our base point. With a combined scale factor of 0.9996, we would have about 0.04' of difference between our grid and project coordinates for every 100 feet we get away from our base point. This error affects both the northing and easting, so a point that is 100' south and 100' east of our base point would have 0.04' of error in the northing and 0.04' of error in the easting, for a total of about .06' of error. For a small construction project that only measures a couple hundred feet across and has normal tolerances, 0.06' of error is probably insignificant. The net result is, with this system, it is possible for someone to have absolutely no knowledge of how to use a grid system and still get the job done. The surveyor can even completely mix up grid and project coordinates, and use a combination of both, and may still manage to get the project built without serious issues.

However, this is a relatively "sloppy" method of proceeding. It is far more desirable that the surveyor understands the difference between grid and project coordinates, and uses them correctly. And assuming the surveyor understands the difference, it is far better to use a system that makes it clear which set of coordinate is which, and makes it impossible to "mix up" grid and project coordinates. For this reason, we strongly recommend that stretched coordinates are avoided. Instead, it is far better to use the so-called "truncated" localized coordinates as project coordinates, as described earlier in this paper.

Designing "On the Grid", but Using a Combined Scale Factor

This is another extremely-common engineering error. This comes from a desire to "use grid coordinates", but without fully-understanding them. Usually, the driving factor behind this error is the fact that a review agency wants plans submitted electronically for GIS purposes, which means they want them on the grid. So the engineer just does everything "on the grid".

However, if the project does not require the use of the grid, this is adding unnecessary complications to the project. As we have seen, there are lots of complications with designing a grid-based project. Distances in grid-based projects are grid distances. However, all legal documents, such as deeds and plats, typically use surface distances. This means the property lines for a grid-based project will not measure the same distance as what is called out in their legal documentation, which is another source of confusion for those who do not understand grid-based projects. And of course, we can run into the exact same problems mentioned earlier, where the engineer "forgets" that all distances in the plans are grid distances, and designs that 500-foot building using 500 feet on the grid instead of 500 feet on the surface.

Instead of doing this, if a project site is small enough that it can be designed using project coordinates, then it *should* be designed using project coordinates. This way, many possible sources of confusion are completely avoided. And it is a relatively simple operation to scale the linework down to the grid before sending it on to the GIS professionals. There are some in the industry who advocate using grid coordinates for everything. But I think this paper may have illustrated that, while useful and much

simpler than Geodetic Surveying, a grid-based project still creates lots of possibilities for confusion and error. If we do not *need* to use a grid-based system because of the size or extents of our project site, then there is no good reason to introduce those complications, and localized or project coordinates work better.

“Let Down” by our Software

This one hurts, because it is largely out of our control. If we have a problem with our software, we must rely on our software vendors to fix it, and in the meantime, we can only make the best of what we have.

Historically, grid coordinate systems are something that have not been widely-used. Until relatively recently, grid coordinate systems tended to be used only in long highway and roadway jobs (e.g. the Department of Transportation projects), or by the military. But now that projects are getting larger and larger, and GPS is becoming more and more prevalent, and GIS is becoming more and more important, the grid systems are gradually permeating everything we do. It is quickly becoming imperative that all professionals in the Civil industry fully-understand the topics covered in this paper.

Unfortunately, the designers of our software are in a different profession (software engineering), usually with no direct knowledge of or experience in the Civil industry. They rely wholly upon Civil professionals to tell them what we need our software to do. And since so many Civil professionals have a hazy or incorrect grasp of grid coordinate systems, a fair bit can get “lost in translation”, and our software is often designed in ways that do not really support the tasks we need to support. For example, our software may not fully-support the translation from grid to localized to project coordinates, with the ability to create labels using any of the three coordinate systems. It may not allow us to choose whether we want to design in grid coordinates or project coordinates, and lock us into designing in either one or the other. It may lack the ability to measure or label any element in a grid-based project with any combination of grid, surface, and geodetic distances. It may even lack the ability to correctly determine the floating scale factor for points in a grid-based project, or it may perform survey reductions improperly, or it may suffer any of a number of other issues.

A key problem that strikes much of our software is the lack of “project support”. For example, we may be able to configure all options correctly in our data collector (e.g., are we using Grid or Project coordinates; if we are using Grid coordinates, what is our coordinate zone; if we are using Project coordinates, what is the combined scale factor; etc.). But if these configurations are stored in the same job file as our survey observations, those options must be configured properly every time we create a new job file for that project. Any error in these settings can have terrible repercussions. So, our data collectors should have the concept of a “Project”, which we can create once, containing all settings for our project. We should then be able to create individual job files each day, so that we can daily dump all collected data from our data collectors and start a new data collection file, but without re-entering all the project options. This is a feature that is not regularly supported in current data collection software.

Of course, we cannot expect the software engineers to get our software right if we in the Civil industry do not understand what we want. As long as Civil professionals are abusing and misunderstanding and mis-implementing grid coordinates, we cannot really expect software designers to get our software to work well. So hopefully, this paper will help clarify the issues in the Civil industry, so we start giving the right feedback to our software providers. And hopefully, this paper will also clarify the issues for the software experts, so that they can understand what we are asking for.

References

- Moffitt, F.H., and Bouchard, H. (1975). *Surveying (6th Ed.)*. Harper & Row
- James P. Reilly, Ph.D. (2000). State Plane Coordinates vs. Surface Coordinates, Parts 1-9. *Point of Beginning*. Retrieved from <http://www.pobonline.com>.
- National Geodetic Survey web page and FAQ. Retrieved from <http://www.ngs.noaa.gov/faq.shtml#Datums>.
- Schwarz, C. R., Editor. (1989). Relation of NAD 83 to WGS 84. *NOAA Professional Paper NOS 2, Article 22*. Retrieved from <http://www.mentorsoftwareinc.com/resource/Nad83.htm>.
- The State Plane Coordinate System. Warner College of Natural Resources, Colorado State University. Retrieved from http://welcome.warnercnr.colostate.edu/class_info/nr502/lg3/datums_coordinates/spcs.html.